Published in final edited form as:
Int J Radiat Oncol Biol Phys. 2013 October 1; 87(2): 246–260. doi:10.1016/j.ijrobp.2013.03.036.

Does Quality of Radiotherapy Predict Outcomes of Multicentre Cooperative Group Trials? A Literature Review

Alysa Fairchild, MD, FRCPC[1,*], **William Straube, MSc**[2], **Fran Laurie, BSc**[3], and **David Followill, PhD**[4]

[1]Department of Radiation Oncology, Cross Cancer Institute, Edmonton, AB, Canada

[2]Advanced Technology Consortium, Imaged-guided Therapy QA Centre, St Louis, MO, USA

[3]Quality Assurance Review Centre, Lincoln, RI, USA

[4]Radiological Physics Centre, University of Texas MD Anderson Cancer Centre, Houston, TX, USA

Abstract

Central review of radiotherapy (RT) delivery within multicentre clinical trials was initiated in the early 1970's in the USA. Early quality assurance (QA) publications often focused on metrics related to process, logistics and timing. Our objective was to review the available evidence supporting correlation of RT quality with clinical outcomes within cooperative group trials. Medline search was performed to identify multicentre studies which described central subjective assessment of RT protocol compliance (quality). Data abstracted included method of central review, definition of deviations, and clinical outcomes. Seventeen multicentre studies (1980–2012) were identified, plus one Patterns of Care Study. Disease sites were hematologic, head and neck, lung, breast and pancreas. Between 0% and 97% of treatment plans received an overall grade of acceptable. In seven trials, failure rates were significantly higher after inadequate versus adequate RT. 5/9 and 2/5 trials reported significantly worse overall and progression-free survival after poor quality RT, respectively. One reported a significant correlation and two reported non-significant trends towards increased toxicity with non-compliant RT. Although more data are required, protocol-compliant RT may decrease failure rates and increase overall survival and likely contributes to the ability of collected data to answer the central trial question.

Keywords

quality assurance; clinical outcomes; trial design; cooperative group

[*]Corresponding Author: Dr A Fairchild, Dept of Radiation Oncology, Cross Cancer Institute, 11560 University Avenue, Edmonton, AB T6G 1Z2, Canada, alysa.fairchild@albertahealthservices.ca, P: +1 780 432 8516, F: +1 780 432 8380.

Conflict of Interest notification: There are no conflicts of interest to declare.

Presented in part at the International Conference on Translational Research, 27 Feb-2 Mar 2012, Geneva, CH.

Introduction

Central review of radiotherapy (RT) delivery within multicentre clinical trials was initiated in the early 1970's in the USA. The initial purpose was to increase the number of patients accrued to studies who were ultimately evaluable as specified by each protocol (1). Early quality assurance (QA) performed by the CALGB and other clinical trial groups revealed non-uniformity of treatment strategies and suggested a need for monitoring, particularly when participants were asked to employ a regimen that differed from their standard technique (2). Aside from the suggestion of increased patient evaluability with central QA office intervention, initial publications focused largely on metrics related to process, logistics, and timing (3).

QA is the systematic effort to monitor performance, to compare quality to a predefined standard, and to implement corrective measures if performance does not meet requirements within a certain margin (4). QA in radiotherapy (RT) includes procedures that ensure the consistent and safe fulfillment of an RT prescription with regard to dose to the target and normal tissues, minimization of exposure of personnel, and patient monitoring aimed at determining the results of treatment (5). A QA program should define the range of acceptable deviations, detect potential causes, and develop mechanisms of action for correction and prevention of deviations. Hence QA in RT must oversee each step from patient simulation to treatment planning, to beam production and patient follow-up (6).

According to van Tienhoven et al, trial-specific QA procedures should be: feasible; capable of quantifying variations in treatment parameters; able to detect protocol variations early in the course of a trial; able to contribute to correction of deviations; and finally, able to demonstrate an impact on the final outcome of the study (7). Central patient-specific RT plan review, also known as "rapid review", instituted by many cooperative clinical trial groups and QA centres, is suitable to evaluate both patient- and tumour-related eligibility as well as application of protocol RT (8–9).

The objective of this work was to review available evidence for correlation of RT quality with clinical outcomes within multicentre cooperative group clinical trials.

Methods

Medline search was performed to identify candidate multicenter studies with no restrictions on date of publication but restricted to the English language (Appendix 1). Eligibility criteria included multicentre trials which accrued adults only, were published in full, and led by any cooperative clinical trial study group. Included studies described central subjective and/or objective assessment of external beam RT protocol compliance (quality) and reported correlations between RT quality and clinical outcomes. Additional studies were identified from reference lists of retrieved papers and review articles. Data were abstracted from QA publications and, where necessary, companion clinical publications, concerning the methodology of central evaluation, deviation rates and clinical outcomes. Correlation of compliance of RT delivered per-protocol with response rates (RR), locoregional or distant failure, progression-free survival (PFS), overall survival (OS), and toxicity was analyzed. Although source publication terminology differed, "deviation" has been used throughout this analysis for consistency.

Results

General

Seventeen multicentre trials described in 16 articles published over three decades (1980–2012) and one Patterns of Care Study (PCS) were identified (2–3,10–24)(Table 1). In addition to the PCS, there were eight cooperative trial groups represented, based in North America, Europe and Australia. Disease sites were hematologic (13, 17–21), head and neck (3,14–15,23), lung (10–12), breast (2,16) and pancreas (22,24). The central trial question was RT-related in ten trials, and in 11, chemotherapy was to have been delivered to at least half of patients accrued (induction chemotherapy [17,19–21]; induction + adjuvant [10,12]; concurrent [3,13,15]; concurrent + maintenance [22]; and induction + concurrent + adjuvant [24]). The trial's primary endpoint was negative in 16/18 trials and either unknown or not applicable in the remainder (13,23). The maximum number of institutions participating in central review was 212 (19). Reported median follow-up for patients analysed within the QA publications ranged from 2.8 – 8.1 years (N=10 trials).3/18

Methodology of Central Review

The methodology and level of detail of each RT plan review procedure varied (Table 2). In 16/18 trials, all patients receiving RT were to have been reviewed. In the PCS, 181 patient records were randomly chosen from five facilities with a large experience in Hodgkin lymphoma. In German Hodgkin's Study Group (GHSG) HD8, the cohort was not explictly described. The proportion of plans evaluable, however, varied from <50% to >95% of patients who actually received RT. In five trials, central assessment of RT quality was reported as blinded to outcomes. In 13/18 trials, it was performed only after RT completion; in the remaining five, evaluation took place both prior to RT start and retrospectively, including three GHSG trials where RT plans were designed prospectively by a central reference panel.

The definition of a deviation varied significantly between groups, as well as over time within the same group (e.g. adequate versus inadequate; acceptable, minor or major deviation)(Table 3). What constituted a deviation was defined *a priori* in five trials (2,14,15,24). Overall protocol deviation grades assigned for RT quality (Table 4) could be reconstructed in 16/18 studies. Between 0% and 97% of treatment plans reviewed were graded as acceptable. Nine trials assigned an overall case grade of major deviation which was received by up to 47% of plans. Analysis of the correlation of RT quality with outcomes was controlled for other prognostic factors in approximately half of the protocols (2,3,12,13,15,18, 21–22,24).

Outcomes

RR with respect to RT quality were evaluated in two trials and no statistical correlation was observed in either, although in one (12), the complete response (CR) rate was higher for those cases with a minor or no deviation (44% [38/85]) as compared to cases scored as a major deviation (39% [16/41]; p=0.68). In RTOG 7301, response rates (55–59%) were comparable between patients with RT plans with a major deviation at the primary site versus those without (11).

Fourteen trials reported failure rates (locoregional +/− distant) and in seven, these were significantly higher when RT was judged to be inadequate (3,10,13,15,18,23)(Table 5). In another two, failure rates were higher, but not significantly so, after inadequate RT, and in the remaining five trials, there were no significant differences in failure rates reported. PFS was examined in five trials, of which two reported significantly worse PFS in patients with poor quality RT (3,18). Nine trials examined the correlation of adequacy of RT with OS;

five suggested that compliant RT significantly increased OS (3,12,15,24). The TROG 0202 trial, for example, reported a 20% improvement in OS between patients with major deficiencies versus those whose RT was initially protocol compliant.

Reports from three studies evaluated toxicity in relation to RT quality. In RTOG 0411, there was a significant correlation between major deviations and the incidence of grade three or higher gastrointestinal toxicity during both the concurrent chemoRT (45% vs 18%; [p=0.05]) and maintenance chemotherapy phases (45% vs 13%; [p=0.01])(22). Two additional trials described non-significant trends towards increased toxicity with non-compliant RT (14,24). In RTOG 7102, there were 22.2% (10/45) serious complications with protocol-required three field treatment vs 9.8% (8/82) with two fields (p=NS). Additionally, 36.3% (4/11) of patients treated with non-protocol source-to-skin distances (SSD) sustained serious soft tissue complications, compared to 10.8% (14/130) treated with the correct SSD (no p value reported). In RTOG 9704, there were no significant differences in grade three or higher acute toxicities in the 5FU arm based on RT plan QA score, but in the gemcitabine arm, there was a trend towards increased toxicity for non-protocol compliant RT plans for both hematologic (p=0.08) and non-hematologic (p=0.06) grade four toxicity.

Discussion

Despite concerted efforts by cooperative clinical trial groups to ensure uniformity of protocol radiotherapy, variation in its administration persists (24). Potential reasons for the inability to comply with protocol RT are listed in Table 6. Although the trials reviewed reveal varying degrees of compliance, it is of course not known how these results compare with other studies where less comprehensive or no QA took place (25). The validity of conclusions of past studies may be called into question depending on the rigor of central QA performed (24,26) since the magnitude of the detrimental effect of non-compliant RT can be larger, in some cases significantly, than the anticipated benefit of the interventions studied (3,24). Understanding the incidence, types and reasons for deviations contributes significantly to our understanding of the application and limitations of RT on and off-study (27,28).

There was significant variation in the reported proportions of randomized, eligible and treated patients with RT plans evaluable for central review. This observed variation impacts cross-trial comparisons and generalizability of conclusions. If RT plans are not systematically reviewed in a consistent manner, bias may be introduced in the number and type of deviations seen. For example, in the EORTC 22991 trial, retrospective central plan review was conducted only for major participating sites, which due to institutional experience and available resources probably positively affected the results (9). In trials reviewed here, plans may not have been available or sites may have been non-compliant in submitting data at the appropriate time. Patients were taken off study prior to starting RT (12), and excluded if they did not complete all (10) or a certain proportion of planned RT (e.g. ≥85% of prescribed dose [3]). In NSABP B04, patients who started RT greater than six weeks after surgery were excluded from the quality analysis (2). Those with gross geographic misses or other unacceptable variations were considered inevaluable by SECSG76 investigators (10). In the combined RTOG 7913/7915 analysis, patients with less than 90 days of follow-up and those with major deviations were excluded from assessment of RT protocol compliance (15). Entire sites have been removed from quality analysis in past publications if they did not accrue a minimum number of patients. The exclusion of these patients before they could ever be evaluated may in fact dilute published results.

There were also significant differences in central review methodology for the trials included, which represent progression from the paper/hard copy film to the digital era. Some groups

instituted general QA such as site visits, machine calibration checks and baseline credentialing. The specific components of RT plans actually reviewed (CT simulation films, port films, contours, DVHs, isodose curves) also varied, as well as whether case report forms and baseline diagnostic imaging was concurrently evaluated. Dosimetric recalculation was reported by five trials. The composition of the review committee could be the study chair alone, an independent panel, or a cooperative group QA office. In the GHSG protocols, a panel designed each patient's RT plan and then evaluated its subsequent delivery by the local site. Central analysis of RT was performed retrospectively and/or prospectively, impacting whether the prescribing physician had an opportunity to modify the plan prior to completion of treatment. Review may or may not have been blinded to clinical outcomes such as local recurrence. This spectrum of QA procedures further increases the uncertainty in the comparability of correlation of outcomes between trials. Efforts to harmonize QA practices with the formation of the Advanced Technology Consortium in the USA and the Global Harmonization of RT QA in Clinical Trials group are underway (29).

Trials evaluated also represent significant evolution in treatment delivery capability. Although it is apparent that historical findings regarding RT quality cannot be directly extrapolated to current practice and technology, they have been included in this review for several reasons. They impact the confidence with which we now interpret the results of those (and contemporary) trials, some of which still influence practice today. They confirm quantitatively what has always been intuitively believed to be true; for example, that geographic miss is of major concern, and they have contributed to the evolution of optimal treatment in many disease sites. Similar to the ongoing *clinical* reliance on many trials with outdated RT, until there is a body of evidence directly applicable to today's advanced techniques, in order to make informed decisions on most appropriate utilization of resources for QA, this collection of historical and current data must suffice.

There are several potential reasons why the quality of RT was not found to correlate with clinical outcomes in some trials reviewed. It is not likely that all assessed parameters (eg target volume, normal tissue dose constraints, beam arrangement) have an equal probability of affecting clinical outcomes, and yet in most trial reports, different deviations are weighted equally. The required degree of compliance to avoid assignment of deviations may be too strict or not strict enough (16,23,30). In the EORTC 22922 trial, variation in planning methods between centres and individualization of RT to patient anatomy were not considered deviations as long as planning aims were achieved (31), but this leniency has not been commonly permitted by other clinical trial groups. RT treatment quality may not be perfect, but may not correlate with outcomes if it in fact meets minimal criteria for acceptability; alternatively, it may be so inadequate that the degree to which it is deficient does not materially affect outcomes. Additionally, quality may not appear to predict outcomes if an insensitive endpoint is chosen for analysis or if there has been insufficient follow-up (16).

Historically, the impact of poor quality RT treatment was presumed to be addressed by the process of randomization and/or washed out by the use of systemic therapy. However, some trials do report differences in distribution of deviations by treatment arm. In terms of the impact of chemotherapy on protocol deviations, we have found no obvious trends. In HD7, more patients received the protocol deviation "too little" RT after chemotherapy ($p=0.035$) (21). Concurrent chemotherapy with RT in 26 patients in the PCS study eliminated the adverse effect of inadequate margins with no significant increase in recurrence (13). In the seminal TROG 0202 protocol, Peters et al describe the possibility that the higher cisplatin dose in the control arm could have compensated to some degree for poor RT (3). Since chemotherapy was introduced into early stage GHSG trials, no significant influence of RT quality on freedom from treatment failure has been evident (32). Chemotherapy delivery has

rarely been rigorously monitored, and has usually been accepted at face value as delivered per protocol without significant variations. Administration of chemotherapy, however, should not allow what constitutes acceptable RT to be less rigorously defined (10,13).

RT quality may correlate with clinical outcomes without being causative. There may be a differential rate of toxic deaths by treatment arm which may be confounded by the effects of other modalities. Some primary sites and stages are more difficult to plan and treat; the possible impact of primary and nodal disease stage was investigated in TROG 0202 with no significant differences found (3). Tumour location, size, and number of positive nodes also failed to explain deviation rates in NSABP B04 (2). While potentially explanatory correlations between patient/disease characteristics and protocol variations should be ruled out before ascribing differences in outcomes to RT quality (11), only half of trials reviewed reported multivariate analysis. In one of these, RTOG 9704, the authors report an inability to find other possible explanations (24). This lack of apparent correlation may also relate to the inability to fully blind RT studies. In an early GHSG trial, there was a difference between treatment arms, with five versus 13 volume-related protocol deviations, leading the authors to suggest that volumes may have been more generously designed in the setting of lower prescribed doses (33).

The impact of institutional experience in terms of number of patients randomized and treated has been extensively reported. It is likely that a learning curve exists (23) such that the more patients accrued to a trial, the better the quality of later submissions. This improved quality may be a result of increased communication and training related to QA mechanisms implemented for the trial. In TROG 0202, the proportion of RT plans with predicted adverse impact on outcome varied significantly by both country and number of patients accrued (<5 versus 20 or greater)(3). For the HD4 trial, the proportion of patients with a protocol deviation was significantly lower in centres that contributed ≥10 cases ($p=0.005$) but type of centre was not predictive (18). Sites which entered >30 patients demonstrated much lower rates of protocol deviations compared to those with a lower accrual rate in both NSABP B04 (2) and GBSG I (16). However, no correlation with QA score was seen by number of patients enrolled per institution in RTOG 9704 (24).

Other factors contributing to institutional experience include the duration of participation on a particular study and receipt (and application) of feedback from central QA review. Early in the NSABP B04 study, more than 20% of RT plans had a major protocol variation, but as the study progressed, this decreased to <10% and then 0% in the last 6 months of accrual (2). In EORTC 22922, which performed central plan evaluation of subsequent patients from the same institution, target volume delineation was missing in 67% in the first round of review, decreasing to 47% subsequently (34). Continuing medical education (semi-annual meetings with case reviews) and quality control screening improved RT quality during the GBSG I study. During the initial period (1983–1985) only 1/3 of RT was protocol-compliant, increasing to 50% between 1987–1989 (16).

A subset of the trials included in this study has been recently reviewed by authors from the EORTC, who concluded that results of nine prospective phase II/III studies (1994–2012) support the compulsory imposition of credentialing in RT trials since failure to adhere to protocol treatment reduces OS, local control and potentially increases toxicity (35). The challenges associated with QA programs within cooperative group trials were described, along with reasons for resistance to participation and implementation of recommendations. Additionally, a meta-analysis of study-level data from eight cooperative group clinical trials, encompassing both adult and pediatric patients, reported an approximate 75% increase in both overall mortality and aggregated secondary endpoints of local control and event-free survival associated with RT deviations, based on unadjusted hazard ratios (36). The authors

found no evidence of publication bias although conceded that the ability to detect it may have been limited by sample size. The effects of minor versus major deviations were not differentiated, and studies evaluated were primarily representative of historical treatment planning methods: 6/8 used 2D treatment planning and no protocols using IMRT were included (36). The focus of this review is on differences in methodology, reporting and analysis to discern specific factors influencing published QA results. Identification of potential predictors allows interpretation of reported correlations with outcomes in the proper context, and in relation to potential sources of bias. We agree that non-protocol compliant RT in multicentre trials may waste time, effort and money (35), but arrive at more cautious conclusions about the impact on clinical outcomes based on broader literature review.

The variation in definitions of deviations discovered is noteworthy: they may be described quantitatively or qualitatively, may be related to target volumes or normal tissues, or may reflect appropriateness of contours, dosimetry or logistics. Definitions and frequencies of protocol deviations must be interpreted in the context of the technological era from which they originate. They reflect not only what was felt to be important for quality RT and presumably clinical outcomes at the time the trials were being designed, but also the degree of sophistication of the tools available for central RT plan assessment. The specificity with which protocol compliance can be evaluated has evolved concomitantly with advances in treatment planning, RT delivery, digital data transfer and ability to perform geographically distributed review. Similar to the fact that past findings regarding RT quality cannot be directly extrapolated to current practice and technology, historical absolute deviation rates probably cannot be directly compared to those derived from modern publications.

There are many questions remaining regarding how deviations should be defined; indeed, this aspect of multicentre clinical trial QA could benefit urgently from harmonization. Based on the data currently available, making specific recommendations for quantitative criteria would be premature. What constitutes a major versus minor deviation must first be standardized to the extent possible, since a ‘one-size-fits-all’ approach may not apply to all disease sites, RT techniques or settings. For example, an overall treatment time exceeding 10% of protocol specifications may reasonably be considered a major deviation for all external beam treatment of unresectable head and neck cancer, but would this also apply to stereotactic body RT for inoperable early stage lung cancer? The strictness with which they are described should probably depend on whether the central question of the trial is RT-related, especially if the main issue is one of radiation dose (7). Definition should also focus on the likelihood of an associated detrimental impact on clinical outcomes without adversely affecting required patient accrual. Plans may not be strictly compliant, but still compatible with reasonable standard of care off-protocol (3). In EORTC 22922, some techniques were recognized to potentially increase late toxicity, but were allowed since they were not foreseen to impact the primary endpoint of overall survival (37). Moreover, the minimum number of patients and composition of the dataset to be reviewed to adequately evaluate protocol compliance has not yet been determined. The optimal timing and type of QA procedures for a given multicenter clinical trial cannot yet be determined with certainty (38), and any recommendations would be incomplete without data on relative cost-effectiveness, which are sorely lacking at present. Evidence from which to derive quantitative deviation criteria will require the establishment of globally harmonized RT QA procedures, a case QA repository with links to clinical outcomes, and further secondary analyses of previously published trials which included various levels of QA (8,38). Finally, as suggested by Bekelman et al, prospective studies evaluating different levels of RT QA as part of the trials themselves are urgently needed (38).

Efforts to improve the quality of RT delivered within cooperative clinical group protocols are ongoing. As advanced technologies are incorporated into and evaluated by multicentre clinical trials, the time and resources required to maintain a rigorous QA program have increased substantially, and are increasingly seen as barriers to accrual (35,38). Cooperative groups would like to avoid the dedication of significant financial resources and time to a large, multicentre trial which then fails to change clinical practice due to perceived flaws, such as the absence of sufficient quality assurance (39). The concern for undue burden on local investigators, however, must be balanced with the need to ensure safe delivery of care on clinical trials (38). Study participation is now more often limited to sites that can contribute a specific number of patients (3). Some cooperative groups have implemented practice cases, or dummy runs (DR), to try to increase rates of compliance with protocol RT, with early data suggesting DR participation significantly increases the likelihood of successful completion of subsequent QA procedures within the same or later trials (40). The decision to require successful completion of credentialing or other QA requires compromise between the spectrum of institutions that will be permitted to participate in the trial, and the uniformity and quality of the data to be collected (41). By definition, QA will constrain the participating institutions to those that can pass the requirements, which may reduce the number of patients enrolled (41). However, these institutions are more likely to comply with the protocol, achieve a lower rate of major deviations, and submit evaluable patient data (41,42). It also appears likely that participating in a clinical trial improves quality of RT delivered to patients off-protocol due to the training of personnel and institutional implementation of best practices (38,39).

In the tiered approach suggested by the NCI's Working Group on RT QA, if the main trial objective necessitates advanced planning or RT techniques, more intense QA is required than a trial incorporating standard RT (8,38). For example, real-time rapid review may not be required for all protocols (39). A standardized submission process, automated where possible, that builds on participation in previous RT QA procedures would increase efficiency (38,39). Overburdening sites with requirements for data submission not directly related to study aims should be minimized (38). Implementation of appropriate levels of QA per study should also be guided by evidence-based predictors of RT quality, such as institutional accrual patterns and the local site's QA results from previous trials.

Conclusions

It is challenging to draw definitive conclusions from the available literature due to differences in central review methodology, definitions of deviations, timing and blinding, median follow-up and statistical analysis. Although more data are required, current reports suggest that protocol-compliant RT tends to decrease failure rates and increase overall survival, and likely contributes to the ability of collected data to answer the central trial question.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: Supported by NCI grants U24 CA81647, CA29511 and CA10953.

References

1. Reinstein L, Peachey S, Laurie F, Glicksman A. Impact of a dosimetry review program on radiotherapy in group trials. Int J Radiat Oncol Biol Phys. 1985; 11:1179–1184. [PubMed: 3888935]
2. Fisher B, Montague E, Redmond C, et al. Findings from NSABP Protocol No B-04: Comparison of radical mastectomy with alternative treatments for primary breast cancer. Cancer. 1980; 46:1–13. [PubMed: 6992972]
3. Peters L, O'Sullivan B, Giralt J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 0202. J Clin Oncol. 2010; 28(18):2996–3001. [PubMed: 20479390]
4. Bentzen S, Bernier J, Davis J, et al. Clinical impact of dosimetry quality assurance programmes assessed by radiobiological modelling of data from the thermoluminescent dosimetry study of the European Organization for Research and Treatment of Cancer. Eur J Cancer. 2000; 36:615–20. [PubMed: 10738126]
5. World Health Organization. Quality Assurance in Radiotherapy. Geneva: WHO; 1998.
6. Budiharto T, Musat E, Poortmans P, et al. Profile of European radiotherapy departments contributing to the EORTC Radiation Oncology Group (ROG) in the 21st century. Radiother Oncol. 2008; 88:403–10. [PubMed: 18538427]
7. Van Tienhoven G, Mijnheer B, Bartelink H, Gonzalez D. Quality assurance of the EORTC trial 22881/10882: Boost versus no boost in breast conserving therapy. An overview. Strah Onkol. 1997; 173(4):201–7.
8. Melidis C, Bosch W, Izewska J, et al. Quality assurance for clinical trials in radiotherapy [Abstract]. Radiother Oncol. 2012; 103(1):S546–7.
9. Matzinger O, Poortmans P, Giraud J-Y, et al. Quality assurance in the 22991 EORTC ROG trial in localized prostate cancer: Dummy run and individual case review. Radiother Oncol. 2009; 90:285–90. [PubMed: 19038468]
10. Perez C, Krauss S, Bartolucci A, et al. Thoracic and elective brain irradiation with concomitant or delayed multiagent chemotherapy in the treatment of localized small cell carcinoma of the lung. Cancer. 1981; 47:2407–2413. [PubMed: 6268269]
11. Perez C, Stanley K, Grundy G, et al. Impact of irradiation technique and tumour extent in tumour control and survival of patients with unresectable non-oat cell carcinoma of the lung. Cancer. 1982; 50:1091–9. [PubMed: 6286087]
12. White J, Chen T, McCracken J, et al. The influence of radiation therapy quality control on survival, response, and sites of relapse in oat cell carcinoma of the lung. Cancer. 1982; 50:1084–90. [PubMed: 6286086]
13. Kinzie J, Hanks G, MacLean C, Kramer S. Patterns of care study: Hodgkin's disease relapse rates and adequacy of portals. Cancer. 1983; 52:2223–2226. [PubMed: 6640492]
14. Ortiz H, Marcial V, Hanley J, et al. Protocol variations and their influence on the final outcome: Critical review of the split-course base of tongue study of the RTOG. Int J Radiat Oncol Biol Phys. 1985; 11:885–892. [PubMed: 3886608]
15. Pajak T, Laramore G, Marcial V, et al. Elapsed treatment days: A critical item for radiotherapy quality control review in head and neck trials. Int J Radiat Oncol Biol Phys. 1991; 20:13–20. [PubMed: 1993621]
16. Seegenschmiedt M, Sauerbrei W, Sauer R, et al. Quality control review for radiotherapy of small breast cancer: Analysis of 708 patients in the GBSG I trial. Strah Onkol. 1993; 169(6):339–350.
17. Fabian C, Mansfield C, Dahlberg S, et al. Low-dose involved field radiation after chemotherapy in advanced Hodgkin's disease. Ann Intern Med. 1994; 120:903–912. [PubMed: 8172436]
18. Duhmke E, Franklin J, Pfreundschuh M, et al. Low-dose radiation is sufficient for the noninvolved extended-field treatment in favourable early-stage hodgkin's disease: Long-term results of a randomized trial of radiotherapy alone. J Clin Oncol. 2001; 19(11):2905–2914. [PubMed: 11387364]
19. Engert A, Schiller P, Josting A, et al. Involved-field radiotherapy is equally effective and less toxic compared with extended-field radiotherapy after four cycles of chemo in patients with early stage

unfavourable Hodgkin's lymphoma: Results of the HD8 trial of the German Hodgkin's Study Group. J Clin Oncol. 2003; 21(19):3601–8. [PubMed: 12913100]

20. Aleman B, Girinsky T, van der Maazen R, et al. Quality control of involved-field radiotherapy in patients with advanced Hodgkin's lymphoma (EORTC 20884). Int J Radiat Oncol Biol Phys. 2005; 63(4):1184–1190. [PubMed: 15936157]
21. Engert A, Franklin J, Eich H, et al. Two cycles of Doxorubicin, Bleomycin, Vinblastine, and Dacarbazine plus extended-field radiotherapy is superior to radiotherapy alone in early favourable Hodgkin's lymphoma: Final results of the GHSG HD7 Trial. J Clin Oncol. 2007; 25:3495–3502. [PubMed: 17606976]
22. Crane C, Winter K, Regine W, et al. Phase II study of bevacizumab with concurrent capecitabine and radiation followed by maintenance gemcitabine and bevacizumab for locally advanced pancreatic cancer: Radiation Therapy Oncology Group 0411. J Clin Oncol. 2009; 27:4096–4102. [PubMed: 19636002]
23. Eisbruch A, Harris J, Garden A, et al. Multiinstitutional trial of accelerated hypofractionated intensity-modulated radiation therapy for early-stage oropharyngeal cancer (RTOG 0022). Int J Rad Oncol Biol Phys. 2010; 76:1333–1338.
24. Abrams R, Winter K, Regine W, et al. Failure to adhere to protocol specified radiation therapy guidelines was associated with decreased survival in RTOG 9704: A phase III trial of adjuvant chemotherapy and chemoradiotherapy for patients with resected adenocarcinoma of the pancreas. J Clin Oncol. 2012; 82:809–16.
25. Rischin D, Peters L, O'Sullivan B, et al. Tirapazamin, cisplatin, and radiation versus cisplatin and radiation for advanced squamous cell carcinoma of the head and neck (TROG 0202, HeadSTART): A phase III trial of the Trans-Tasman Radiation Oncology Group. J Clin Oncol. 2010; 28(18):2989–95. [PubMed: 20479425]
26. Poortmans P, Venselaar J, Struikmans H, et al. The potential impact of treatment variations on the results of radiotherapy of the internal mammary lymph node chain: A quality assurance report on the dummy run of EORTC phase III randomized trial 22922/10925 in stage I-III breast cancer. Int J Radiat Oncol Biol Phys. 2001; 49(5):1399–408. [PubMed: 11286848]
27. Glicksman A, Reinstein L, Laurie F. Quality assurance of radiotherapy in clinical trials. Can Treat Rep. 1985; 69(10):1199–1205.
28. Deutsch M, Bryant J, Bass G. Radiotherapy review on National Surgical Adjuvant Breast and Bowel project (NSABP) phase III breast cancer clinical trials: Is there a need for submission of portal/simulation films? Am J Clin Oncol. 1999; 22(6):606–8. [PubMed: 10597746]
29. Santanam L, Hurkmans C, Mutic S, et al. Standardized naming conventions in radiation oncology. Int J Radiat Oncol Biol Phys. 2012; 83(4):1344–9. [PubMed: 22245204]
30. Fairchild A, Weber D, Bar-Deroma R, et al. Quality Assurance in the EORTC 22033-26033/CE5 Phase III Randomised Trial for Low Grade Glioma: The Digital Individual Case Review. Radiother Oncol. 2012; 103(3):287–292. [PubMed: 22560713]
31. Musat E, Poortmans P, van den Bogaert W, et al. Quality assurance in breast cancer: EORTC experiences in the phase III trial on irradiation of the internal mammary nodes. Eur J Cancer. 2007; 43(4):718–24. [PubMed: 17240136]
32. Muller R, Eich H. The development of quality assurance programs for radiotherapy within the German Hodgkin Study Group. Strah Onkol. 2005; 181:557–66.
33. Duhmke E, Diehl V, Loeffler M, et al. Randomized trial with early-stage Hodgkin's disease testing 30Gy vs 40Gy extended field radiotherapy alone. Int J Radiat Oncol Biol Phys. 1996; 36(2):305–10. [PubMed: 8892452]
34. Poortmans P, Ataman F, Davis J, et al. Quality assurance in the EORTC phase III randomized 'boost vs no boost' trial for breast conserving therapy: Comparison of the results of two individual case reviews performed early and late during the accrual period. Radiother Oncol. 2005; 76:278–84. [PubMed: 15919127]
35. Weber D, Tomsej M, Melidis C, Hurkmans C. QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy. Radiother Oncol. 2012; 105(1):4–8. [PubMed: 22985777]

36. Ohri N, Shen X, Dicker A, Doyle L, Harrison A, Showalter T. Radiotherapy protocol deviations and clinical outcomes : A meta-analysis of cooperative group clinical trials. J Natl Cancer Inst. 2013; 105(6):387–393. [PubMed: 23468460]
37. Lievens Y, Poortmans P, Van den Bogaert W. A glance on quality assurance in EORTC study 22922 evaluating techniques for internal mammary and medial supraclavicular lymph node chain irradiation for breast cancer. Radiother Oncol. 2001; 60(3):257–65. [PubMed: 11514005]
38. Bekelman J, Deye J, Vikram B, Bentzen S, Bruner D, Curran W, et al. Redesigning radiotherapy quality assurance: Opportunities to develop an efficient, evidence-based system to support clinical trials – Report of the National Cancer Institute Work Group on Radiotherapy Quality Assurance. Int J Radiat Oncol Biol Phys. 2012; 83(3):782–790. [PubMed: 22425219]
39. Goodman K. Quality assurance for radiotherapy: A priority for clinical trials. J Natl Cancer Inst. 2013; 105(6):376–377. [PubMed: 23468459]
40. Fairchild A, Collette L, Hurkmans C, et al. Does quality of the EORTC dummy run predict outcomes of multicentre clinical trials? Eur J Cancer. 2012; 48:3232–9. [PubMed: 22766515]
41. Ibbott G, Followill D, Molineu A, et al. Challenges in credentialing institutions and participants in advanced technology multi-institutional clinical trials. Int J Radiat Oncol Biol Phys. 2008; 71:S71–75. [PubMed: 18406942]
42. Followill D, Urie M, Galvin J, Ulin K, Xiao Y, Fitzgerald T. Credentialing for participation in clinical trials. Front Oncol. 2012; 2:198–204. [PubMed: 23272300]
43. Poortmans P, Kouloulias V, Venselaar J, et al. Quality assurance of EORTC trial 22922/10925 investigating the role of internal mammary-medial supraclavicular irradiation in stage I–III breast cancer: The individual case review. Eur J Cancer. 2003; 39(14):2035–42. [PubMed: 12957458]
44. Davis J, Reiner B, Dusserre A, et al. Quality assurance of the EORTC trial 22911. A phase III study of post-operative external radiotherapy in pathological stage T3N0 prostatic carcinoma: The dummy run. Radiother Oncol. 2002; 64:65–73. [PubMed: 12208577]
45. Tai P, Van Dyk J, Yu E, et al. Variability of target volume delineation in cervical esophageal cancer. Int J Radiat Oncol Biol Phys. 1998; 42(2):277–88. [PubMed: 9788405]
46. Van der Schueren E, Horiot J, Leunens G, et al. Quality assurance in cancer treatment. Report of a Working Party from the European School of Oncology. Eur J Cancer. 1993; 29A(2):172–81. [PubMed: 8422278]

Table 1

Summary of trials.

Trial (Ref)	Disease	Open	Accrued	Trial Randomization
Heme				
PCS (13)	Stage I–III HL	1976*	253	Extended survey of treatment and outcomes after RT alone or concurrent chemoRT
SWOG 7808 (17)	Stage III–IV HL	1978 – 1988	278	Randomized those with complete response to induction chemo -> 20Gy involved field RT vs no RT
GHSG HD4 (18)	Favourable stage I–II HL	1988 – 1994	382	40Gy extended field RT vs 30Gy extended field + 10Gy boost to involved field (no chemotherapy)
GHSG HD7 (21)	Favourable stage I–II HL	1994 – 1998	650	30Gy extended field RT + 10Gy involved field boost vs multiagent induction chemo -> same RT
GHSG HD8 (19)	Unfavourable stage I–II HL	1993 – 1998	1204	Patients without PD after induction chemo randomized to extended vs involved field RT (doses as in HD4)
EORTC 20884 (20)	Stage III–IV HL	1989 – 2000	421	Randomized those with complete response to induction chemo -> 24Gy involved field RT vs no RT
H&N				
RTOG 7102 (14)	Base of tongue	1971 – 1976	145	66Gy/30 continuous vs split course RT (30Gy/10 -> break - > 30Gy/10)
RTOG 7913 (15)	Inoperable H&N	1979 – 1983	210	66-73.8Gy/33-41 continuous vs 60Gy/50 via 1.2Gy delivered twice per day
RTOG 7915 (15)	Inoperable H&N	1979 – 1983	306	66-73.8Gy/33-41 RT +/− radiosensitizer misonidazole
RTOG 0022** (23)	T1-2 N0-2 Oropharynx	2001 – 2005	69	Feasibility study of moderately accelerated radical SIB IMRT alone
TROG 0202 (3)	Stage III–IV head & neck^	2002 – 2005	861	Concurrent chemoRT versus concurrent chemoRT + radiosensitizer tirapazamine
Lung				
SWOG 7628 (12)	Limited stage SCLC	1976 – 1979	298	2 × 2 randomization: two different multiagent chemo regimens; BCG or no BCG; all received chest and brain RT
SECSG 76 (10)	Limited stage SCLC	1976†	70	RT (chest/brain) + concurrent multiagent chemo vs same RT alone
RTOG 7301 (11)	Unresectable stage III NSCLC	1973 – not spec	481	40Gy split course RT vs 40Gy/20 vs 50Gy/25 vs 60Gy/30 continuous
Breast				
NSABP B04 (2)	Operable breast	1971 – 1974	1765	cN0: Radical mast vs total mast vs total mast + regional RT; cN+: Radical mast vs total mast + regional RT
GBSG I§ (16)	pT1 N0 breast	1983 – 1989	1119	After lumpectomy + axillary lymph node dissection, randomization to RT or mast
GI				
RTOG 9704 (24)	Pancreas T1-4 N0-1 post-GTR	1998 – 2002	538	Adjuvant chemo§ ->50.4Gy RT + concurrent 5FU -> further chemo§; § randomization 5FU vs gem
RTOG 0411** (22)	Pancreas T1-4 Nx-1 unresectable	2005 – 2006	94	RT (50.4Gy/28) + concurrent capecitabine/ bevacizumab -> maintenance gemcitabine + bevacizumab until progression

* Retrospective review.

** Phase II trial.

^ Excluding T1-2N1 and M1.

[†] Terminated early due to poor accrual.

[§] Randomized trial changed to prospective observational multicentre study due to low accrual.

Abbreviations: chemo – chemotherapy; cN0 – clinically lymph node negative; cN+ - clinically lymph node positive; gem – gemcitabine; GTR – gross total resection; HL – Hodgkin lymphoma; IMRT – intensity-modulated radiotherapy; mast – mastectomy; NSCLC – non-small cell lung cancer; PCS – Patterns of Care Study; PD – progressive disease; RT – radiotherapy; SCLC – small cell lung cancer; SIB – simultaneous integrated boost.

Table 2

Methodology of quality assurance review.

Trial (Ref)	General QA	Patient-Specific QA	Plan Review Performed
NSABP B04, 1980 (2)	Site visit by RPC and TMC to evaluate general procedures, machine calibration, output	RT plans and photos examined to verify RT was delivered according to protocol requirements; 1D dosimetry recalculated by RPC & RT reviewed (including port films) by TMC	• Retrospectively * • Clinical outcomes probably known based on the 1977 report
SWOG 7628, 1982 (12)	Site visits by RPC to check machine calibration	1D Dosimetric reconstruction by RPC and port films reviewed and graded by study coordinator	• Retrospectively • With respect to when clinical outcomes were known : NS
SECSG 76, 1981 (10)	Machine calibration monitored by RPC	RT records, dose distributions, isodose curves, simulation and port films reviewed by study coordinator and Quality Control Centre	• Retrospectively • With respect to when clinical outcomes were known: NS
RTOG 7301, 1982 (11)	Machine calibration monitored by RPC	Port films, daily RT dose records and isodose computations evaluated and 1D dosimetric reconstruction performed by RPC	• Retrospectively • Blinded to clinical outcomes
PCS, 1983 (13)	NS	RT records, hospital records, diagnostic x-rays, simulator films, and port films reviewed by a team of staff radiation oncologists and trained data managers during site visits Dose recalculation performed by physicist team members	• Retrospectively • Not blinded to clinical outcomes
RTOG 7102, 1985 (14)	Machine calibration monitored by RPC	Dosimetry, technique of each case reviewed by RPC	• Retrospectively • With respect to when clinical outcomes were known : NS
RTOG 7913, 1991 (15)	Machine calibration monitored by RPC	Dose recalculation data reviewed by RT staff; study chair reviewed each case for accuracy of treatment plan (simulation/port films), dosimetry and normal tissue	• Retrospectively • With respect to when clinical outcomes were known : NS
RTOG 7915, 1991 (15)	Machine calibration monitored by RPC	Dose recalculation data reviewed by RT staff; study chair reviewed each case for accuracy of treatment plan (simulation/port films), dosimetry and normal tissue	• Retrospectively • With respect to when clinical outcomes were known : NS
GBSG I, 1993 (16)	Semi-annual meetings held for review of ongoing studies	4 reference centres examined RT records for protocol compliance, completeness	• Retrospectively • With respect to when clinical outcomes were known : NS

Trial (Ref)	General QA	Patient-Specific QA	Plan Review Performed
SWOG 7808, 1994 (17)	Machine calibration monitored by RPC	Port films, dose calculations and treatment records reviewed by QA centre, RPC and study coordinators	• Retrospectively • With respect to when clinical outcomes were known : NS
GHSG HD4, 2001 (18)	Information provided to participating RO at annual meetings for quality control	Expert panel designed each RT plan based on staging provided by site **. After RT completed, simulation and port films submitted to reference centre for review.	• Retrospectively ** • Blinded to clinical outcomes
GHSG HD8, 2003 (19)	Same as HD4 (presumed)	Expert panel designed each RT plan based on staging provided by site **. After completion, simulation and port films submitted to reference centre for review by 4 ROs.	• Retrospectively ** • With respect to when clinical outcomes were known : NS
GHSG HD7, 2007 (21)	Same as HD4 (presumed)	Appropriate RT planned centrally by expert RO panel ** Retrospective review details NS	• Retrospectively • With respect to when clinical outcomes were known : NS
EORTC 20884, 2005 (20)	NS	Imaging, RT charts, simulation and port films, and treatment planning calculations reviewed by RO panel	• Retrospectively • Blinded to clinical outcomes
RTOG 0022, 2010 (23)	Credentialed before participation by IMRT plan review + phantom	Delineation evaluated by study chair on initial cases from each site, followed by spot checks. ITC reviewed initial port films, DRRs & dose distributions from each site then spot checks. All plans submitted to ITC.	• Prospectively + retrospectively • With respect to when clinical outcomes were known : NS
TROG 02.02, 2010 (3)	NS	Sites submitted diagnostic imaging & RT plans to QARC by the end of the first week of RT. After RT completion, full documentation reviewed by TMC for protocol compliance with secondary review of non-compliant plans for predicted impact on outcome	• Prospectively (687/820) + retrospectively • Blinded to clinical outcomes
RTOG 9704, 2012 (24)	Machine calibration monitored by RPC	One PI reviewed submission; if RT was per protocol, no further review done; if RT was less than PP, other PI reviewed to confirm.	• Retrospectively • Blinded to clinical outcomes
RTOG 0411, 2009 (22)	Machine calibration monitored by RPC	Diagnostic imaging, GTV contours, DRRs, simulation films, DVHs, dosimetry including isodose distributions evaluated	• Retrospectively • With respect to when clinical outcomes were known : NS

Abbreviations: 1D – one-dimensional dose reconstruction to the prescription point specified by each protocol. IMRT – intensity-modulated radiotherapy; NS – not specified; QA – quality assurance; RO – radiation oncologist; RPC – Radiological Physics Centre; TMC – Trial Monitoring Committee.

*Retrospective review refers to review subsequent to the completion of protocol radiotherapy.

**Central prescription of RT described in some publications as prospective QA.

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2014 October 01.

^ Submission of materials prospectively was requested but not completed.

Table 3

Summary of trial deviation definitions. Although source publication terminology differed, deviation has been used in this table for consistency.

Trial (Ref)	Major Deviation	Minor Deviation
Heme		
PCS (13) *^	• Inadequate upper, axillary, mediastinal-hilar, or abdominal margins • Block consistently touched or overlapped tumour • Block covered uninvolved LN which were to be treated • Geographic miss of LN which were to be treated	
SWOG 7808 (17)	• Failure to give RT to a previously involved site • Concomitant administration of chemotherapy with RT	• Dose infractions * • Failure to complete RT for any reason
GHSG HD4 *^ (18)	"Relevant" protocol deviations: • Incomplete coverage of tumour or inadequate safety margin • Total dose <90% • <1.8Gy/day • >2 weeks delay during treatment to one volume • >4 weeks interval between large fields • Technical: lack of megavoltage equipment or large-field method Other: • Excessive coverage • Total dose >110%	
GHSG HD7 * (21)	• Volume too small or too large • RT too protracted in time • Dose too low or too high	
GHSG HD8 (19)	Not defined	Not defined
EORTC 20884 (20)	• Omission of or incomplete RT to an originally involved area (except omission of spleen if involvement of PA LN or omission of PA LN with involvement of spleen) • Dose <90%	Not defined
H&N		
RTOG 7102 ^ (14)	Major: • Dose or OTT within +/−16 to 20% • >2 but <4 week interruptions Unacceptable: • Dose or OTT >20% • >4 week interruptions (except for split course)	• Dose or OTT within +/−6% to 15% • >1–2 week interruptions
RTOG 7913 and RTOG 7915 (15)	Major deviation acceptable: • Total dose +/− 16–20%	• Total dose +/− 6–15% • 6–15% variation from specified fractionation

Trial (Ref)	Major Deviation	Minor Deviation
	• 16–20% variation from specified fractionation • Partial miss of primary tumour • OTT within 8–14 days of protocol Major deviation unacceptable: • Total dose >+/−20% • >20% variation from specified fractionation • Primary tumour omitted from field • OTT >14 days from protocol specification	• Tight field margins around primary • OTT within 4–7 days of protocol specification
RTOG 0022 (23)	For PTVs: • Not meeting criteria for either no or minor deviation For the parotid glands: • Dose goals not met but >60% of each gland received >30Gy	• Prescription criteria are not met, but all of the following are fulfilled: For PTV66: • 60Gy isodose covers ≥99% of PTV66, 66Gy isodose surface covers ≥90%, and 72.6Gy isodose covers ≤25% For PTV60: • 52Gy isodose covers >90% of PTV60, and 72.6Gy isodose covers ≤20% except when it coincides with PTV66 For PTV54: • 47Gy isodose covers ≥99% of PTV54, 54Gy isodose covers ≥90%, and 72.6Gy isodose covers ≤20% of PTV54 For the parotid glands: • Dose goals are not met but <60% of one gland received >30Gy
TROG 0202 (3)	• Dose not 2Gy/fraction • Gross disease receiving <66.5Gy (except LN <2cm in size) • >10% of PTV receiving <66.5Gy (<57Gy for small LN) or >75Gy • OTT >9 weeks • Max spinal cord dose >50Gy • Excessive volumes or doses to uninvolved normal tissues *	Not defined
Lung		
SWOG 7628 (12)^	• 5–10% underdosage of any involved area • ≥10% underdosage of any area • ≥10% overdosage of any critical normal structure • Incorrect daily dose • Delivery of only 1 field/day to chest • Omission of brain or supraclavicular areas	Not specified in publication
SECSG 76 (10)	• Dose >11% • Inadequate margin on primary tumour, ipsilateral hilum or adjacent mediastinum *	• Dose within +/− 6–10% • Exclusion of the contralateral hilum

Trial (Ref)	Major Deviation	Minor Deviation
	• Incomplete RT or gross geographic miss considered not evaluable	• Margin around tumour of 1cm
RTOG 7301 (11)	Major: • Dose within +/− 11–15% • Margin on primary: 0 • Elective target: partially treated Unacceptable: • Dose >16% • Primary tumour geographic miss	• Dose within +/− 6–10% • Margin on primary: 0.5cm • Margin on elective volume: 0
Breast		
NSABP B04 (2)[^]	• Different dose limits of acceptability defined depending on target volume and nodal status • No posterior axillary portal • Use of kV photons except for treatment of scar • 1–2 fractions/week • 3 fractions/week with OTT <29d • 4–6 fractions/week with OTT <22d	• Different dose limits of acceptability defined depending on target volume and nodal status • Use of electrons • 3 fractions/week with OTT of >28d • 4–6 fractions/week with OTT of 22–28 days, ≥50 days or >63 days depending on axillary boost
GBSG I[§] (16)	• RT started ≥57 days post-surgery • Designed portals yielded an unacceptable dose distribution/target volume coverage • Dose delivered <90% or >110% • Inhomogeneity >+/−10% (linac) or >+/−20% (cobalt) • OTT ≥57d, <4 or >5 fractions/week • Treatment interruption ≥7 days • Missing parasternal, infra- or supraclavicular portal • Any other protocol deviations resulting in unacceptable over- or undertreatment[*]	• RT started 43–56 days post-surgery • Designed portals yielded an acceptable dose distribution/target volume coverage but not as per protocol • Dose not delivered as per protocol but within +/−10% • Inhomogeneity +/− 5–10% (linac) or +/−20% (cobalt) • OTT 50–56d, 4–5 fractions/week • Treatment interruption 4–6 days • Normal parasternal portal but missing infra- or supraclavicular portal • Any other protocol deviations resulting in acceptable over- or undertreatment[*]
GI		
RTOG 9704 (24)[^]	For APPA fields: • Length: <4 or >5 VB • Distance from 1° tumour bed: <1cm or >4cm • Distance from VB edge: <1cm or >3cm For lateral fields: • Length: <4 or >5 VB • Distance from anterior tumour bed: <0.5cm or >3cm • Distance from rest of tumour bed: <1cm or >4cm • Posterior edge: >1cm away from mid VB	For APPA fields: • Distance from primary tumour bed: >3cm to <4cm; 1cm to <2cm • Distance from VB edge: >2cm to <3cm For lateral fields: • Distance from anterior tumour bed: 0.5cm to <1.5cm; >2–3cm • Distance from rest of tumour bed: 1cm to <2cm; >3–4cm • Posterior edge: within 1cm of mid VB • Distance from anterior field edge to VB: 2.5 to <3.5cm; >4–5cm Boost fields:

Trial (Ref)	Major Deviation	Minor Deviation
	• Distance from anterior field edge to VB: <2.5cm or >5cm Boost fields: • Distance from tumour bed: <1cm or >3cm Other: • Dose delivered: not within +/−10% • OTT: >14 days' break	• Distance from tumour bed: 1 cm to <1.5cm; >2–3cm Other: • Dose delivered: not within +/−5% but within +/−10% • OTT: 8–14 days' break
RTOG 0411 ** (22)	• GTV 5cm greater than actual tumour size in any dimension • Inability to contour GTV • Use of block margin >5cm	Not defined

* Not further specified.

^ Publication listed other parameters assessed for RT quality but definitions of minor versus major deviations not specified.

Abbreviations : GTV – gross tumour volume; LN – lymph node s; max – maximum; OTT – overall treatment time; PA – para-aortic; PTV – planning target volume; RT – radiotherapy; VB – vertebral body.

Table 4

Deviation rates of reviewed radiotherapy plans.

Trial (Ref)	N Evaluable for Review	Acceptable	Minor Deviation	Major Deviation
NSABP B04 (2)	646 patients were randomized to RT (352 cN0, 294 cN+), 543 evaluable (298 cN0, 245 cN+) from 34 sites	253/543 (46.6%) • cN0 : 62.4% (186/298) • cN+ : 27.3% (67/245)	35.5% (193/543) * • cN0 : 19.5% (58/298) • cN+ : 55.1% (135/245)	17.9% (97/543) • cN0 : 18.1% (54/298) • cN+ : 17.6% (43/245)
SWOG 7628 (12)	277/298 patients eligible of which 140/277 were evaluable for RT	68.6% (96/140) with no or minor deviation		31.4% (44/140)
SECSG 76 (10)	53 evaluable	66% (35/53) fully compliant	26.4% (14/53)	7.5% (4/53)
RTOG 7301 (11)	346/378 evaluable	87.7% (277/316) at primary site	NS	12.3% (39/316) at primary site
PCS (13)	181/253 port films from 4 sites	63.5% (115/181) adequate	36.5% (66/181) inadequate	
RTOG 7102 (14)	142/145 eligible pts were evaluable from 22 sites	74.4% (102/137) **	14.6% (20/137) **	10.9% (15/137) major or unacceptable variation **
RTOG 7913 (15)	187 eligible patients (94 hyperfractionated arm, 93 control arm) of which 163 were evaluable for RT plan review (83 hyper-fractionated, 80 control)	76% of control pts were either per protocol or with a minor deviation versus 72% of hyperfractionated arm		14% (11/80) major variation rate for control RT vs 24% (20/83) for hyperfractionated RT
RTOG 7915 (15)	297 patients eligible of which 263 were evaluable for RT (130 misonidazole, 133 control)	75% of control pts were either per protocol or with a minor deviation versus 65% of misonidazole arm		23% (31/133) major variation rate for control RT versus 30% (39/130) for misonidazole + RT
GBSG I (16)	708/773 pts who got RT were evaluable from 69 sites	41.2% (292/708)	41.0% (290/708)	17.8% (126/708)
SWOG 7808 (17)	135 were randomized to RT, of which 104 received it and 96% of patients receiving RT were reviewed (100 presumed)	86% (86/100) had RT without a major deviation		NS
GHSG HD4 (18)	368/376 eligible patients were reviewed	63.0% (232/368) had no PD 65.8% (242/368) had no relevant PD for RFS analysis§		37.0% (136/368) had ≥1 PD 34.2% (127/368) had ≥1 relevant PD for RFS analysis§
GHSG HD8 (19)	1064 eligible patients started RT from 212 sites Number reviewed NS	NS No difference between study arms		
GHSG HD7 (21)	529/627 eligible patients reviewed from 189 sites	34.2% (181/529) had no PD	NS	65.8% (348/529) had ≥1 type of PD
EORTC 20884 (20)	135/172 patients randomized to RT were evaluable from 42 sites	53.5% (72/135) had no major deviation		46.7% (63/135) had ≥1 major deviation
RTOG 0022 (23)	53/67 eligible pts were evaluable from 14 sites	0% (0/53)	88.7% (47/53)	11.3% (6/53)

Trial (Ref)	N Evaluable for Review	Acceptable	Minor Deviation	Major Deviation
TROG 02.02 (3)	818/853 patients were reviewed by the TMC, of which 780 were reviewed for outcomes from 81 sites	74.6% (610/818) compliant on 2° review No difference between trial arms	NS	25.4% (208/818) non- compliant by TMC review of which 47.1% (97/206) had an expected major adverse impact on local control
R TOG 9704 (24)	416/451 eligible patients were evaluable	51.9% (216/416) PP No difference between trial arms	48.1% (200/416) were less than PP No difference between trial arms	
			42.1% (175/416) variation acceptable	6.0% (25/416) variation unacceptable
RTOG 0411 (22)	82/94 patients eligible from 36 sites; all evaluable for RT	86.6% (71/82) no unacceptable deviation		13.4% (11/82) unacceptable deviation

*Including intermediate

**Based on NSD ratio

§Numbers from Figure 4 in source publication. Only relevant PDs in patients attaining a complete response are included in RFS analysis.

Abbreviations : cN0 – clinically lymph node negative; cN+ - clinically lymph node positive; NS – not specified; PD – protocol deviation; pp – per protocol; RFS – recurrence-free survival; RT – radiotherapy; TMC – Trial Management Committee.

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2014 October 01.

Table 5

RT plan quality correlation with outcomes. RTOG 0411 did not report any of these outcomes and therefore was excluded from the table.

Trial (Ref)	Locoregional and/or Distant Failure	Progression-Free Survival	Overall Survival
NSABP B04 (2)	• cN0 : 23.2% (26/112) with deviation, 25.8% (48/186) without • cN+ : 42.1% (75/178) with deviation, 50.7% (34/67) without • cN0/cN+ : No significant difference *	NR	• cN0 : 25.0% (28/112) died with deviation, 24.7% (46/186) without • cN+ : 39.3% (70/178) died with deviation, 44.8% (30/67) without • cN0/cN+ : No significant difference *
SWOG 7628 (12)	• 52% had LRR (50/96) if minor viol/no viol vs 64% (28/44) with major viol (p=0.27) • For pts with CR, corresponding #s are 34% (13/38) and 69% (11/16) with p=0.042	NR	• MS 60 wks if acceptable/minor deviation vs 40 weeks if major deviation (p=0.002) for patients receiving RT • cN0 contralateral hilum: better OS if full compliance or minor deviation vs major (p=0.017)
SECSG 76 (10)	• 69.2% (9/13) with inadequate portals had LRR vs 32.5% (13/40) with adequate (p=0.026)	NR	NR
RTOG 7301 (11)	• Primary site: 45.8% (127/277) with adequate RT had intrathoracic recurrence vs 53.8% (21/39) with major viol (p=NS *) • Nodal sites: No significant difference in intrathoracic recurrence between adequate RT vs major deviation	NR	• MS of 42–44wks comparable in patients who had major deviation at the primary site and those who did not
PCS (13)	• Overall: 14.8% (17/115) with adequate RT relapsed vs 50.0% (33/66) inadequate (p=0.0001) • Overall: 7.8% (9/115) with adequate RT had infield or marginal recur vs 31.8% (21/66) inadequate (p=0.001) • RT + chemo: 17.6% (3/17) with adequate relapsed vs 22.2% (2/9) inadequate RT (p=NS *) • RT + chemo: 11.8% (2/17) with adequate had infield or marginal recur vs 22.2% (2/9) inadequate (p=NS *) • RT only: 14.3% (14/98) with adequate had any recur vs 54.4% (31/57) inadequate (p=0.0001) • RT only: 7.1% (7/98) with adequate had infield or marginal recur vs 33.3% (19/57) inadequate(p=0.001)	NR	NR
RTOG 7102 (14)	NR	NR	2 year OS with no evidence of disease **: • 21.5% (22/102) acceptable

Trial (Ref)	Locoregional and/or Distant Failure	Progression-Free Survival	Overall Survival
			• 20.0% (4/20) minor variation • 20.0% (2/10) major variation • 0% (0/9) unacceptable variation (p=NS *)
RTOG 7913 & 7915 (15)	• The 13 patients with major deviation unacceptable (TD, borders, fract) had poorer LRC than those without (p=0.06) • 3yr LRC rate 13% if major deviation unacceptable vs 27% if acceptable (p=0.01) • 53 pts with major deviation unacceptable (TD, OTT, borders, fract) had poorer LRC than those without (p=0.001)	NR	• 13 pts with major deviation unacceptable (TD, borders, fract) had significantly decreased OS (p=0.0009) • 3 year OS 26% if acceptable vs 13% if unacceptable (p=0.01) • 53 pts with major deviation unacceptable (TD, OTT, borders, fract) had significantly worse OS (p=0.0003)
GBSG I (16)	NR	No significant difference in DFS for no/minor vs major deviation *	NR
SWOG 7808 (17)	• Of patients who relapsed, 56% had a major RT deviation and 12.5% had a minor RT deviation * • 9/86 (10%) without a major deviation relapsed	NR	NR
GHSG HD4 (18)	• Presence of a relevant PD significantly increased the risk of relapse (p=0.009)	• Relevant PD significantly decreases 7 year RFS (72% vs 84%) (p=0.004)	NR
GHSG HD8 (19)	• PD had no influence on failure (numbers NR)	• No influence on FFTF (numbers NR)	• PD had no influence on OS (numbers NR)
GHSG HD7 (21)	NR	• No influence on FFTF (numbers NR; p=NS) *	NR
EORTC 20884 (20)	• 15.3% (11/72) without major deviations vs 9.5% (6/63) with major deviation relapsed (p=0.31 **) • No relation between presence of a major deviation and local failure	NR	NR
RTOG 0022 (23)	• Of 7 those with LRR: 2 had major deviations and 2 had inevaluable plans • 2/4 with major deviations (PTV dose) had local failure vs 3/49 LRR with no major deviation in PTV dose (p=0.04)	NR	NR

Trial (Ref)	Locoregional and/or Distant Failure	Progression-Free Survival	Overall Survival
TROG 02.02 (3)	• Patients with major deficiencies (N=87) had inferior outcomes vs those whose RT was initially protocol compliant (N=502): 2 year freedom from LRR 54% vs 78% (p<0.001)[§]	• Major deficiencies had worse 2 yr failure-free survival vs initially compliant RT, number NR	• Patients with major deficiencies (N=87) had inferior outcomes vs those whose RT was initially protocol compliant (N=502): 2 year OS 50% vs 70% (p<0.001)[§]
RTOG 9704 (24)	• PP score associated with first failure (171/347) significantly less often than plans less than PP (176/347), p=0.016 • 45/69 PP patients had no failure vs 24/69 patients with plans less than PP • Isolated LRR occurred in 59 pts: 49% with less than PP and 51% PP (p=NS)[*]	NR	• For all patients, OS increased for PP vs less than PP (MS 1.7 vs 1.5yrs; p=0.008) • Also true for the subset of pancreatic head tumours (MS 1.7 vs 1.5yrs; p=0.03)

[*] No p value given in source publication.

[**] Based on NSD ratio.

[^] P value for log-rank test on the difference between cumulative relapse rates.

[§] 780 patients with evaluable plans who received at least 60Gy were analyzed for outcomes.

Abbreviation : DFS – disease-free survival; FFTF – freedom from treatment failure; LRC – locoregional control; LRR – locoregional recurrence; MS – median survival; NR – not reported; OS – overall survival; PD – protocol deviation; PP – per protocol; RFS – recurrence-free survival.

Table 6

Factors affecting ability to comply with protocol-specified radiotherapy.

Reason	References
• Misinterpretation of or ambiguities in protocol	43–44
• Institutional differences in treatment planning eg margins required for set-up reproducibility; modelling of build-up regions; calculation algorithms	7,13,44–45
• Differences in quality or interpretation of staging	18,20,22
• Changes introduced deliberately to prevent or treat toxicity	20,45
• Interphysician variation in target volumes	44–45
• Deliberate deviation from a protocol which is perceived as too radically different from standard practice	44
• Radiotherapy trials cannot be blinded	18
• Implementation of an unfamiliar treatment technique	26,37
• Specifics of patient case mix	46
• Investigator experience	2,3,16,18,23,34