A systematic review of radiotherapy toxicity reporting in randomized controlled trials of rectal cancer: A comparison of patient-reported outcomes and clinician toxicity reporting

Running Title: Review of rectal radiation toxicity
Abstract

The use of multimodal treatments for rectal cancer has improved cancer-related outcomes but makes monitoring toxicity challenging. Optimizing future radiotherapy regimens requires collection and publication of detailed toxicity data. This review evaluated the quality of toxicity information provided in randomized controlled trials (RCTs) of radiotherapy in rectal cancer and focused on the difference between clinician-reported and patient-reported toxicity.

Medline, EMBASE and the Cochrane Library were searched (January 1995-July 2013) for RCTs reporting late toxicity in patients treated with regimens including preoperative (chemo)radiotherapy. Data on toxicity measures and information on toxicity reported was extracted using QUANTEC (Quantitative Analyses of Normal Tissue Effects in the Clinic) recommendations. International Society for Quality of Life Research (ISOQOL) standards on patient-reported outcomes (PROs) were used to evaluate the quality of patient-reported toxicity.

21 RCT publications met inclusion criteria out of 4144 articles screened. All PRO papers reported higher rates of toxicity symptoms than clinician-reported papers and reported on a wider range and milder symptoms. No clinician-reported paper published data on sexual dysfunction. 55% of clinician-reported papers grouped toxicity data related to an organ system together (e.g. ‘Bowel’) and 45% presented data only on more severe (≥grade 3) toxicity. In comparison, all toxicity grades were reported in 79% of PRO publications and all studies (100%) presented individual symptom toxicity data (e.g. bowel urgency). However, PRO reporting quality was variable. Only 43%
of PRO studies presented baseline data, 28% did not use any psychometrically validated instruments and only 29% of studies described statistical methods for managing missing data.

Analysis of these trials highlights the lack of reporting standards for adverse events and reveals the differences between clinician and patient-reporting of toxicity. Recommendations for improving the quality of adverse event data collection are provided with the aim of improving critical appraisal of outcomes for future studies.
Introduction

Rectal cancer is diagnosed in approximately 40000 people annually in the United States [1]. The approach to rectal cancer treatment is multimodal with the majority of patients receiving either preoperative radiation or chemoradiation. The use of multimodal treatments has improved cancer-related outcomes, however it has also led to an increase in substantial immediate and late adverse events/toxicity[2,3].

Reliable collection and analysis of adverse event data in oncology is challenging as complex multimodal regimens, such as in locally advanced rectal cancer, involve not only different treatments but also variations in dose intensity and duration[4]. Methods for toxicity data capture and reporting in oncology were developed from other disciplines which employ treatments with a different, and often less toxic profile, such as antibiotics[4]. Adverse events in oncology may be inadequately captured by these methods and are often underreported[5]. A number of international reports, including QUANTEC (Quantitative Analysis of Normal Tissue Effects in the Clinic), have highlighted that in order to optimize future radiation treatment regimens a systematic approach to the collection and publication of detailed toxicity data is required[5]. Newer radiotherapy techniques such as Intensity Modulated Radiotherapy (IMRT) aim to reduce toxicity by reducing the amount of normal tissue exposed to high doses of irradiation. One important consequence is that more organs are exposed to a low dose of irradiation than in conventional
treatment. This has an unknown impact on late toxicity and rigorous toxicity reporting methods are required to capture this data[6].

The clinician-reported Common Terminology Criteria for Adverse Events (CTCAE) version 4 has recently been accepted as the preferred instrument for collection of adverse event data in cancer trials[7]. However, patient-reported outcomes (PROs) included in trials are increasingly used as a surrogate measure of late toxicity, usually as a secondary outcome. Using PROs has been found to increase the number and variety of adverse events recorded and highlighted discrepancies between clinician and patient reporting[8,9]. The inclusion of PROs in clinical trials may therefore provide additional information to better inform clinical decision-making. However, a number of reviews of PROs in clinical trials have revealed concerns regarding the methodological quality and reporting of the results[10-12]. Two recently published internationally developed guidelines highlight this area of concern[13,14].

Previous reviews of radiotherapy treatment in rectal cancer have focused on survival outcomes and descriptions of late adverse events or functional outcomes in a variety of different trial settings, including retrospective studies[15-19]. This review focuses on randomized controlled trials (RCTs) in rectal cancer, as the research gold standard, with the following objectives: (1) to establish the clinician and patient-reported toxicity instruments used; (2) to assess the methodological quality of the studies and quality of PRO reporting;
and (3) to report a summary of the percentage of toxicity reported by treatment received and compare differences in clinician and patient reporting. The review concludes with recommendations for improving adverse event data collection from clinicians and PROs and describes the impact of reporting quality of adverse events on the ability to establish safe dose constraints for normal tissues in future optimization studies for radiation treatments.

**Methods**

**Search strategy**

Medline, EMBASE and the Cochrane Library were searched from January 1995 to July 2013 for RCTs reporting late toxicity in patients treated with regimens including preoperative (chemo)radiotherapy. The search followed Centre for Reviews and Dissemination recommendations for undertaking systematic reviews[20] and PRISMA guidelines[21] (Appendix 1.1). Only English language publications were included. Relevant studies listed as references were hand searched.

**Selection criteria**

All Phase II and III RCTs in adult patients with a localized resectable rectal cancer were eligible if patients were randomized to at least one arm of preoperative radiation or chemoradiation. Studies of patients treated only with postoperative radiation were excluded unless in a comparison study with a preoperative radiotherapy arm. Studies of surgery alone, intraoperative
radiation or brachytherapy were not eligible. Conference abstracts were excluded.

**Outcome measures examined**

Studies including clinician-reported toxicity and/or patient reporting on symptoms or some other aspect of health-related quality of life (HRQOL) as a primary or secondary outcome were considered. PROs were defined as any reports coming directly from the patient[22]. Late toxicity was defined by side effects present from three months post radiotherapy treatment[23]. Any secondary analysis papers of late toxicity were reviewed in conjunction with the original publication. Multi-dimensional PRO measures (for example a measure covering different aspects of functioning such as physical, emotional or cognitive function) or single-item health outcomes were included if patient-reported. Clinician conducted interviews, structured using PRO questionnaires, were considered as clinician-reported. Studies reporting post-operative complications or patient satisfaction were excluded.

**Data extraction and type of information extracted**

The identified RCTs were assessed using a predefined data extraction form adapted from a published checklist to include clinician-reported toxicity studies[11]. Data on toxicity measures and detailed information on how toxicity was reported was extracted using QUANTEC recommendations[24]. Three reviewers (XX, XX, XX) independently screened the titles and abstracts of all retrieved studies. In cases of disagreement the full articles were revisited
to reconcile differences and achieve consensus. XX and XX independently extracted and analyzed the data from all chosen articles. Differences were reconciled through discussion.

Data was extracted into a predefined database for each RCT on (1) basic trial demographics (e.g. publication year, trial phase, design); (2) clinical demographics (e.g. overall sample size, sample size for toxicity reporting, treatment regimens, primary endpoints); (3) adverse event reporting (e.g. toxicity measure(s) used, grade and percentage of toxicity reported) and (4) methodological quality (e.g. quality of PRO-reporting, risk of bias assessment, statistical analysis and presentation of results).

In trials with multiple publications the results are presented separately when data on different side effects and/or time points was presented or the methodological reporting quality varied.

**Quality assessment of RCTs and PRO reporting**

Internal validity was assessed by applying the Cochrane Risk of Bias tool to evaluate: adequacy of sequence generation; allocation concealment; blinding; incomplete outcome data; and selective reporting[25]. PRO quality assessment was adapted from the recently published ISOQOL recommended standards [14].
Results

The search yielded 5682 records (figure 1). 4144 records were screened after duplicates and articles published before 1995 were removed. 21 publications representing data from 13 different RCTs fulfilled the inclusion criteria (Table 1)[26-46]. The median duration of follow up for all studies was 5 years. Toxicity was a secondary endpoint in all but one trial. Park et al[45] included toxicity as part of multiple primary endpoints. In the studies where statistically significant cancer outcomes were achieved (Stockholm/Swedish; Dutch Total Mesorectal Excision (TME) and CRO7 trials), these were associated with deterioration in some aspect of patient or clinician-reported late toxicity (see table 1 for details)[47,48]. Only one trial disclosed industry funding/affiliations[46].

Insert figure 1

Insert table 1

Methods of toxicity reporting: Clinician-reporting versus PROs

Table 2 summarizes the data extracted from the publications using QUANTEC recommendations for toxicity reporting. In total, 15 different PRO instruments were used in 14 publications and seven different clinician-reported instruments in 11 publications. RTOG/EORTC was used most commonly for late toxicity clinician-reporting (n=4) followed by the CTCAE (n=2). EORTC-QLQ core questionnaire (C30) and colorectal cancer-specific module (CR38) were the most commonly used validated PROs (n=4/n=2
respectively). Baseline symptoms alongside acute and/or late PRO toxicity were reported in 6 out of the 7 longitudinal PRO studies [34,35,38-40,46]. The remaining eight PRO studies used a cross sectional design to assess late toxicity or HRQOL at a single time-point. Only one of the 11 longitudinal clinician-reported papers published baseline symptoms[46]. In almost half of the clinician-reported papers only the more severe grades of toxicity, ≥grade 3, were reported (n=5; 45%). In comparison 79% (n=11) of the PRO publications, published data on the full range of toxicities (from no symptoms to severe toxicity).

**Insert table 2**

The most frequently reported late adverse event in any RCT was related to bowel toxicity (84% of publications) followed by urinary dysfunction (40%) (Table 2). *None* of the 11 clinician-reported papers reported on sexual dysfunction. 50% (n=7) of the PRO publications covered sexual dysfunction, 43% (n=6) also reported on HRQOL, mainly using the EORTC QLQ-C30 (n=4). Skin toxicity (n=5; 45% clinician-reported papers) and hematological toxicity (n=5; 45%) are reported in the clinician-reported papers and not in the PRO publications.

The majority of clinician-reported publications grouped symptoms referable to the bowel or bladder as a single organ, reporting on ‘small/large bowel’ or ‘bowel’ or ‘bladder’ toxicity or only reporting all ≥grade 3 toxicities (n=7; 64%).
In comparison, all PRO studies (n=14; 100%) reported a breakdown of individual symptoms, for example fecal incontinence, or a combination of individual symptoms and a summary score of multiple items as implemented in the EORTC-QLQ system.

**Frequency of symptomatic toxicity: Clinician-reporting versus PROs**

Table 3 shows the frequency of toxicity reported as a range of percentages separated by treatment received. Higher rates of toxicity symptoms were described in the patient-reported papers in comparison with clinicians. In the clinician-reported papers ≥grade 3 bowel toxicity was reported at rates ranging from 1.4-9%. Fecal incontinence and diarrhea were reported at rates of around 9%. Bladder toxicity ≥grade 3 was reported at lower rates between 1-2%.

In the patient-reported papers, fecal incontinence rates varied between 8-50% for solid stools and 24-72% for liquid (or non-specified) stools. Urinary incontinence rates were between 18-45%. None of the clinician-reported papers mentioned sexual dysfunction, which was reported in seven PRO papers. Between 70-80% of male patients reported a decline in sexual function, with 71% reporting erectile dysfunction in one study[46]. Another study reported severe dysfunction scores using the International Index of Erectile Function questionnaire[33]. EORTC-QLQ-CR38 mean scores for sexual dysfunction ranged from 40.8 to 65.7 (with a higher score, up to 100, indicating more problems). In women, 41-52% reported a decline in sexual
function[42] and EORTC-QLQ-CR38 mean scores ranged from 29.9 to 50[35]. 86-100% vaginal dryness and dyspareunia ranging between 50-86% was reported in one study[33]. Two PRO studies were unable to report in detail on sexual dysfunction outcomes due to a paucity of response data[30,40].

The results of the 22921-EORTC trial reveal the potential differences in clinician-reported toxicity and PRO data over and above the clear differences in symptom frequency reported. The clinician-reported paper did not detect/find any significant differences in toxicity between the four different treatment arms[29] however the cross-sectional PRO study using the EORTC QLQ-C30/CR38 found an increase in diarrheal symptoms in patients treated with chemotherapy at any stage as well as lower social and role functioning and overall global QOL[30].

Insert table 3

**Quality assessment of RCTs and PROs**

The RCTs varied little in the Cochrane Risk of Bias tool assessment with 16 studies with an overall low risk of bias assessed (76%; table 2). The response rates for the studies including PROs varied widely. The response rates for single cross-sectional assessments (n=6) varied between 55 and 90% and from 49% to 89% in longitudinal studies with 2 to 5 year follow up (n=7). Paper data collection was used in seven studies and was not explicitly stated in the remaining seven studies.
Table 4 shows the evaluation of PRO quality using the recently published ISOQOL recommended standards[14]. The quality of the reporting was highly variable. We considered three previously recommended key methodological criteria[49]: reporting of baseline data; statistical methodology for missing data; and the use of validated instruments. Only 43% of PRO studies presented baseline data[34,38,40,50-52] and 29% of studies described statistical methods for managing missing data[34,40,50,51]. One of the main difficulties was that eight of the PRO RCTs had used a cross-sectional design and thus could not provide baseline data [26-28,36,42]. 28% (n=4) studies did not use any psychometrically validated PRO instruments. The remaining studies either used solely psychometrically validated instruments (n=5; 36%) or a combination of validated and non-validated PROs or modified instruments (n=5; 36%).

*Insert table 4*
Discussion

This review describes the toxicity outcomes for some 8800 patients enrolled in 13 different RCTs, with 21 papers considering the impact of radiotherapy toxicity following rectal cancer treatment over the past two decades. The outcomes of these trials have determined clinical practice and the summary of reported toxicities by clinicians and patients presented in this review is relevant to all clinicians treating rectal cancer. Analysis of these trials highlights the lack of standards for adverse event reporting, both patient and clinician-reported, in cancer clinical trials and raises a number of questions about how future treatment may be optimized on the basis of past RCT results. The results support the complementary nature of the two different methods of reporting. Detailed information is more readily available from validated PROs and information on observable adverse events, such as skin reactions, available only from clinician reports.

The clinician-reported papers tend to report only the more serious toxicities (≥grade 3) and group symptoms relating broadly to a single organ unit together. The frequency of adverse event symptom reporting was consistently lower than those reported using PROs. There was also a lack of clinician-reported data on sexual dysfunction (an important clinical issue) and baseline symptoms were rarely reported. Although the approach used in the clinician-reported papers allows an overview of the adverse events that may be expected following treatment, the details are lacking and may lead to a paucity of clinically meaningful information. The lack of detail on the adverse events experienced will not only impact on the knowledge of the true
incidence of complications but may also impact research into improving treatment with radiation, effective interventions for symptoms and limit research in areas such as radiogenomics[53,54].

“Mild” symptoms (such as a CTCAE grade 1/2 diarrhea) experienced over a lifetime following treatment may have a significant impact on daily life and require intervention. Currently this data is rarely available in the clinician-reported papers. It not usual practice for clinician-reported RCTs to publish baseline symptom data, even if it has been collected pre-randomization. This finding is of particular importance when modifications to dose-volume constraints using radiobiological modeling are based on the presence, or not, of complications in particular organs at risk. If baseline symptoms are not routinely reported or considered in the analysis it may not be clear if a patient’s symptoms were present prior to treatment and thus may not be a true ‘complication’.

The international review of dose-volume-outcome data from the QUANTEC Group highlights challenges with the current systems of adverse event reporting. One of their key concerns was the impact of poor quality outcome data on the ability to improve future radiotherapy treatments by failing to provide sufficiently detailed information on which to define dose-volume constraints[24]. To improve the quality of clinician-reported data published in future studies we recommend that all grades of toxicity and individual toxicity symptoms, including sexual dysfunction, be published in clinician-reported
papers using the CTCAE with consideration of the change from patients’ baseline symptoms.

Inclusion of prospectively collected PRO data in clinical trials may offer additional benefits. Information on a wider range and milder side effects including sexual dysfunction is reported. Many validated instruments also have data on what constitutes a clinically important difference in symptom and function scores (e.g. EORTC-QLQ systems)[55]. This feature of PROs may offer some benefits over the use of the CTCAE if PRO data is used in modeling normal tissue complication probability (NTCP), as the CTCAE is not formally validated as an instrument to measure differences in adverse event severity[56]. However, the review findings raise some important methodological issues that need to be addressed to improve PRO incorporation in future clinical trials. In line with PRO-CONSORT/ISOQOL recommendations, a number of key features require consideration: (1) the choice of a validated PRO instrument; (2) methods of data collection; and (3) statistical methods to manage missing data.

(1) Choice of validated PRO instrument

The choice of instrument will depend on the outcome of interest. The clinical trials in this review have used PRO data as a surrogate measure of late toxicity and mainly included a combination of a generic instrument to cover health-related concepts, such as physical function (e.g. FACT-G; EORTC-QLQ-C30) and disease-specific instruments to cover symptoms related to the
disease and treatment (e.g. FACT-C; EORTC-QLQ-CR29). Consideration should also be given to the availability of different translations and whether any copyright costs or permissions are required. A detailed description of PRO instruments is beyond the scope of this article, however, guidance to assist selection of a suitable validated instrument is available from different sources, including PROQOLID[57], and the National Institutes’ of Health (NIH) Patient-Reported Outcome Measurement Information System (PROMIS)[58].

It is important that the validated PRO selected covers the adverse events expected with different treatment regimens. This point raises interesting issues for recommending the incorporation of PRO data collection into all Phases of clinical trials research, including Phase I. Currently, Phase I trials in rectal cancer focus mainly on the maximum tolerated dose relating to the new agent and rarely focus on the radiation-related toxicity or incorporate PRO data. Using PROs could enable data collection of milder toxicities (providing a more accurate description of patient’s subjective experience in all aspects of the treatment, including radiation), and enable the validation of new PRO symptom-related items/questionnaires for the new treatments evaluated.

(2) Methods of data collection

Traditionally PRO data has been collected in clinical trials using paper methods, as found in this review. However, a companion study (RTOG-0828) to the RTOG-0415 RCT comparing hypofractionated to conventionally
fractionated prostate radiotherapy highlighted the benefits of electronic PRO systems using Internet-based PRO data collection in a subset of patients[59]. The completion rates using the paper version of the PRO measure EPIC (Expanded-Prostate-Index-Composite), were 36% at one year as compared to 82% using Internet-based technology. This study also made use of real-time data collection and email reminders to patients when items or forms were incomplete. Electronic methods could offer additional benefits through immediate, real-time reporting of serious adverse events for patients in clinical trials.

A number of electronic PRO platforms exist within cancer care[60]. The National Cancer Institute (NCI) has developed the patient-reported extension to the CTCAE, PRO-CTCAE, for use with an electronic-based system[61]. The patient-reported items generated are mapped onto the CTCAE grading, which will provide key information about how clinician-reported toxicity grading relates to the patient-reported equivalent. The NIH’s PROMIS initiative provides free access to standardized PRO questionnaires using a Web-based platform and international extension is in development[58].

(3) Managing missing data

Only 29% of the RCTs reported on statistical methods for managing missing data. Various reasons for missing items or forms are reported: treatment or disease-related illness; being too busy; poor administration; or not wishing to complete data on sensitive issues[62,63]. Internet-based collection may
improve questionnaire administration allowing patients to complete them at home, privately, at a convenient time. However, consideration for managing missing data should be given when establishing sample size (if patient-reported toxicity is an important endpoint) and during analysis, described in detail in previous papers[64]. During analysis, ‘imputing’ is a commonly used method for managing missing items scored as part of a group of items (subscale). Provided over half of the items in the subscale have been reported the mean value of these items may be substituted for the missing item.

The PRO-CONSORT and ISOQOL guidelines provide further details on reporting standards for trials with PROs as primary or important secondary outcomes[13,14]. Key recommendations, in addition to reporting on PRO psychometrics and statistical management of missing data, include:

- identification of the PROs in the abstract as primary or secondary endpoints;
- description of the PRO-related hypothesis; reporting on PRO-specific limitations and relating results to cancer outcomes and clinical practice.

This review has limitations. Despite the use of a comprehensive search strategy, it is still possible that some RCTs reporting on toxicity were missed. Articles published after the cut-off date of this systematic literature search are not included in this review. To our knowledge only one eligible paper by Wiltink and colleagues [65] reporting on the 14 year HRQOL following the Dutch TME trial has been published since the electronic search was completed. Although this review focused on RCTs as the gold standard for
clinical decision-making, RCTs of radiation treatments may not necessarily be
the most efficiently designed studies for prospective collection of high quality
toxicity data. It could be argued that through good observational cohorts more
detailed toxicity data collection and analysis of dose-volume constraints using
radiobiological modeling may be possible.

This paper also has a number of strengths. PRO trials were evaluated using
the most up to date methodological evaluation criteria[14] and the details and
frequency of the toxicities reported were considered in relation to the
consensus developed QUANTEC recommendations[24]. These findings were
then synthesized in a way to consider the impact of toxicity data reporting on
development of future clinical trials and clinical decision-making in routine
clinical practice.

In conclusion, this review highlights the lack of reporting standards for
adverse events in both clinician and patient-reported RCTs, and describes the
inconsistency within and between clinician and patient reporting of toxicity.
The results of the review will help clinicians treating rectal cancer in designing
future trials and support consultation with patients about expected toxicities in
routine clinical practice. To significantly improve the quality of toxicity outcome
data for future studies these findings recommend greater adherence to key
guidelines in this area[13,14] for the collection and reporting of PRO data and
for more detailed publication of clinician-reported adverse event data using
the CTCAE version 4 as the current gold standard.
References


50. Lange MM, Marijnens CAM, Maas CP, Putter H, Rutten HJ, Stiggelbout AM, Meershoek-Klein Kranenburg E, van de Velde CJH, Cooperative


