Acute and late adverse events associated with radical radiation therapy prostate cancer treatment: A systematic review of clinician and patient toxicity reporting in randomized controlled trials.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/108747/

Version: Accepted Version

Article:
Holch, P, Henry, A orcid.org/0000-0002-5379-6618, Davidson, S et al. (7 more authors) (2017) Acute and late adverse events associated with radical radiation therapy prostate cancer treatment: A systematic review of clinician and patient toxicity reporting in randomized controlled trials. International Journal of Radiation Oncology*Biology*Physics, 97 (3). pp. 495-510. ISSN 0360-3016

https://doi.org/10.1016/j.ijrobp.2016.11.008

© 2016 Elsevier Inc. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International http://creativecommons.org/licenses/by-nc-nd/4.0/

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Acute and late adverse events associated with radical radiotherapy prostate cancer treatment: A systematic review of clinician and patient toxicity reporting in randomised controlled trials (RCT)

Patricia Holch1,2, Ann M Henry2,3, Susan Davidson4, Alexandra Gilbert2, Jacki Routledge4, Leanne Shearsmith2, Kevin Franks3, Emma Ingleson2, Abigail Albutt2, Galina Velikova2.

1. Psychology group, School of Social Sciences, Leeds Beckett University, Calverley Building, Room CL 815 City Campus, Leeds, LS1 9HE, UK

2. Patient Reported Outcomes Group, Leeds Institute of Cancer Studies and Pathology, University of Leeds, Bexley Wing, St James's Hospital, Beckett street, Leeds, LS9 7TF, UK

3. Leeds Teaching Hospitals NHS trust, St James's Institute of Oncology, Bexley Wing, St James's Hospital, Beckett Street, Leeds, LS9 7TF, UK

4. The Christie NHS Foundation Trust, 550 Wilmslow Rd, Manchester M20 4BX, UK

Contact for corresponding author: Patricia Holch: email: T.Holch@Leedsbeckett.ac.uk; Tel: 0113 8124950; Address: Psychology Group, School of Social Sciences, Leeds Beckett University, Calverly Building, Room CL 815 City Campus, Leeds, LS1 9HE, UK

Acknowledgements

This work was conducted at the Patient Reported Outcomes Group, Leeds Institute of Cancer Studies and Pathology, University of Leeds, St James's Hospital, Leeds UK, LS9 7TF

This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Reference Number RP-PG-0611-20008). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
Summary
Curative radiotherapy for prostate cancer requires balancing current treatment regimens against adverse event (AEs) risk. Systematic recording of AEs would inform strategies to reduce the impact on patients. To determine clinician and patient-reported-outcome (PRO) instruments used to report AEs we systematically reviewed radical prostate cancer radiotherapy RCT through MEDLINE, EMBASE and COCHRANE. For acute AEs only clinician-reported instruments were used. For late AEs 48% of studies used PROs; however a definitive instrument was not evident.

Abstract
Purpose
This review aimed to: 1) To determine the clinician and PRO instruments currently used in RCT of radical radiotherapy for non-metastatic prostate cancer to report acute and late AEs. 2) Review the quality of methodology and PRO reporting. 3) Report the prevalence of acute and late AEs.

Materials and methods
MEDLINE, EMBASE and COCHRANE databases were searched (April-August 2014) according to the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) statement. Identified reports were reviewed according to the PRO Consolidated Standards of Reporting Trials (CONSORT) guidelines and the COCHRANE Risk of Bias tool. 1149 records were screened and 21 papers included in the final review.
Results
We determined the acute and late AE for 9,040 patients enrolled in 15 different RCTs. Only clinician reported instruments were used to report acute AE < 3 months (e.g. RTOG & CTCAE). For late clinician reporting LENT-SOMA & RTOG were used and often augmented with additional items to provide comprehensive coverage of sexual functioning and anorectal symptoms. Some late AE were reported (48% papers) using PROs (e.g. ULCA-PCI [University of California, Los Angeles Prostate Cancer Index], FACT-G & P [Functional assessment of Cancer therapy General & Prostate Module], EORTC QLQC-30 +PR25 [European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire & Prostate Module] however a definitive 'preferred' instrument was not evident.

Discussion
Our findings are at odds with recent movements towards including patient voices in AEs reporting and patient engagement in clinical research. We would recommend including PRO to evaluate radical radiotherapy before, during and after the treatment to fully capture patient experiences, and support the development of predictive models for late effects based on the severity of early toxicity.

Conclusion
Patient reporting of acute and late AEs is under-represented in radiotherapy trials. We recommend working towards a consistent approach to PRO assessment of radiotherapy related AEs.
Introduction

Curative or radical radiotherapy (RT) for non-metastatic prostate cancer is a commonly used primary treatment. Men have good outcomes in terms of cancer control but if RT AEs occur they may persist for a significant period of time, if not permanently [1]. Up to 20% will have gastrointestinal (GI) side effects, and 30-45% post-RT sexual dysfunction [2]. However, the true extent of late effects may be underestimated as many studies report only clinician assessed toxicity and AEs are often under-reported [3, 4]. Current best practices for radical prostate cancer treatment include hormone and chemotherapy treatments, improved RT delivery such as intensity modulated radiotherapy (IMRT) and image guided radiotherapy (IGRT), and hypofractionated regimens leading to higher cure rates balanced against AEs risk [5-7]. These recent advances make monitoring and accurate recording of AEs an ever more desirable goal. Systematic recording of both acute and late AEs would ultimately inform and enable strategies to reduce the long-term impact on patients [8] and allow the development of models relating RT dose to AEs.

Furthermore, with better knowledge of AEs, patients may receive appropriate specialist support for late AEs management.

Clinician-reported AEs in radiotherapy is usually achieved via the Radiation Therapy Oncology Group (RTOG) scale (acute) and for late effects the Late Effects on Normal Tissues–Subjective, Objective, Management and Analytic (LENT-SOMA) scale, now incorporated into the into the Common Terminology Criteria for Adverse Events (CTCAEs v 4.0) from the National Cancer Institute (NCI) in the US [9]. These established systems are symptom and safety focussed. Recently the subjective
symptoms from the CTCAEs scale have been adapted for patients to self-report (NCI PRO-CTCAE) [10].

It is now recognised that asking patients to report their own symptoms or AEs via patient reported outcome measures (PROs) can achieve added value by improving symptom control, physician-patient communication and decision making [11-14]. Traditionally this data has been collected in the form of health related quality of life (HRQOL) instruments often referred to as PROs. For trials when only clinician reporting is used it can result in under-reporting of lower grade morbidity and the downgrading of symptom severity [3, 15-19]. This is especially true for subjective AEs [20] which are best captured by patient self report [21]. We also know low inter-rater reliability exists between clinician and patient-reporting and more recently there is increasing emphasis on the importance of incorporating patient-reported toxicity into clinical trials [20]. In further support of PROs Basch and colleagues [3] found that while longitudinal clinician toxicity assessments better predicted significant clinical events, PROs were better at reflecting daily health status. This highlights the complementary nature of both clinician and patient perspectives to gain a holistic view of patient status during treatment and beyond.

In 2008 in the UK the Darzi report [22], recommended PRO data as an essential component of health care evaluation. The United States Food and Drug Administration (FDA) also recognise the importance of incorporating PROs in drug development [23]. Following UK Department of Health (DOH) guidelines [24], England leads with advanced use of PROs in a public health service [25].
In 2009, a review of PROs in prostate oncology practice could not definitively recommend a single tool but suggested piloting the SF-36 and the EQ-5D (European Quality of Life) [26]. A recent review focussing on the quality of PROs reporting in prostate cancer trials (2004-2012) showed that when reported in sufficient methodological detail PROs can be used by policy makers and clinicians to appraise treatment efficacy [27].

Electronic reporting of PROs in health research and practice has risen during the past 20 years for conditions including cancer [28-36] with many using the Internet as a reporting medium. Comparability between electronic and paper reporting is well established [29]. A recent study in prostate cancer patients has shown that completion of the Expanded Prostate Cancer Index Composite (EPIC) at 1 year was superior for electronic reporting (82% compared to 36% using paper forms), thus reducing the risk of bias via missing data[35]. The benefits of electronic reporting have been confirmed in a recent meta-analysis [37].

We have developed a feasible cost-effective model to enable electronic PRO reporting of treatment related AEs from home via the internet linked to the patients’ electronic patient record (EPR) system. eRAPID (Electronic patient self-Reporting of Adverse-events: Patient Information and aDvice) is a web-based electronic patient reporting system including severity linked alerts and /or self-management advice [36]. The system has been successfully used in patients treated systemically for breast, gynaecological and lower gastro-intestinal cancer and is currently being developed to support patients undergoing pelvic RT (including radical prostate cancer treatment) in two large cancer centres in the UK St James's Institute of
Recent reviews of PRO use in gynaecological and prostate cancer found electronic PRO capture and integration with electronic health records to be the most effective method for seamless assimilation into existing care pathways [39, 40].

We aimed to gain a definitive understanding of the type, frequency and severity of AEs associated with modern radical radiotherapy treatments. Further we wanted to identify the instruments used by clinicians and patients to report AEs in the acute treatment phase (defined as during and up to 12 weeks post treatment) and late post-treatment follow-up phase beyond 12 weeks. The specific purpose was to inform the selection of suitable self-reported questionnaires (PROs) for routine monitoring in the eRAPID programme in prostate RT. However, the results of the systematic review would inform a wider clinical audience on the current state of knowledge and standards of reporting acute and late AEs of radical prostate cancer radiotherapy.

This systematic review aimed to establish:

1. Which clinician- and patient-reported instruments or systems are currently used to report acute and late AEs associated with radical radiotherapy treatment for prostate cancer?
2. What is the methodological quality of the studies and quality of PRO reporting?
3. What is the prevalence of the acute and late AEs associated with radical radiation treatment for prostate cancer?
Methods and materials

Search strategy

A systematic literature search was conducted of AEs and clinician and PRO reporting instruments used in RCTs in patients treated with radical radiotherapy for prostate cancer published since 2010. The search was limited to trials reporting from 2010 to incorporate the latest radiotherapy regimes utilising IMRT and IGRT. The protocol is published on the NIHR-PROSPERO website: CRD42014009223. MEDLINE, EMBASE and COCHRANE databases were searched using the Centre for Reviews and Dissemination guidance from April 2010 to August 2014 [41]. For the electronic search strategy used See Appendix 1. Only English language publications were included and all titles and abstracts were screened for inclusion. All documentation followed PRISMA guidelines [42]. In this paper the terms toxicity and adverse events (AEs) are used interchangeably, similarly HRQOL and PROs.

Selection criteria

Patients

The population comprised adult patients with a radical intent management plan including patients having either RT or brachytherapy with or without concomitant hormonal therapy (HT) or chemotherapy. Patients not treated with curative intent and with castration resistant prostate cancer were excluded.

Intervention

We studied the reporting methodology of AEs from the clinician and patient perspective.

Outcomes
Instruments

To address aims 1) and 2) we recorded any clinician and patient reported toxicity/PRO instruments used. Quality assessment of the PRO details was extracted following CONSORT (CONsolidated Standards of Reporting in Trials) guidelines for PRO reporting [43, 44].

Toxicity

To address aim 3) we recorded the type of treatment and the severity of acute and late toxicity associated with radical prostate radiotherapy. We also recorded the number and percentage of patients with AEs and methods of PRO toxicity reporting.

Studies

We selected RCTs as they are considered the gold standard to assess treatment effectiveness and inform clinical decision-making. Toxicity/PRO instruments are more likely to be recorded in an RCT. Meta-analyses, case series, case-reports, case-controlled, editorials, letters, practice guidelines, literature reviews and conference abstracts were excluded.

Data extraction, selection and coding

All RCTs were assessed using criteria from a published checklist, modified to include clinician reported toxicity studies [27]. The principal researcher PH and three reviewers’ (AG, EI, and AA) independently screened the titles and abstracts of all studies retrieved through the electronic search to identify studies that potentially met the inclusion criteria. In cases of disagreement the full articles were discussed to achieve consensus. Both PH (principal researcher) and LS (a second researcher)
independently completed the data extraction on all the chosen articles. Differences were reconciled and consensus reached through discussion with a third reviewer AH.

To assess the quality and data for each RCT, a data extraction form was developed. Extracted information included: study setting, population and participant information, treatment details and basic trial demographics (e.g. first author, year of publication, trial phase, design, funding source). We included detail of the radiation dose and technique, control conditions, study methodology and the follow up duration, timing of toxicity/PRO measurement and clinical demographics (e.g. overall sample size, sample size for PRO reporting, treatment regimes, primary endpoints). We recorded the type of AEs: patient versus clinician reporting, type of toxicity/PRO measures used, grade of toxicity reported. We applied the Cochrane Risk of Bias tool [45] to assess the internal validity of all RCTs.

**Results**

1147 records were screened and 21 papers publishing data from 15 different RCTs were included in the final review (Figure 1) [46-73]. In studies reporting secondary and late effects analyses on data from trials completed before the cut-off date (2010) we report basic trial demographics and assessment of bias from the original trial. Overall the RCTs varied little in the Cochrane Risk of Bias tool assessment and the overall risk was deemed ‘low’.

#insert figure 1#

The trial sizes ranged from 2028 to 41 participants with 80% of the trials enrolling more than 200 patients, the number of participants involved in toxicity follow up
ranged from 1979 to 39. 9/15 (60%) trials were multi-centre and 6/15 (40%) were international collaborations. Primary outcomes included: toxicity 5/15 (33%), freedom from biochemical relapse 5/15 (33%), overall survival 3/15 (20%), local failure 1/15 (6%), and QOL 1/15 (6%). All RCTs were of longitudinal design. (Table1). Out of 5 trials reporting toxicity as the primary outcome 2 showed a significant difference. Of the 5 trials reporting freedom from biochemical relapse 4 reported a significant difference. All 3 trials reporting overall survival as a primary outcome demonstrated a significant benefit, similarly the 1 study reporting local failure. In the trials reporting an improvement in the primary outcome 2/10 (20 %) showed this improvement was at the expense of increased late GI and GU toxicity. Of the remaining 8/10 trials reporting an improvement in primary outcomes no differences in clinician rated toxicity were demonstrated [49, 50, 58, 60, 63-65, 69-72] (Table 1).

#insert table 1#

**Reporting instruments**

We assessed clinician and patient reported instruments for acute and late reporting. Table 2 summarises this information according to QUANTEC recommendations [74]. 48% (10/21) of the papers reported QOL or toxicity using PROs, in 3 papers 14% (3/21) PROs were the focus of the paper [46, 61, 65].

#insert table 2#

**Clinician reporting (acute)**
RTOG was the preferred reporting tool for acute symptoms (N=9), used alone [58, 63, 67, 72] or with other measures e.g. CTCAE [60, 64]. RTOG was also extended to include additional items [49, 75] including erectile function, assessed via an own developed question [57]. An adapted LENT-SOMA scale was also used to include urgency of defecation and micturition [70, 71, 73].

Comparing conventional and hypofractionated deliveries Dearnaley et al [58] predicted and found no difference in acute GU and GI toxicity between treatments despite 1.5% of those receiving 59Gy reporting GI toxicity at grade 4 and acute events occurring earlier in the hypofractionated group. Similarly Archangeli et al [49] found no significant differences between the conventional and hypofractionated arms despite a greater GI incidence of grade 2 toxicity for the latter. Yeoh et al [70] found at 1 month both GI and GU symptom scores were significantly worse in the patients treated with hypofractionated RT compared with the conventional schedule however these differences did not persist (see below). For (IMRT) Michalski and colleagues [64] found significantly less acute grade 1 and ≥ grade 2 combined GI and GU toxicity in the intensity modulated radiotherapy (IMRT) arm compared to conventional 3D conformal RT. Less acute GI ≥ Grade 2 toxicity was found for IMRT but this was not statistically significant. Comparing a dose escalation of 79.2 Gy units with 70.2 Gy Zietman et al [72] found a non-significant increase in acute grade 2 toxicity in the dose escalated arm. Grade 3 toxicities were rare and of similar rates in both arms.

Combining hormone therapy (HT) to 3D conformal RT resulted in no difference in acute toxicity (grade 3 GI 1% and 2% GU in each group) [63]. Similarly more hepatic events were greater in the hormone group but not significantly so.
In terms of GU toxicity resulting from brachytherapy (BT) Crook et al [55] found that introduction of a non-steroid anti-inflammatory drug at either 1 week prior to BT or post-procedure did not reduce catheterisation rates or oedema at 1 or 3 months measured by clinicians using the International Prostate Symptom Score (IPSS) score. The authors concluded that baseline prostate volume remains the primary predictor of post implant urinary retention. For summary see (Table 4).

#Insert table 4# Online supplementary material

**Patient reporting (acute)**

No PRO measures for acute RT treatment toxicities were used, however 9/10 (90%) of papers including PROs collected baseline PRO data. It is especially important as patients often better detect baseline symptoms than clinicians [76] and recording of baseline PRO AEs enables differentiation between pre-existing and treatment related symptoms [77]. Furthermore, collection of baseline data is important to demonstrate similarity in treatment groups, allow subgroup analyses on baseline differences, and analyse whether baseline variables influence patient outcomes [78]. Interestingly a PRO instrument the International Prostate Symptom Scale (IPSS) was used as a clinician reported tool to record baseline data [55].

**Clinician reporting (late)**

All studies reporting late AEs had pre-planned longitudinal design (range 1.5-10.5 years). The preferred instrument was LENT-SOMA, used in 9 studies alone [49, 53, 59, 67] in one it was adapted [70] and 2 studies used it alongside the Royal Marsden Hospital Scale (RMH) [51, 58] and [68] combined LENT-SOMA, RMH with RTOG. The IPSS (clinician rated) [55] was also combined with LENT-SOMA. RTOG was
used alone [56, 63, 72] and combined with CTCAE (V2) [64]. The Dische scoring system was used in one study [62] and an own erectile function assessment tool was reported by Daly [57].

For fractionated vs conventional deliveries GU and GU did not differ at 2 years [58], 3 years [49] and 5 years [67, 70]. Studies where a difference in primary outcome was demonstrated include [49, 70] e.g. freedom from biochemical relapse (FFBR). Subsequent 5 year findings of the non-inferiority trial [79] have confirmed no significant differences in the proportion, cumulative incidence of toxicities between hypofractionated high-dose RT and conventional treatment at 5 years. In terms of dose escalated regimens in studies demonstrating a difference in the primary outcome (FFBR) [53] found significantly more ≥ grade 2 GU and global bladder toxicity and GI were observed in the 80Gy arm (RTOG) at 5 years. Similarly [68] in a follow up from the Medical Research Council (MRC) RT01 Trial reported a statistically significant increase in the 5 year cumulative incidence of GU (rectal bleeding and less commonly diarrhoea and proctitis), but not GU toxicity in the dose-escalation arm (74Gy) versus the conventional (64Gy) arm. However Zietman and colleagues [72] found and no significant differences in late (10 year) GU and GI toxicity between high and low dose regimes. In an exploration of the impact of clinical factors on late toxicity in the MRC RT01 trial Barnett et al [51] found acute GI symptoms were associated with increased late RT toxicity (e.g. proctitis and stool frequency). Acute GU symptoms also predicted late toxicity however when assessing GU function as a change from baseline this effect was negated. Severe toxicities were comparable across treatment groups and looking at PRO data (see table 6) the dose escalation was well tolerated.
For high dose intensity modulated radiotherapy (IMRT) in a preliminary toxicity analysis of the RTOG 0126 trial, Michalski et al. [64] found significant reduction in late GI toxicity (proctitis and bleeding) for IMRT compared to 3D conformal RT. Again acute grade 2 GI toxicity was a predictor of grade 3 late toxicity. Subsequent results have shown [80] greater FFBR at 5 and 10 years for the higher dose (but not survival) however significantly more ≥ grade 2 GI and GU toxicity was demonstrated in the IMRT high dose arm.

Neoadjuvant hormone therapy plus RT, was found to improve overall survival at 2 [63] and 7 years [69] with no differences in late toxicity. Jones et al. [63] found an expected increased grade 1 and 2 GU toxicity (manageable diarrhoea and rectal bleeding) and late side-effects higher than grade 3 were rare in both groups [69, 81]. Hepatic toxicity was greater in the hormone group [63] but not significantly so; however 1% of patients in the HT group only had cardiac toxicity up to and including 2 years later. Hirano et al. [60] found that neoadjuvant ADT plus estramustine phosphate (EMP) and RT afforded greater FFBR for intermediate and high risk patients. Although insufficiently powered for meaningful analysis and severe toxicities were not evident in either group, 45% (nine patients) EMP plus RT patients had grade 1 gynecomastia and 10% developed grade 2 GI toxicity. However the length of neoadjuvant androgen deprivation (4 and 8-month regimes) did not have a bearing on FFBR nor promote differences in toxicity at 7 years.

For erectile dysfunction (ED) Daly et al. [57] (ICORG 97-01 trial) and Beckendorf and colleagues [53] assessed potent patients and found no difference in ED nor survival between arms at 7 and 5 years respectively and conclude that 26% of men on the ADT plus RT arm could expect to retain sexual function at 5 years. Beckendorf et al.
[53] found sexual function to be maintained in 34.5% patients at 70Gy and in 33% who were given the higher dose (80Gy).

For brachytherapy (BT) studies Hoskin et al [62] found patients given EBRT HDR plus (BT) had improved FFBR but this came with ED. Commencement of meloxicam (a non-steroidal anti-inflammatory) 1 week prior to BT or starting it on the evening after the procedure did not reduce catheterization rates or 1-month oedema factor, nor did it improve scores on the International Prostate Symptom Score (IPSS) at 1 and 3 months [55]. The authors conclude that larger baseline prostate volume remains the primary predictor of post-implant urinary retention.

Barnett et al [51] reported GU and GI toxicity as a change from baseline and explored which treatment related and patient factors contributed to late toxicity development post RT at 5 years. Acute GI at 6 weeks was associated with reporting of proctitis, sphincter control and increased stool frequency. Similarly acute GU toxicity predicted late urinary frequency. Inflammatory bowel disease was a predictor of increased reporting of loose stools and increased age was a predictor of increased rectal bleeding and nocturnal urinary frequency. However patients with hypertension were less likely to report poor urinary flow at 5 years. For a summary of clinician reporting see table 5.

#insert table 5# online supplementary material

Patient reporting (late)
Ten publications used PRO for late AEs reporting, 9 collected data longitudinally (range 1.5-10.5 years) and one was cross-sectional [65]. The most popular measures were the Functional assessment of cancer therapy general & prostate modules (FACT-G & FACT-P) and University of California, Los Angeles Prostate Cancer Index (ULCA-PCI) used in 3/10 (30%) and the European organisation for research & treatment of cancer, quality of life questionnaire & prostate module (EORTC QLQ-C30 +PR25) used in 2/10 (20%) of papers. FACT-G and FACT-P were generally used to report a global QOL score and bowel and urinary functioning [61, 68]. In one study separate social, physical, emotional and functional QOL domains were employed along with erectile function from a single FACT-P question [61]. Similarly for EORTC QLQ-C30 +PR25 bowel and urinary toxicity were mainly reported [53, 69] along with a global QOL score [53]. Mostly PRO measures were used in combination; for example the FACT-G & FACT-P was combined with EORTC QLQ-C30 +PR25 [69] and with the UCLA-PCI [68]. The EORTC QLQC-30 +PR25 was combined with the UCLA-PCI to assess bowel toxicity related to dose volume in two studies [51, 59].

Some own-developed PRO measures were utilised: for example a QOL measure developed by researchers from the MOS Sexual Problems (MOS–SP) scale and the Short Form (36) Health Survey (SF-36) [82] was used to assess long term bowel and sexual function [65]. The SAQ was utilised to assess erectile dysfunction [63]. A self-administered version of the SF-36 was also utilised to report both global QOL and domains [46].

#insert table 3#
Out of the 10 PRO papers 4 (40%) reported GI symptoms and 1 (10%) reported GU symptoms. 2 (20%) papers reported maintenance of sexual function. Moving away from specific toxicity the most commonly reported PRO was Global QOL in 7 (70%). For clinician reporting in the acute phase out of 21 papers 6 (28%) reported non-specific GU and 8 (38%) non-specific GI symptoms. 19% of papers reported late erectile dysfunction (ED) including 2 PROs papers.

For late clinician reporting out of 21 papers the most common AEs were: GU 16 (76%) and GI 15 (71%). In 1 paper GU reporting was broken down into cystitis, haematuria and urinary obstruction. 14 (66%) reported non-specific GI symptoms however others were more specific, reporting urgency, bleeding, proctitis, tenesmus and ‘bowel bother’. All grades of symptoms were reported in 10 (48%) of papers and 1 paper (4%) reported haematology issues (table 4). However for late reporting the time scales differed with some studies reporting late toxicity data at 10, 7, 5, 3, and 2 years and some as little as 6 months and a year.

The highest grade of acute and late toxicity was reported in conventional 3D conformal RT treatment groups (+/- hormones) with grade 3 and above GI and GU toxicity [51, 56, 60, 61, 63, 69, 72, 81] with [58] reporting some grade 3 associated with both conventional and hypofractionated intensity modulated RT. However the proportions and prevalence was low. Grade 4 GU [53] GI [68] and erectile dysfunction [57] was also noted (see tables 4 and 5).

Patient reported data are summarised in Table 6. Syndikus et al [68] found mild diarrhoea and abdominal cramps were reported via PROs during and post RT by
both groups. Interestingly moderate to severe diarrhoea was more commonly reported by clinicians in the high dose arm of the trial. No significant difference was found between the treatment arms in terms of mild or moderate rectal urgency (UCLA-PCI) although the (74Gy) arm patients were significantly more likely to report severe symptoms.

Only one third of patients returned the PRO questionnaires, due to organisational flaws in the Beckendorf et al [53] study. However for both of the treatment groups combined GU and GI symptoms measured by the QLQC-30 PR25 were increased from baseline by 4 and 3 points respectively from baseline at 3 years. In the 40 patients who answered the question, libido was unchanged however sexual functioning had decreased by 19 points in both groups (combined data). In terms of sexual function Beckendorf et al [53] found 34.5% of 55 potent patients (70Gy) and in 33%of 48 pts at 80Gy were able to maintain sexual function with no significant different overall. Overall there was significant increase in bladder toxicity at 80Gy but no difference in QOL scores. Warde et al [69] by contrast got good compliance with completion of PRO measures and reported the mean score change from baseline scores from the EORTC QLQC30 and the PR13 (prostate-cancer module) and the FACT-P at 6 and 36 months. Only physical function showed a clinical significant difference of more than 10 points. GI measures were captured only by the EORTC PR-13, revealing short-term (6 month) but not long-term (36 month) differences between groups. Overall QOL and physical function scores showed a general deterioration of physical function in both groups, consistent with ADT suppression. No long term differences at 36 months between groups in PRO
recorded toxicity or QOL, although the FACT-P scores showed a small increase in GI symptom scores at 6 months in patients in the RT and ADT group.

Barnett et al [51] reported PRO data from the MRC RT01 trial [52] faecal urgency and loose stools were reported (as a change form baseline) using the UCLA-PCI to establish the contribution to late effects at 5 years. Findings showed that existing inflammatory bowel disease was significantly associated with the toxicity end point loose stools (but not faecal urgency). In terms of dosimetry there were significant associations between faecal urgency and dose volume of 70Gy and 65Gy respectively. In an earlier dose volume analysis of this trial [59] the UCLA-PCI was used to map the relationship between dose volume and toxicity. The authors stress the importance of adjusting for pre-existing symptoms reported prior to RT as the effects can mimic the effects of pelvic RT. Nguyen et al [65] explored the relationship between rectal dose-volume histogram parameters and QOL, using an own-developed PRO measure the Massachusetts General Hospital scale. This constituted a composite bowel score of 0 (minimum dysfunction) to 100 (maximum dysfunction) [82]. A trend was found towards associations between increased long-term PRO GI dysfunction and dose volume. This was especially true of patients who had higher RT volume than the median value of V60, V70, or V75 for the anterior rectal wall. In terms of V60 this was significantly different.

RT plus Brachytherapy (BT) afforded greater FFBR over a 10.5 year follow-up and no difference in FACT-G, FACT-P or Trial Outcome Index (TOI) scores was seen between the RT and RT plus BT arms of the trial Hoskin et al [61]. Mean erectile function scores were similar in both arms but both were lower than baseline. For
both treatments on the FACT QOL domains there was no evidence that QOL deteriorated over time.

Jones et al [63] found a survival benefit for the addition of short term ADT for men and assessed erectile function with just one question “When sexually excited, are you able to get an erection?” with five response options. The respective rates at 1 year were 21% for RT alone and 31% for RT and HT. Scores at 1 year, (compared with pre-treatment) were improved in 9% of the patients, remained the same in 33%, and worse in 58%. However there were no significant differences between the groups.

At 3-year follow-up, using the Short form QOL questionnaire (SF-36) summary scores for physical and mental health showed no significant differences Al-Mamgani [46]. A statistically significant deterioration in QOL scores over time was registered in both arms in six scales including mental, physical and general health, bodily pain and vitality (these were more pronounced in the 78Gy than the 68Gy arm). However, only role physical (Physical functioning role limitations due to physical health) and physical functioning showed a clinically relevant deterioration.

PRO CONSORT assessment

Out of the 21 papers, 10 (48%) reported PROs and 5 identified PROs as an outcome measure [46, 51, 61, 65, 68]. The remaining 4 specified PROs in the body of the text. A background, rationale and PRO hypothesis were mentioned in 6 [46, 51, 53, 61, 65, 68]. All papers evidenced the reliability and validity of the PRO instruments.
Methods of PRO data collection and effect sizes were not specified in any study. The sample size for PRO analysis was stipulated in all studies, but only 2 described statistical approaches for missing data [46, 61]. The transparency of PRO data reporting over time was evident in 8 studies [46, 51, 53, 61, 63, 67, 68]. All studies stipulated each primary and secondary outcome and recorded multidimensional PRO results at each time point. PROs were discussed in relation to clinical practice and outcomes but only 3 papers discussed the PRO limitations [46, 61, 63] (table 3).

Discussion

This review determined the commonly used methods for capturing acute and late AEs for patients undergoing radical prostate cancer radiation therapy from 21 papers reporting from 15 high quality RCTs enrolling 9,040 patients. A key finding was that acute AEs were described only using clinician reported instruments (with RTOG most used) without any patient reporting. This practice is at odds with recent movement towards including patient voices in AEs reporting and patient engagement in clinical research. We would recommend including PRO to evaluate radical radiotherapy before, during and after the treatment to fully capture patient experiences, and support the development of predictive models for late effects based on the severity of early toxicity. For description of late effects LENT-SOMA was the preferred clinician reporting measure (now integrated into CTCAE v 4.0) [9]. Almost half of the trials used PROs to report late toxicity and used a PRO to record baseline data. A range of PRO instruments was used, but no single preferred instrument emerged. To ensure coverage of AEs existing instruments (both clinician and patient reported) were occasionally augmented with additional items on sexual functioning and anorectal symptoms. The International Consortium for Health
Outcomes Measurement (ICHOM) is working towards creating a standardised ‘set’ of PRO measures for clinical practice to ensure comparability across countries [83]. Similarly using the Core Outcome Measures for Effectiveness Trials (COMET) model Maclennan and colleagues [84] are developing a standard set for clinical trials. Interestingly although the ICHOM recommend the Expanded Prostate Cancer Index Composite EPIC (short form) to capture late effects of localised prostate cancer EPIC was not used in any of the RCT we reviewed. However the full EPIC questionnaire was used in the 2 and 5 year follow up in the CHHiP trial [79, 85] replacing UCLA-PCI, SF-36 and FACT-P. Ideally, similar PRO instruments should be employed in both clinical trials and clinical practice, and hopefully these recent international efforts will be a step in this direction.

An overall generic/global QOL was most frequently reported in PRO papers followed equally by bowel and urinary AEs. Generally when QOL was the focus of the paper PRO specific subscales and domains were reported. Bowel and urinary toxicity were the most frequently clinician-reported symptoms. Interestingly few papers reported erectile dysfunction despite this being an inevitable consequence of treatment and a cause of significant patient distress.

The acute clinician reported symptoms were generally grouped into “Genitourinary” or “Gastrointestinal symptoms” with no detailed differentiation. However late AEs were more likely to be specified e.g. gastrointestinal (urgency, bleeding, proctitis, tenesmus) and genitourinary (cystitis, retention, haematuria, bladder side effects). Generally all grades of toxicity were recorded however 3 papers (14%) only recorded grade 2 or above toxicity. The most common grades reported for both (GU and GI)
toxicity, were grade 2 or above. Interestingly in 7 (33%) of papers reporting all
grades of toxicity ≥ grade 2 toxicity was used for analysis purposes. Some collapsed
AEs into moderate/ severe or combined numerical grades for display purposes.

These findings are similar to those of a recent review of toxicity reporting in RCT in
rectal cancer where 55% of clinician-reported papers grouped toxicity data related to
an organ system (e.g. bowel) [86]. An approach focusing on serious AEs and
grouping together a range of organ symptoms, may underestimate the impact of low
grade persistent toxicity and fail to provide patients with specific information on
expected side effects. Further, we must highlight that the different classification
systems used (RTOG vs. CTCAE) will have influenced the variability in AE
prevalence recorded. We would expect greater consistency in future trials after
successful adoption of the CTCAE.

Assessing the standards of PRO reporting using PRO Consort guidelines [43, 44] it
was encouraging that the majority of studies provided evidence of the reliability and
validity of PROs. Generally PROs were identified as primary and secondary
outcomes, discussed in relation to clinical outcomes and practice and a rationale for
PRO use was included in 50% of the studies. However improving the clarity and
transparency of PRO collection and reporting would be welcomed. The methods of
PRO data collection were absent in many papers and for analysis effect sizes and
strategies for dealing with missing data were generally absent. However we would
expect these standards to improve in the future as more researchers recognise the
value of PROs adopt the PRO Consort guidelines when planning trials.
On the whole improved patient outcomes were not at the cost of increased toxicity although some conventional treatments +/- HT and at high does did demonstrate increased GU [53] and both GU and GI toxicity [51, 80]. Generally the newer RT treatment regimens (e.g. IMRT and hypofractionated delivery) were well tolerated with no meaningful differences between groups. Indeed the most recently published 5 year PRO analysis from the CHHiP trial [79] found the incidence of patient-reported bowel symptoms was low and similar between hypofractionated RT delivery and conventional fractionated RT control group 5 years after radiotherapy.

However, evaluating treatment effects on AE can be problematic if different grading criteria are employed and if late toxicity is recorded at different time points. Therefore working towards a standardised model of assessment and documentation during follow-up would seem a desirable goal.

Strengths of this review include the identification of both PRO and clinician measures to report AEs during and beyond treatment. Furthermore the review determined the type of AEs associated with radical prostate treatments with some of the newer treatment regimens. The review also assessed the methodological quality of the trials and the quality of PRO reporting adding to the valuable previous work from the EORTC group [27]. A limitation of the review is that due to the volume of prostate publications relating to new treatment regimens since completion of our review we were unable to include RCT published from 2014. Recent excellent reviews of RCT focussing on survival outcomes, toxicity and economic impact of hypo-fractionated treatment delivery [87, 88] were beyond the scope of our study. Nevertheless our summary of toxicity reporting by treatment regimen will be of interest to clinicians.
We also recognise that our findings are in a sample of trials planned and conducted mainly before the Darzi [22] and FDA recommendations [23] advocating the use of PROs. Indeed it would be useful to look at how the recommendations have influenced the use of PROs in clinical trials from 2014 onwards. We would recommend a second review of RCT up to and including 2019 to assess the influence of these recommendations on PRO reporting. However questions remain why in this review PROs were adopted to monitor late and not acute effects.

Conclusions

The finding that trials reporting acute AEs did not utilise PROs and the use of additional items to ensure coverage of late AEs suggests that considerable work needs to be done to develop items for patients undergoing radical prostate cancer radiation treatment to self-report in the acute phase and immediately beyond. To assess coverage of PROs with regard to identified AEs during radical prostate treatment future work will focus on mapping potential items from PROs identified here (e.g. ULCA-PCI, FACT-G & P, EORTC QLQC-30 +PR25) along with others (e.g. EPIC and male pelvic questionnaire [89], PRO-CTCAE [10] to assess coverage. To augment the questionnaire if required we will use consensus methodology to select appropriate items with clinical team members and patients. We agree that affording equal rating to patients and clinicians in the consensus process is imperative [90].

We welcome the initiatives to develop a core set of outcomes in clinical practice [83] and trials [84] and we hope our findings will be of interest and support this work. We would welcome the opportunity to contribute to the international consensus process.
required to develop a core set of PRO outcomes for the acute phase of treatment and immediately beyond.
References


## Appendix 1 Search strategy and key words used in searching of three key databases

<table>
<thead>
<tr>
<th>EMBASE and MEDLINE</th>
<th>COCHRANE TRIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td># 1. (Prost* adj2 (cancer* or carcinoma* or neoplasm* or tumo?r*)) .mp.</td>
<td># 1 prostate carcinoma</td>
</tr>
<tr>
<td># 2. Radiotherap* 8. EBRT .mp.</td>
<td># 2 prostate cancer</td>
</tr>
<tr>
<td># 3. Brachytherap*.mp.</td>
<td># 3 (Prost* adj2 (cancer* or carcinoma* or neoplasm* or tumo?r*))</td>
</tr>
<tr>
<td># 4. Hormon* therap*.mp.</td>
<td># 4 Prostatic neoplasms</td>
</tr>
<tr>
<td># 5. androgen therap*.mp.</td>
<td># 5 or #2 or #3 or #4</td>
</tr>
<tr>
<td># 6. chemotherapy*.mp.</td>
<td># 6 Radiotherapy</td>
</tr>
<tr>
<td># 7. radiation.mp.</td>
<td># 7EBRT</td>
</tr>
<tr>
<td># 8. chemoradiation.mp.</td>
<td># 8 Brachytherap*</td>
</tr>
<tr>
<td># 9. chemoradiotherap*.mp.</td>
<td># 9 Hormone therap*</td>
</tr>
<tr>
<td># 10. androgen antagonists.mp.</td>
<td># 10 androgen therap*</td>
</tr>
<tr>
<td># 11. androgen receptor agonists.mp.</td>
<td># 11chemotherap*</td>
</tr>
<tr>
<td># 12. androgen*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]</td>
<td># 12radiation</td>
</tr>
<tr>
<td># 13. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12</td>
<td># 13radiotherap*</td>
</tr>
<tr>
<td># 14. toxicit* criteria.mp.</td>
<td># 14brachytherap*</td>
</tr>
<tr>
<td># 15. toxicity* grading.mp.</td>
<td># 15hormon*</td>
</tr>
<tr>
<td># 16. Acute toxicit* .mp.</td>
<td># 16 chemoradiation</td>
</tr>
<tr>
<td># 17. late effect*.mp.</td>
<td># 17chemoradiotherap*</td>
</tr>
<tr>
<td># 18. adverse event*.mp.</td>
<td># 18MeSH descriptor: [Androgen Antagonists] explode all trees</td>
</tr>
<tr>
<td># 19. adverse event.mp.</td>
<td># 19MeSH descriptor: [Androgens] explode all trees</td>
</tr>
<tr>
<td># 20. HRQol.mp.</td>
<td># 20 MeSH descriptor: [Chemoradiotherapy] explode all trees</td>
</tr>
<tr>
<td># 21. QOL.mp.</td>
<td>#21 #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20</td>
</tr>
<tr>
<td># 22. PROs.mp.</td>
<td># 22 toxicity criteria</td>
</tr>
<tr>
<td># 23. PROMs.mp.</td>
<td># 23 toxicity grading</td>
</tr>
<tr>
<td># 24. CTCAE.mp.</td>
<td># 24 acute toxicit*</td>
</tr>
<tr>
<td># 25. RTOG.mp.</td>
<td># 25 late effect*</td>
</tr>
<tr>
<td># 26. RTOG late morbidity scores.mp.</td>
<td># 26 adverse event*</td>
</tr>
<tr>
<td># 27. EPCI.mp.</td>
<td># 27 HRQol</td>
</tr>
<tr>
<td># 28. FACT-P.mp.</td>
<td># 28 QOL</td>
</tr>
<tr>
<td># 29. International index of erectile function questionnaire.mp.</td>
<td># 29 PROs</td>
</tr>
<tr>
<td># 30. International prostate symptom score.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]</td>
<td># 30 PROMs</td>
</tr>
<tr>
<td># 31. 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30</td>
<td># 31 CTCAE</td>
</tr>
<tr>
<td># 32. randomised controlled trial.mp.</td>
<td># 32 RTOG</td>
</tr>
<tr>
<td># 33. RCT.mp</td>
<td># 33 EPCI</td>
</tr>
<tr>
<td># 34 random?ed controlled trial.mp</td>
<td># 34FACT-P</td>
</tr>
<tr>
<td># 35 controlled clinical trial.mp.</td>
<td># 35 International index of erectile function questionnaire</td>
</tr>
<tr>
<td># 36. randomi?ed.mp.</td>
<td># 36 international prostate symptom score</td>
</tr>
<tr>
<td># 37. randomized controlled trial.mp.</td>
<td>#37 #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36</td>
</tr>
<tr>
<td># 38. placebo.mp.</td>
<td>#37 #5 and #21 and #37</td>
</tr>
<tr>
<td># 39. drug therapy.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]</td>
<td>Publication Date from 2010 to 2014</td>
</tr>
<tr>
<td># 40. 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40</td>
<td></td>
</tr>
<tr>
<td># 41. (animals not humans).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]</td>
<td></td>
</tr>
<tr>
<td># 42. (#40 not #41).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]</td>
<td></td>
</tr>
<tr>
<td># 43. 13 and 31 and 40</td>
<td></td>
</tr>
<tr>
<td># 44. limit 43 to yr=&quot;2010 -Current&quot;</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: PRISMA Flow Diagram

Records identified through database searching
- Medline: 355
- Embase: 545

Additional records identified through other sources (n = 2)

Records after duplicates removed (n = 1070)

Records screened by 2 researchers (n = 1070)

Records excluded (n = 974)
- Full-text articles excluded, with reasons (n = 75)
  - Non radiotherapy (39%)
  - Metastatic disease (9%)
  - No toxicity reported (7%)
  - Secondary treatment (26%)
  - Not RCT (11%)
  - Surgical (1%)

Full-text articles assessed for eligibility (n = 96)

Studies included in qualitative synthesis (n = 21)