

Reporting on quality of life in randomised controlled trials in gastrointestinal surgery

Valérie Bridoux^{1 2}, Grégoire Moutel², Benoit Lefebure¹, Michel Scotte¹, Francis Michot¹, Christian Hervé², Jean-Jacques Tuech^{1 2 *}

¹ Service de chirurgie générale et digestive CHU Rouen, 1 rue Germont, 76031 Rouen Cedex, FR

² LEM, Laboratoire d'éthique médicale et médecine légale Réseau Inserm de Recherche en éthique médicale, Université Paris Descartes, Faculté de médecine, 45 rue des Saints-Pères, Paris 75006, FR

* Correspondence should be addressed to: Jean-Jacques Tuech <jean-jacques.tuech@chu-rouen.fr>

Abstract

Background

Although health-related Quality of life (HRQOL) has become an important outcome measure in surgical trials, questions still remain about the quality of its reporting. The aim of this study was to evaluate HRQOL assessment methodology of randomized clinical trials concerning gastrointestinal surgery.

Methods

All articles published in the calendar years 2006 and 2007 that purported to assess quality of life as end points or make some conclusion about quality of life were chosen for review from eight general surgical journals and four medical journals. Identified eligible studies were selected and then evaluated on a broad set of predetermined criteria.

Results

Twenty-four published RCTs with an HRQOL component were identified. Although most trials exhibited good-quality research, some methodological limitations were identified: only 21% of the studies gave a rationale for selecting a specific HRQOL measure, 46 % of the studies failed to report information about the administration of the HRQOL measure, and 37% did not give details on missing data.

Conclusions

Although it is clear that HRQOL is an important end point in surgical RCTs because the information helps to influence treatment recommendations, a number of methodological shortcomings have to be further addressed in future studies.

Author Keywords quality of life ; randomized controlled trial ; surgery

MESH Keywords Digestive System Surgical Procedures ; Health Status ; Humans ; Quality of Life ; Randomized Controlled Trials as Topic ; standards

INTRODUCTION

This century we have witnessed significant progress in the diagnosis and treatment of disease. The effects of disease and its treatment on patients have traditionally been assessed in terms of pain scores, duration of hospital stay, and return to normal activities. These outcomes, however, are dependent much on external factors, such as on local habits and social security matters. Thus, the application of quality-of-life instruments, which measure recovery in a patient-centered manner, has become more popular in recent times and has been accepted more and more as a solid primary outcome measure in scientific studies [1].

At the present time, there is no single definition of HRQOL. Nevertheless, there is a broad consensus that it refers to the physical, psychologic, and social functioning of patients and the impact of disease and treatment on their abilities and daily functioning [2–5].

There are several valid measures of HRQOL that are suitable to use in surgical research. Generic measures (such as the short form health survey SF-36 [6]) broadly assess physical, mental, and social health and can be used to compare conditions and treatments. Measures specific to illnesses (such as the Gastro-Intestinal Quality of Life Index GIQLI [7]) can supplement generic measures or can be used independently [8].

Although these instruments are widely available, careful application of the tools in clinical studies is needed to produce reliable and clinically useful results. These range from the accurate selection of the most appropriate instrument for the particular trial objective, to the handling of missing data and accurate interpretation of outcomes [9–11]. Unless standards for measuring HRQOL are adhered to in clinical trials, the data that are collected will be difficult to interpret and unlikely to make clinical sense [11].

Previous review on randomised controlled clinical trials (RCTs) including an HRQOL evaluation in oncology have shown overall a number of methodological shortcomings [12–19].

However, to date, no detailed systematic methodological review of the quality of the conduct and reporting of QOL results from RCTs for gastrointestinal surgery has appeared. Therefore, the aim of this study was to evaluate the quality of HRQOL methodologic assessment in randomized controlled clinical trials (RCTs) involving the gastrointestinal surgery, and determine how improvements can be made.

Because they are considered the optimal study design for evaluating the effects of different surgical interventions [20], we limited our search to « randomized controlled trials » and to recent articles published between January 2006 and December 2007.

METHODS

Search strategy for identification of studies

Twelve journals were chosen for review: eight English-language surgical journals (American Journal of Surgery, Annals of Surgery, Archives of Surgery, Journal of the American College of Surgeons, Surgery; British Journal of Surgery, European Journal of Surgical Oncology) and four English-language medical journals (New England Journal of Medicine, Lancet, British Medical Journal, and Journal of the American Medical Association).

To identify eligible articles, all issues of these journals were hand-searched.

Studies included for review had to be randomised-controlled gastro-intestinal surgical trials, phases III published between 01/01/2006 and 31/12/2007.

All randomised-controlled trials (RCTs) comparing different treatment were eligible, regardless of the intervention type. No restrictions were performed on trial location, number of patients enrolled in the trial, treatment modalities and sponsor of trial.

The exclusion criteria were: (a) trials published as a letter, abstract, or short article; (b) randomised phase II trials; and (c) non-experimental (observational) studies.

The search was restricted to RCTs as they represent the gold standard by which health care professionals make decisions about treatment effectiveness [20].

Characteristics assessed

Two reviewers (V.B. and J.J.T.), who were not involved in any of the identified studies, analysed the identified RCTs independently. Any disagreement was resolved through discussion between the two reviewers.

As Quality of Life (QOL) was the main outcome measure sought, any studies including assessing quality of life as an end point or making some conclusion about quality of life were considered.

The standardised protocol was based on a checklist (available from the authors). The items to be included were: country of origin, industry funded (yes versus no), number of patients randomised, multicenter studies (yes versus no), informed consent reported (yes versus no), approval of a research ethics committee reported (yes versus no) and Health Related Quality of Life (HRQOL) difference between treatment arms (yes versus no). The latter was defined as any statistical difference between treatment arms at any given time point assessment during the trial (even if this only occurred in one HRQOL domain).

The selected articles were evaluated for trial quality and quality of reporting on HRQOL.

Trial quality was evaluated with the Jadad scale [21]. The maximum possible score was 13 points using an 11 item instrument. This was considered to be good when the score was more than 9 points and poor when the score was equal to or less than 9 points. Items related directly to the control of bias using the Jadad scale are:

- Was the study designed as randomised?
- Was the study designed as double blind?
- Was there a description of withdrawals and drop outs?

Other markers not related directly to the control of bias:

- Were the objectives of the study defined?

- Were the outcome measures defined clearly?
- Was there a clear description of the inclusion and exclusion criteria?
- Was the sample size justified (for example, power calculation)?
- Was there a clear description of the interventions?
- Was there at least one control (comparison) group?
- Was the method used to assess adverse effects described?
- Were the methods of statistical analysis described?

Items are scored as follows:

- Give either a score of 1 point for each « yes » or 0 points for each « no ». There are no inbetween marks.
- Give 1 additional point if, for question 1, the method to generate the sequence of randomisation was described and was appropriate (table of random numbers, computer generated, etc.) and/or if, for question 2, the method of double blinding was described and was appropriate (identical placebo, active placebo, dummy, etc.).
- Deduct 1 point if, for question 1, the method to generate the sequence of randomisation was described and was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc.) and/or if, for question 2, the study was described as double blind but the method of blinding was inappropriate (for example, comparison of tablet versus injection with no double dummy).

The criteria used to evaluate quality of reporting on HRQOL were based on those proposed by Efficace et al [11] (table 1)

This 11-item checklist was developed on the basis of good practice in conducting a HRQOL evaluation and it was specifically aimed at evaluating the reported quality of the HRQOL assessment methodology in a clinical trial setting. The checklist items were devised to have a dichotomous answer; these can be scored as « yes » (giving a score of 1) or « no » (giving a score of 0), the higher the score the higher the considered robustness of the outcomes. This checklist addresses the basic and essential issues that a given trial should report to have methodologically sound outcomes.

The original checklist also included whether the measure covered, at least, the main HRQOL dimensions relevant for a generic cancer population. This criterion has been built into this review automatically.

Studies scoring at least seven on this checklist including three mandatory items (i.e. baseline compliance, missing data and psychometric properties reported) could be considered as 'probably robust'. Hence, all studies were classified into « probably robust » (as defined above), 'limited' (scoring higher than three but either lower than seven or not including all three mandatory items), and 'very limited' (all other studies, i.e. scoring three or lower on the checklist score).

When an article provided explicit reference to a related paper reporting additional data, this was retrieved as well. When more than one paper reported HRQOL data of the same trial, information was pooled to be reported in the tables.

Ethical aspects

In accordance with French regulations, this study was exempted from IRB approval.

RESULTS

The appendix list all articles reviewed.

According to the eligibility criteria, a total of 26 citations were identified that included HRQOL outcomes in 24 randomized clinical trials. Besides these, three other studies were also retrieved, but excluded from trial analyses (with the consensus of all authors). One of these studies met our criteria, but did not report any details about the methodology used to assess HRQOL and the remaining two was excluded because it was impossible to check for the HRQOL measure used.

Demographics and trial design characteristics

The studies were conducted across a variety of countries: 16 (66,6%) in European countries, two (8,3%) in the USA, two (8,3%) in Asia, one (4,2%) in Australia, one (4,2%) in Burkina Faso. and two (8,3%) were conducted on an international setting.

Eight (33,3 %) of the 24 studies were industry sponsored, as identified by author affiliation with a company or by a statement regarding commercial funding.

Half of the trials were multicenter studies.

The number of patients enrolled into the trials varied considerably, ranging from 27 to 700 patients, with a total of 3476 patients.

A total of 23 trial reports (95.8%) stated that a research ethics committee had approved the research and reported that informed consent from patients had been requested from the participants.

Methodological quality

The methodological quality score (Jadad scale) ranged from 8 to 13, with a mean of 10.6 \pm 1.07. Only 5 articles, in which the study was described as double blind and could therefore score an extra 2 points, were eligible for the maximum possible score of 13, and just 2 achieved this score.

The methodological quality was insufficient (score < 9) for only one trial (table 2). All trials were randomized, but three trials (12,5%) were randomized without the method of randomization being specified. All but one trial detailed trial inclusion and exclusion criteria.

HRQOL assessment

The 24 RCTs identified were classified according to the predefined checklist. One of these could be considered as very limited in terms of methodological design according to previously defined criteria (4%). Twelve trials (50%) were considered limited while eleven (45,8%) were evaluated as being probably robust . The overall level of reporting is provided in table 3 .

14 distinct QOL questionnaires (5 generic, 9 specific) were used in the 24 studies analyzed. The most frequently used instrument (in conjunction or not with other tools) was the Short Form 36-item questionnaire in 13 trials, and the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 in four trials.

In two studies SF-36 was administered in a modified version, attempting to give a more comprehensive HRQOL assessment but altering the psychometric properties of the original tool.

45 % (11/24) of the reports described the questionnaires, i.e. the number of dimensions and their contents, such as the number of items per dimension, and the minimal and maximal scores. Only the name of the questionnaire was given in 13 others.

QOL was a primary endpoint in only six (25%) studies. QOL was a secondary endpoint in 75% (18/24) of the analyzed reports. In two of them, trial results were published in several different articles, one relating the QOL findings and the other(s) the clinical outcomes. Only five (20,8 %) studies reported the use of power calculations for HRQOL aspects.

Only ten studies (41,6%) reported a priori hypotheses and only five studies (20,8%) provided a rationale for selecting a HRQOL measure.

Information about the administration of the HRQOL questionnaire was mentioned in 13 (54,1%) reports.

Three RCTs (12,5%) did not provide the absolute number or the percentage of patients who completed the questionnaire before commencing the trial.

Methods of health-related quality of life analysis and results

The response rate for quality of life end points was given in 19 of the studies, with response rates ranging from 14,2 % to 100%.

Nine RCTs (37,5%) did not provide any details about HRQOL missing data during the course of the trial. Furthermore, in those trials where an indication of HRQOL missing data was provided, only one trial undertook a detailed statistical exploration of the biases due to missing data. The remaining studies did not investigate this issue.

Reporting the level of missing data and the reasons why the data is missing (ie, random or systematic) are factors critical to understanding any possible source of bias in determining HRQOL significance

However, all studies provided details about the timing of HRQOL assessment and 14 (58,3%) discussed somehow the HRQOL outcomes in the paper.

All RCTs, with the exception of one, applied a statistical test for determining a HRQOL difference between treatment arms.

Of the 23 eligible studies, 11 (47,8%) found some significant difference on HRQOL scales between arms.

Obtaining a statistical difference in terms of HRQOL outcome does not necessarily imply a clinically meaningful difference from a patients' perspective (24). But only six (25%) discussed the related HRQOL outcomes in terms of clinical significance from a patient's perspective. This issue is closely related to the difficulty in interpreting the HRQOL data from a given measure.

DISCUSSION

Inadequate reporting of randomised controlled trials is common and hampers the appraisal of the validity and generalisability of results [22 , 23]. To overcome such problems, the Consolidated Standards for reporting of Trials (CONSORT) Group developed the CONSORT statement [24] in 1996, which was followed by a revised version in 2001[25].

This can explain the high level of the mean Jadad score found in our study.

The main objective of this article was to evaluate the methodological quality of RCTs with a HRQOL component in gastrointestinal surgery.

Using the stated selection and eligibility criteria, we found 24 RCTs with HRQOL assessment which included some 3476 patients. HRQOL was a secondary end point in most trials (75%).

The aim of our study was not to compare results with other medical or surgical specialities but to describe how digestive surgeons use HRQOL in their trials and to find how to improve. However, compared with other HRQOL studies in other disease sites and treatments [12 –19 , 26 , 27] (table 4) the overall quality of the reported trials is good, although there are a number of shortcomings with regard to the reporting of the HRQOL design and results.

Among the studies reviewed, there is generally poor details about the rationale for selecting a specific measure and instrument administration.

A justification for selecting the HRQOL measure was given in only five studies (20,8%). In others reviews this justification was given in 9,7 to 90% of the trials (table 4). This point has to be improved as well in surgery as for others specialities.

This is regarded as important because instrument selection is critical for reliability, validity, and reproducible results.

It is also important to standardize the instructions and completion procedures of assessment material administered to patients, particularly in RCTs, because often many researchers and institutions are involved in collecting HRQOL data. Thus, standard procedures help to ensure adequate data quality and minimize any possible bias in data collection. (For example, will patients answer the questionnaire at a clinic visit, by telephone, by mail? Is it to be done by the investigator, his colleagues, or a research nurse?)

Our study was in accordance with previous studies (table 4) where instrument administration was reported in 0 to 66,6%.

A major methodological drawback was a lack of a priori hypothesis about possible HRQOL changes before commencing the trial. Only ten RCTs (41,6%) explicitly stated an a priori hypothesis thus limiting spurious HRQOL results due to multiple significance testing. This results was similar and even better to the majority of previous study where the a priori hypothesis was stated in 13 to 72,7% of the trials (table 4).

A key consideration for future studies is the selection of a limited number of HRQOL indicators before commencing the trial, possibly basing this selection on previous related trials, or on a specific a priori research hypothesis about the impact of a given therapy.

One major issue is the reporting of compliance at baseline and the documentation of missing data. Three RCTs (12,5 %) did not provide the absolute number or the percentage of patients who completed the questionnaire before commencing the trial and 9 (37,5%) did not provide any details about HRQOL missing data during the course of the trial.

This result was similar to previous studies where missing data were documented in only 48,4 to 74,8% (table 4).

Although the majority of trials started with reasonable sample sizes, many were plagued with problems of patient drop-out. Such attrition often limits the general robustness of the results and reduces confidence in the HRQOL conclusions. Data are generally not missing at random and therefore bias can be introduced [17 , 28 –30]. The benefit of a intervention may be overestimated by comparison of group means as only individuals who remain well enough to fill in questionnaires provide data. The unreported details of missing data is a frequent problem in studies where HRQOL is measured [31] and previous works already proposed procedures to address this issue [32 , 33]. More attention to improving compliance and reporting in future studies would be valuable.

Of the 23 eligible studies, 11 (47,8%) found some significant difference on HRQOL scales between arms. This would indicate that the HRQOL measures are valuable in providing additional data.

However, although HRQOL differences were observed, it is necessary to remember that, whereas many subscales are often used and compared over treatments and time, not all subscales will show a significant difference.

This underlines the need to declare in advance the HRQOL hypotheses and the importance of careful interpretation of multiple repeated statistical analyses.

A further trap in analysis of HRQOL data is the difference between statistical and clinical significance in changes of scores. It is acknowledged that, although analysis of large samples may reveal small changes that seem to be statistically significant, these changes may not be clinically meaningful to the patient and are, therefore, of limited value to the improvement of patient care.

An effort to determine if such small numerical differences have a clinical meaning from a patient's perspective has been highlighted as an important aspect for determining the impact of a given treatment [28 , 34].

Unfortunately, only six of these studies (26%) examined the clinical significance of apparent differences. It is highly desirable that future studies will routinely include the concept of clinical significance to help evaluate the value of HRQOL results.

Furthermore, in several RCTs, the HRQL results were not formally presented, but the main results were described in the text.

It is possible that this occurred because most trials used HRQOL as a secondary end point. In such trials, it is frequently observed that limited space is given to HRQOL data, with priority given to the primary clinical end point.

Two authors have overcome this difficulty by separately reporting clinical and HRQL trial outcomes.

This is an opportunity for adequate explanation and presentation of what may often be complex results.

However the disadvantage of splitting the HRQL data from the main trial paper is that surgeons are unlikely to read the HRQL paper once the main clinical message of a particularly trial has been published. If this occurs, then during the process of clinical decision making the HRQL impacts of treatment may be overlooked [19]. It is therefore recommended that clinical and HRQL outcomes are published together so that clinical decision making is based upon relevant patient-centered endpoints.

Whilst we identified the above reported methodological limitations, it was impressive that nearly all the studies used HRQOL valid measures and provided details on the HRQOL timing of assessment during the trial.

There were eleven trials (45,8%) with robust HRQL design, and statistically significant differences in HRQL were reported in six of these trials. Only one trial « very limited » (4%) could be invalidated by its lack of rigor in presenting HRQOL data.

Pertinently, the strict methodological approach to the assessment of the patient-based QOL criteria in the evaluation of therapeutic strategies can help, patients and their doctors, in medical decision making.

The Efficace's checklist can be considered as a minimum standard; however, HRQOL design also greatly depends on the context and the specific research question of the trial; hence good reports may have different emphases and some issues might have different relevance according to the specific study questions [11].

If HRQL is considered to be a relevant outcome in a clinical trial and if HRQL is assessed robustly, then it will always contribute to clinical decision making regardless of the direction of the outcomes [11 , 26]. Only if HRQL assessments are flawed (underpowered study, too many missing data items, invalid questionnaires) may they not contribute to clinical decision making.

Our study showed some limitations. We selected randomised controlled trials (RCTs) from 12 journals. This restrictive choice was led by the recognised quality of the four medical journals selected (leading journals that publish research reports in all fields and have a broad readership) and because the eight surgical journals comprised a good sample of surgery around the world. The purpose of this choice was to create a homogeneous group of publications and conditions that allowed standardised analysis. This arbitrary choice may have introduced a bias causing overestimation of the quality of the RCTs analysed. In addition, authors of original articles were not pursued for additional data or for clarification of points that were unclear about trial methodology.

We recognise also that our review is limited by its restriction to RCTs, but the checklist developed by Efficace and al [11] was originally devised only for this type of design. It could be interesting to develop an applicable and useful checklist for not randomized studies.

Despite these potential limitations, this paper suggests that surgeons are interested in HRQL outcomes and that HRQOL assessment in RCT settings has the potential to provide invaluable data for developing new treatments in gastrointestinal surgery.

CONCLUSION

Despite the emphasis on quality of life outcomes, there are still deficiencies in the execution of these studies.

In attempt to improve in the future the use of QOL as an endpoint in RCTs, researchers should include prior statement of hypotheses, better reporting, improved compliance, detailed methods of analysis, and reporting of missing data.

Such improvements will enhance the reliability of future HRQOL investigations with the ultimate aim of giving a comprehensive picture of a given treatment and also facilitate clinical decision-making.

References:

1. Kluivers KB, Riphagen I, Vierhout ME, Brolmann HA, de Vet HC. Systematic review on recovery specific quality-of-life instruments. *Surgery*. 2008; 143: 206 - 215
2. Schumacher M, Olschewski M, Schulgen G. Assessment of quality of life in clinical trials. *Stat Med*. 1991; 10: 1915 - 1930
3. Leplege A, Hunt S. The problem of quality of life in medicine. *Jama*. 1997; 278: 47 - 50
4. Kong SX, Gandhi SK. Methodologic assessments of quality of life measures in clinical trials. *Ann Pharmacother*. 1997; 31: 830 - 836
5. Velikova G, Stark D, Selby P. Quality of life instruments in oncology. *Eur J Cancer*. 1999; 35: 1571 - 1580
6. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol*. 1998; 51: 1055 - 1068
7. Eypasch E, Williams JI, Wood-Dauphinee S, Ure BM, Schmullig C, Neugebauer E, Troidl H. Gastrointestinal Quality of Life Index: development, validation and application of a new instrument. *Br J Surg*. 1995; 82: 216 - 222
8. Guyatt GH, Bombardier C, Tugwell PX. Measuring disease-specific quality of life in clinical trials. *Cmaj*. 1986; 134: 889 - 895
9. Staquet M, Berzon R, Osoba D, Machin D. Guidelines for reporting results of quality of life assessments in clinical trials. *Qual Life Res*. 1996; 5: 496 - 502
10. Avery K, Blazeby JM. Quality of life assessment in surgical oncology trials. *World J Surg*. 2006; 30: 1163 - 1172
11. Efficace F, Bottomley A, Osoba D, Gotay C, Flechtner H, D'Haese S, Zurlo A. Beyond the development of health-related quality-of-life (HRQOL) measures: a checklist for evaluating HRQOL outcomes in cancer clinical trials--does HRQOL evaluation in prostate cancer research inform clinical decision making? *J Clin Oncol*. 2003; 21: 3502 - 3511
12. Efficace F, Bottomley A. Health related quality of life assessment methodology and reported outcomes in randomised controlled trials of primary brain cancer patients. *Eur J Cancer*. 2002; 38: 1824 - 1831
13. Efficace F, Bottomley A, van Andel G. Health related quality of life in prostate carcinoma patients: a systematic review of randomized controlled trials. *Cancer*. 2003; 97: 377 - 388
14. Bottomley A, Efficace F, Thomas R, Vanvoorden V, Ahmedzai SH. Health-related quality of life in non-small-cell lung cancer: methodologic issues in randomized controlled trials. *J Clin Oncol*. 2003; 21: 2982 - 2992
15. Bottomley A, Therasse P. Quality of life in patients undergoing systemic therapy for advanced breast cancer. *Lancet Oncol*. 2002; 3: 620 - 628
16. Goodwin PJ, Black JT, Bordeleau LJ, Ganz PA. Health-related quality-of-life measurement in randomized clinical trials in breast cancer--taking stock. *J Natl Cancer Inst*. 2003; 95: 263 - 281
17. Efficace F, Bottomley A, Vanvoorden V, Blazeby JM. Methodological issues in assessing health-related quality of life of colorectal cancer patients in randomised controlled trials. *Eur J Cancer*. 2004; 40: 187 - 197
18. Efficace F, Kemmler G, Vignetti M, Mandelli F, Molica S, Holzner B. Health-related quality of life assessment and reported outcomes in leukaemia randomised controlled trials - a systematic review to evaluate the added value in supporting clinical decision making. *Eur J Cancer*. 2008; 44: 1497 - 1506
19. Gujral S, Avery KN, Blazeby JM. Quality of life after surgery for colorectal cancer: clinical implications of results from randomised trials. *Support Care Cancer*. 2008; 16: 127 - 132
20. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *Jama*. 2001; 285: 1992 - 1995
21. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996; 17: 1 - 12
22. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care*. 1996; 12: 195 - 208
23. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *Bmj*. 2001; 323: 42 - 46
24. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *Jama*. 1996; 276: 637 - 639
25. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2001; 357: 1191 - 1194
26. Blazeby JM, Avery K, Sprangers M, Pikhart H, Fayers P, Donovan J. Health-related quality of life measurement in randomized clinical trials in surgical oncology. *J Clin Oncol*. 2006; 24: 3178 - 3186
27. 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: 775 - 781
28. 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: 139 - 144
29. Sprangers MA, Moynihan CM, Moynihan TJ, Patrick DL, Revicki DA. Assessing meaningful change in quality of life over time: a users' guide for clinicians. *Mayo Clin Proc*. 2002; 77: 561 - 571
30. Troxel AB, Fairclough DL, Curran D, Hahn EA. Statistical analysis of quality of life with missing data in cancer clinical trials. *Stat Med*. 1998; 17: 653 - 666
31. Bernhard J, Cella DF, Coates AS, Fallowfield L, Ganz PA, Moynihan CM, Mosconi P, Osoba D, Simes J, Hurny C. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med*. 1998; 17: 517 - 532
32. Conroy T, Bleiberg H, Glimelius B. Quality of life in patients with advanced colorectal cancer: what has been learnt? *Eur J Cancer*. 2003; 39: 287 - 294
33. Young T, de Haes H, Curran D. Guidelines for Assessing Quality of Life in EORTC Clinical Trials. The EORTC Quality of Life Group & EORTC Quality of Life Unit, version 2.0. EORTC Publications; Brussels 2002;
34. Osoba D. A taxonomy of the uses of health-related quality-of-life instruments in cancer care and the clinical meaningfulness of the results. *Med Care*. 2002; 40: III31 - 38

Table 1Minimum Standard Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials [¹¹]

HRQOL Issue	Answer	Description
Conceptual		
A priori hypothesis stated	Yes No N/A *	Assessed whether authors had a predefined HRQOL end point and/or stated expected changes because of the specific treatment.
Rationale for instrument reported	Yes No	Assessed whether authors gave a rationale for using a specific HRQOL measure.
Measurement		
Psychometric properties reported	Yes No	Assessed whether a previously validated measure was used or psychometric properties were reported or referenced in the article.
Cultural validity verified	Yes No N/A †	Assessed whether the measure was validated for the specific study population.
Adequacy of domains covered	Yes No	Assessed whether the measure covered, at least, the main HRQOL dimensions relevant for a generic cancer population and/or according to the specific research question.
Methodology		
Instrument administration reported	Yes No	Assessed whether authors specified who and/or in which clinical setting the HRQOL instrument was administered.
Baseline compliance reported	Yes No	Assessed whether authors reported the number of patients providing an HRQOL assessment before the start of treatment.
Timing of assessments documented	Yes No	Assessed whether authors specified the HRQOL timing of assessment during the trial.
Missing data documented	Yes No	Assessed whether authors gave some details on HRQOL missing data during the trial.
Interpretation		
Clinical significance addressed	Yes No	This refers to the discussion of HRQOL data being clinically significant from a patient's perspective and not simply statistically significant.
Presentation of results in general	Yes No	Assessed whether authors discussed the HRQOL outcomes, giving any comments regardless of the results (either expected or not).
Abbreviation: HRQOL, health-related quality of life.		
* If a study explicitly states an exploratory HRQOL evaluation.		
† If the HRQOL measure is validated in the same population as the one of the trial.		

Table 2numbers of trials according to the Jadad score [²¹] (n=24)

Score	≤7	8	9	10	11	12	13
No. trials	0	1	2	7	12	0	2

Table 3

Level of Reporting According to the Minimum Standard Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials

HRQOL Issue	No [*]	%
Conceptual		
A priori hypothesis stated	10/24	41,6
Rationale for instrument reported	5/24	20,8
Measurement		
Psychometric properties reported	22/24	91,6
Cultural validity verified	22/24	91,6
Methodology		
Instrument administration reported	13/24	54,1
Baseline compliance reported	21/24	87,5
Timing of assessments documented	24/24	100
Missing data documented	15/24	62,5
Interpretation		
Clinical significance addressed	8/24	33,3
Presentation of results in general	23/24	95,8

Abbreviation: HRQOL, health-related quality of life.

* Number of articles reporting item/number of articles to which item is applicable.

Table 4

Results of previous review on RCTs reporting HRQOL evaluation by the Efficace's checklist

			Conceptual		Measurement		Methodology			Interpretation		
			A priori hypothesis stated	Rationale instrument reported	Psychometric properties reported	Cultural validity verified	Instrument administration reported	Baseline compliance reported	Timing assessments documented	of Missing data documented	Clinical significance addressed	Presentation of results general
Our study	Gastrointestinal surgery (2006–2007)	24	41,6 %	20,8%	91,6%	91,6%	54,1%	87,5%	100%	62,5%	33,3%	95,8%
Efficace 11]	[Prostate cancer (1980–2001)	24	13%	29%	87%	81%	25%	46%	96%	54%	12%	67%
Efficace 12]	[Brain cancer (1980–2001)	5	20%	80%	NA	50%	0%	100%	100%	60%	40%	80%
Bottomley 14]	[Non-small-cell-lung cancer (1980–2002)	29	31%	34,5%	NA	13/13	34,5%	13,8%	29/29	62%	20,7%	58,6%
Efficace 17]	[Colorectal cancer (1980–2003)	31	25,8%	9,7%	100%	94,7%	9,7%	61,3%	100%	48,4%	12,9%	93,5%
Efficace 18]	[Leukaemia (1980–2007)	9	22,2%	55,5%	100%	100%	66,6%	83,3%	100%	66,6%	44,4%	100%
Gujral 19]	[Surgery for colorectal cancer (1980–2006)	8	50%	87,5%	NA	NA	25%	20%	8/8	50%	66,6%	62,5%
Blazeby 26]	[Surgical oncology (1985–2005)	33	72,7 %	90,9 %	NA	NA	63,6%	45,4%	90,9%	54,5%	NA	NA
Efficace 27]	[Cancer clinical trials (1990–2004)	159	28,9%	47,8%	94,3%	96,6%	30,8%	75,5%	99,4%	74,8%	24,5%	93,1%

NA= not available