HHS Public Access

Author manuscript

Pract Radiat Oncol. Author manuscript; available in PMC 2021 July 01.

Published in final edited form as:

Pract Radiat Oncol. 2020 ; 10(4): 265–273. doi:10.1016/j.prro.2019.11.007.

Patterns of Failure Observed in the 2-Step Institution Credentialing Process for NRG Oncology/Radiation Therapy Oncology Group 1005 (NCT01349322) and Lessons Learned

X. Allen Li, PhD[a,*], **Jennifer Moughan, MS**[b], **Julia R. White, MD**[c], **Gary M. Freedman, MD**[d], **Douglas W. Arthur, MD**[e], **James Galvin, DSc**[f], **Ying Xiao, PhD**[d], **Susan McNulty, BS**[f], **Janice A. Lyons, MD**[g], **Vivek S. Kavadi, MD**[h], **Marc T. Fields, MD**[i], **Melissa P. Mitchell, MD**[j], **Bethany M. Anderson, MD**[k], **Michael I. Lock, MD**[l], **Kristine E. Kokeny, MD**[m], **Jose G. Bazan, MD**[c], **Adam D. Currey, MD**[a], **Tarek Hijal, MD**[n], **Sally B. Cheston, MD**[o], **Frank A. Vicini, MD**[p]

[a]Department of Radiation Oncology, Medical College of Wisconsin, Milwaukee, Wisconsin [b]NRG Oncology Statistics and Data Management Center/ACR, Philadelphia, Pennsylvania [c]Ohio State University Comprehensive Cancer Center, Columbus, Ohio [d]Department of Radiation Oncology, Abramson Cancer Center, University of Pennsylvania, Philadelphia, Pennsylvania [e]Massey Cancer Center, Virginia Commonwealth University, Richmond, Virginia [f]Imaging and Radiation Oncology Core (IROC), Philadelphia, Pennsylvania [g]Case Western Reserve University, Cleveland, Ohi [h]Texas Oncology Cancer Center Sugar Land, Sugar Land, Texas [i]Rohnert Park Cancer Center Kaiser Permanente Medical Group, Rohnert Park, California [j]University of Kansas Cancer Center [k]University of Wisconsin Hospital and Clinics, Madison, Wisconsin [l]London Regional Cancer Program, London, Ontario, Canada [m]Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah [n]Division of Radiation Oncology, McGill University Health Centre, Montreal, Canada [o]Department of Radiation Oncology, Greenebaum Comprehensive Cancer Center, University of Maryland, Baltimore, Maryland [p]21st Century Oncology MHP — Farmington, Farmington Hills, Michigan

Abstract

Purpose: To investigate patterns of failure in institutional credentialing submissions to NRG/RTOG 1005 with the aim of improving the quality and consistency for future breast cancer protocols.

Methods and Materials: NRG/RTOG 1005 allowed the submission of 3-dimensional conformal radiation therapy (3DCRT), intensity-modulated radiation therapy (IMRT), and simultaneous integrated boost (SIB) breast plans. Credentialing required institutions to pass a 2-step quality assurance (QA) process: (1) benchmark, requiring institutions to create a plan with no unacceptable deviations and < 1 acceptable variation among the dose volume (DV) criteria, and (2) rapid review, requiring each institution's first protocol submission to have no unacceptable

[*]Corresponding author: X. Allen Li, PhD; ali@mcw.edu.

Supplementary data

Supplementary material for this article can be found at https://doi.org/10.1016/j.prro.2019.11.007.

deviations among the DV criteria or contours. Overall rates, number of resubmissions, and reasons for resubmission were analyzed for each QA step.

Results: In total, 352 institutions participated in benchmark QA and 280 patients enrolled had rapid review QA. Benchmark initial failure rates were similar for 3DCRT (18%), IMRT (17%), and SIB (18%) plans. For 3DCRT and IMRT benchmark plans, ipsilateral lung most frequently failed the DV criteria, and SIB DV failures were seen most frequently for the heart. Rapid review contour initial failures (35%) were due to target rather than organs at risk. For 29% of the rapid review initial failures, the planning target volume boost eval volume was deemed an unacceptable deviation.

Conclusions: The review of the benchmark and rapid review QA submissions indicates that acceptable variations or unacceptable deviations for the ipsilateral lung and heart dose constraints were the most commonly observed cause of benchmark QA failure, and unacceptable deviations in target contouring, rather than normal structure contouring, were the most common cause of rapid review QA failure. These findings suggest that a rigorous QA process is necessary for high quality and homogeneity in radiation therapy in multi-institutional trials of breast cancer to ensure that the benefits of radiation therapy far outweigh the risks.

Introduction

For eligible patients with stage I and II invasive breast cancer, the addition of whole breast irradiation (WBI) after lumpectomy results in durable long-term local control rates on the order of 90% to 95% and survival outcomes equivalent to mastectomy.[1,2] Given that the use of breast conservation surgery compared with mastectomy may correlate with the ready availability of radiation therapy,[3,4] some patients may forego breast conservation surgery because the 6 to 7 week time frame for conventionally fractionated WBI is too lengthy. Hypofractionated WBI (HWBI) has emerged as a way to reduce the overall radiation therapy (RT) treatment duration.[5] Multiple prospective randomized clinical trials using HWBI have demonstrated that cancer outcomes are not inferior to conventional fractionated WBI.[6–11]

However, these trials did not address the optimal fractionation method for delivery of the boost to the tumor bed or the outcome for those with higher risk breast cancer cases requiring boost. To address these concerns and other issues (eg, the outcome of patients who were not included in the previous HWBI trials), a randomized phase III noninferiority study was developed to compare accelerated HWBI with concurrent boost to standard WBI or HWBI with sequential boost: NRG/RTOG 1005.[12]

Institutions that participated in the NRG/RTOG 1005 study were required to meet specified technology requirements and to have passed a 2-step quality assurance (QA) process before enrolling or treating patients. The first step, hereafter referred to as the “benchmark,” entailed the creation of a treatment plan upon a standard computed tomography (CT) data set with predefined regions of interest (ROI) including targets and organs at risk (OAR). The standard CT and ROI set was provided by the protocol, and the institution’s treatment plan was subject to approval by the protocol study chairs. An approved benchmark demonstrated that the institution had the technology and procedures in place to produce adequate treatment plans. The second step, hereafter referred to as the “rapid review,” required for the

concurrent boost arm (Arm 2), entailed the submission of the planning CT data set, plan ROIs and planned dose distributions for the institution's first 3-dimensional conformal radiation therapy (3DCRT) and first intensity modulated radiation therapy (IMRT) cases randomized to that arm. The rapid review case was reviewed by one of the protocol radiation oncology study chairs for approval. The rapid review assessed the institution's technology and procedures for delineation and for planning. More details on these 2 QA steps can be found in the NRG/RTOG 1005 protocol.[12] After rapid review approval, the rest of the cases entered onto the concurrent boost arm were reviewed in a timely fashion throughout the remaining accrual to the trial. Timely reviews were also done for the cases on the sequential boost arm (arm 1).

Only upon fulfilling the benchmark and rapid review requirements could an institution submit additional cases to the protocol; institutions that failed each QA step were required to resubmit. The NRG/RTOG 1005 study group has compiled statistics indicating resubmission rates and reasons for submission failure for the benchmark and rapid review QA steps among all participating institutions. The objectives of the present study, which is focused on the benchmark and rapid reviews, are to analyze the resubmission frequencies and reasons for resubmission and identify specific contingencies for which QA step failure tended to occur. Understanding these patterns of failure may help to reduce the QA resubmission rate and improve overall quality of data sets for future multi-institutional trials.

Methods and Materials

The NRG/RTOG 1005 protocol accepted for benchmark credentialing submissions of intact breast RT with a concurrent boost or simultaneous integrated boost (SIB) to the postlumpectomy bed. Seven RT techniques were allowed by the protocol:

- 3DCRT WBI with 3DCRT photon concurrent boost
- 3DCRT WBI with IMRT concurrent boost
- 3DCRT WBI with electron concurrent boost
- IMRT WBI with 3DCRT concurrent boost
- IMRT WBI with IMRT concurrent boost
- IMRT WBI with electron concurrent boost
- IMRT WBI with IMRT SIB

The protocol specifically defines the IMRT and SIB techniques.[12] In the Discussion section of the present article, 3 broad categories of plan modalities are defined: "3DCRT" plans featured a concurrent boost and were such that the WBI component was planned using 3DCRT; "IMRT" plans featured a concurrent boost and were such that the WBI component was planned with IMRT; and "SIB" plans featured a simultaneous boost which could only be planned using IMRT.

For several ROIs, the protocol provided the contouring guidelines and a set of planning goals in term of the dosevolume (DV) compliance. Defined for each planning goal was (1) an

"acceptable variation," such that the planning goal for the ROI was not met, but was otherwise considered clinically acceptable by the protocol and thus the case enrollment into the protocol could proceed, and (2) an "unacceptable deviation," such that the case could not be enrolled into the protocol. For example, the protocol requires that at least 95% of the lumpectomy cavity evaluation planning target volume receives at least 95% of the boost prescribed dose and considers the variation acceptable if at least 90% of the lumpectomy cavity evaluation planning target volume receives at least 90% of the boost prescribed dose and, otherwise, unacceptable. The details of the protocol requirements and acceptable variations are provided as Table E1 (available online at https://doi.org/10.1016/j.prro.2019.11.007).

An institution was considered to have failed the benchmark QA if there were any unacceptable deviations or if there were 2 or more acceptable variations. The benchmark QA submission was thus graded more stringently by the protocol than a routine case submission (for which no unacceptable deviations were allowed, but multiple acceptable variations were permitted). An institution was considered to have failed the rapid review QA if there were any unacceptable deviations among the DV and contouring criteria (a guideline that was also applied to subsequent case submissions).

For both QA steps, resubmission was required until approval; there was no limit imposed upon the number of submission attempts for either QA step. DV data for the benchmark failures and DV and contouring data for the rapid review failures were compiled and analyzed. The overall submission rates, number of resubmissions and reasons for resubmission for both QA steps were analyzed.

Results

The NRG/RTOG 1005 trial enrolled QA cases and protocol patient plans from May 24, 2011 to June 20, 2014, with a total of 2,354 cases accrued at closing. The number of institutions participating in the benchmark QA process was 352 and, among that group, 280 rapid review cases were submitted.

Benchmark credentialing

Of the 352 institutions that submitted benchmark QA plans, 74 of them (21%) failed the initial submission (Table 1). Of these, 61 institutions (82.4%) passed the first resubmission, 12 institutions (16.2%) required a second resubmission, and 1 institution (1.4%) required a third resubmission attempt (Table 1). The rates of initial submission success and resubmission for each plan modality (3DCRT, IMRT, or SIB) are also summarized in Table 1. Initial submission failure rates were observed to be 18.0%, 17.0%, and 17.6% for 3DCRT, IMRT, and SIB cases, respectively. Among the institutions failing the first benchmark QA submission, the pass rates for the first resubmission were 81.1%, 88.2%, and 81.0% for 3DCRT, IMRT, and SIB plan modalities, respectively. These results suggest that the reasons for benchmark submission failure do not strongly depend on the plan modality and that other underlying causes may be responsible.

For the 30 3DCRT benchmark QA cases that failed the initial submission and required one resubmission, 9 of them (30.0%) exhibited at least one unacceptable deviation, 17 (56.7%) exhibited at least 2 acceptable variations, and 4 (13.3%) violated both criteria (Table 2). This suggests that the failed benchmark QA plans are more likely to have multiple acceptable variations than to have an unacceptable deviation. Note that a submitted protocol case with at least 2 variations could be enrolled into the protocol, although it would fail the benchmark QA test. The benchmark QA pass criteria were designed to be more stringent to encourage institutions to create higher quality plans. The number of acceptable variations and unacceptable deviations for various ROI categories observed among the 3DCRT benchmark cases are listed in Table 3. Ipsilateral lung DV results exhibited acceptable variations for 21 cases (70.0%) and unacceptable deviations for 2 cases (6.7%). The next most frequent ROI exhibiting acceptable variations for 10 cases (33.3%) was the surgical cavity boost planning target volume (PTV), and 3 cases (10.0%) had unacceptable deviations. This suggests that fulfilling the ipsilateral lung DV criteria is a challenge for most institutions requiring resubmission and that special attention should be devoted to the ipsilateral lung during the planning process. Comparing with other modalities (IMRT and SIB; Tables 4 and 5), 3DCRT had a higher rate of unacceptable deviation on the maximum point dose to the contralateral breast.

For the 15 IMRT benchmark QA cases that failed the initial submission and required one resubmission, 2 of them (13.3%) exhibited at least one unacceptable deviation, 8 (53.3%) exhibited at least 2 acceptable variations, and 5 (33.3%) violated both criteria (Table 2). The percentage of failed IMRT QA submissions exhibiting 2 or more acceptable variations but no unacceptable deviations is similar to the failed 3DCRT benchmarks. The number of acceptable variations and unacceptable deviations for various ROI categories observed among the IMRT benchmark cases are listed in Table 4. Similarly, to the 3DCRT data in Table 3, the most frequently observed ROI DV violation is the ipsilateral lung with the plans from 12 institutions (85.7%) exhibiting an acceptable variation. For institutions that fail the initial benchmark QA submission, lung DV criteria appear to be similarly difficult to meet when using either 3DCRT or IMRT.

Among the 17 SIB benchmark QA cases failing initial submission and requiring one resubmission, 5 of them (29.4%) exhibited at least one unacceptable deviation, 7 (41.2%) exhibited at least 2 acceptable variations, and 3 (17.6%) violated both criteria (Table 2). The slightly lower percentage of benchmark plans otherwise acceptable to the protocol may suggest that unacceptable deviations are more likely with the SIB technique. The number of acceptable variations and unacceptable deviations for various ROI categories observed among the SIB benchmark cases are listed in Table 5. Although ipsilateral lung DV variations and deviations are apparent among SIB plans, acceptable variations for one of the heart dose constraints (constraint 3) is most prominent (8 cases, 47.1%). It may be that SIB planning is able to more easily satisfy the lung DV objectives relative to 3DCRT or non-SIB IMRT planning, but at the expense of heart dose sparing.

Rapid review

Table 6 shows the overall resubmission rate and the number of resubmissions required, respectively, for the 280 rapid review QA submissions. At initial submission, the percentage of cases failing the DV criteria (35.4%) is nearly equal to the percent failing the contour criteria (35.0%). Subsequent resubmission rates for DV-criteria and contour-criteria failures are also similar. This suggests that the stringency of the DV criteria is not considerably different from that of the contour criteria.

Among rapid review submissions that required one resubmission before approval (n = 80), Table 7 shows the incidences of contour-criteria unacceptable deviations for several target ROIs and OARs. Among the target ROIs, contours were deemed as unacceptable deviations in 5.1% (for the surgical bed) to 29.5% (for the PTV boost eval) of submissions. Unacceptable deviations among OARs tended to be considerably lower, ranging from 1.3% (for the ipsilateral lung) to 10.4% (for the thyroid). This general trend was also apparent among rapid review submissions that required 2 submissions before approval.

It is worthwhile to determine whether institutions, which failed the benchmark QA, were more likely to fail the rapid review QA than those that passed the benchmark QA. For all plan modalities and among the 3DCRT rapid review cases, Table 8 shows the contour rapid review resubmission rates for institutions passing the initial benchmark QA and those that required at least one benchmark QA resubmission. In particular, for the 3DCRT plans, the likelihood of an institution failing the initial rapid review QA appeared to be independent of the results of its initial benchmark QA submission.

Discussion

Patterns of failure in institutional credentialing submissions to NRG/RTOG 1005 study were analyzed. Benchmark QA submissions were reviewed to assess the participating institutions' compliance with DV guidelines, and rapid review QA submissions were reviewed to assess compliance with contouring guidelines. Seventy-four (21%) of the participating institutions failed in their initial submissions for the benchmark QA, and 35% of cases failed in the initial submission for the rapid review. The benchmark QA failure appeared to be independent of plan modality. For 3DCRT and IMRT planning without SIB, acceptable variations or unacceptable deviations for the ipsilateral lung dose constraints were the most commonly observed cause of benchmark QA failure, whereas for SIB DV failures were seen most frequently for the heart. In the rapid review submissions, the rates of unacceptable deviations in the contours of the targets were 5% for the surgical bed and 30% for the PTV boost eval, substantially higher than those for OARs, ranging from 1% for the ipsilateral lung to 10% for the thyroid. These findings indicate that a rigorous QA process is necessary for high quality and homogeneity in radiation therapy in multi-institutional trials of breast cancer to ensure that the benefits of radiation therapy far outweigh the risks.

One reason for the observed high failure rates in the QA processes might be that there was a learning curve for some institutions, as NRG/RTOG 1005 was the first breast cancer cooperative group RT trial requiring detailed 3D delineation of the targets and OARs and the restricted DV constraints for whole breast irradiation, and participating in this trial might be

the first time for these institutions to perform 3D treatment plans for breast cancer. The higher failure rate in the contours of the targets compared with those for the OARs may be due in part to the fact that the definitions of the target ROIs may not have been as standardized as were the OAR definitions. This could be due either to inherent difficulties in defining target ROIs for breast patients (due to, for example, the breast and lumpectomy site being in a predominantly soft-tissue environment, variable use of surgical clips or seroma formation), or to a less rigorous definition in the protocol for target ROIs, in particular their CTV definitions and PTV expansions. It should be noted that large interobserver variability in structure delineation has been reported for breast RT planning.[13]

The relatively high failure rate on OAR DV constraints underlines the importance of quality assurance to ensure the quality and uniformity of treatments in each arm, so that the comparison between the arms can be meaningful in such a large cooperative group trial. The challenge for meeting OAR constraints would encourage institutions to consider approaches leading to small PTV margins, such as the use of robust patient positioning (eg, supine vs prone) or delivery techniques (eg, respiration gating, deep inspiration breath hold), which, in turn, would promote good clinical practice. The tumor control and toxicity outcome data from this trial may demonstrate whether the contouring variation revealed in the rapid reviews QA process is important.

The protocol required DV constraints are believed to be reasonable based on the existing practice in the study chairs' institutions. The fact that NRG/RTOG 1005 was the fastest accruing breast cancer trial in the history of RTOG and met all required accrual 2 years before the proposed closing date demonstrates that the protocol requirements were practically doable. However, the benchmark criteria considering 2 or more acceptable variations as failure might have been too stringent. Practically, treatment planners may consider multiple factors to balance the target coverage versus OAR sparing. Moreover, the current analysis indicates that there is no correlation between the failures of benchmark and rapid review QA, implying that making benchmark QA stringent did not result in improvement in the rapid review submission. Based on these considerations, the criterion of considering 2 or more acceptable variations as failure in future trials should not be used. If unacceptable deviation was only considered as failure, the failure rate would be reduced by approximately 50% (Table 2).

Although the transition from 2-dimensional to 3D planning for breast cancer has been challenging and the learning curve for the 3D treatment planning could be steep, the authors firmly believe that only with the volumetric information of the targets and OARs can it be ensured that the targets are adequately covered and the doses to OARs are minimized, thus maximizing the therapeutic ratio.

Incomplete volumetric approaches, eg, 2.5D, could not guarantee optimal treatment. As shown in Table 7, nearly 30% of institutions requiring one resubmission in the rapid review had unacceptable deviations on the contouring of PTV boost eval. Such incorrect delineation would increase the risk of recurrence and/or toxicity. Table 3 indicates that near 26% of institutions requiring one submission of their 3DCRT benchmark plans had unacceptable

deviations in the contralateral breast doses. Without the volumetric information, the opportunity to minimize the dose to an OAR would be missed.

The contributing factors identified for the failures of these QA processes would help to optimize the QA process in future trial designs that aim to minimize heterogeneity in the treatment planning and delivery within a treatment arm between institutions, thus, to maximize seeing any existing therapeutic effects between arms. It has been shown that poorly executed treatment delivery in clinical trials may potentially obscure the benefits of a new treatment technique[14] or significantly affect the treatment outcome such as that reported by Hansen et al[15] and those reviewed by Ohri et al.[16] Radiation therapy reduces the absolute risk of breast cancer mortality but can cause secondary cancer or heart disease decades later.[17] A rigorous QA process is necessary for high quality and homogeneity in radiation planning and delivery in multi-institutional trials of breast cancer RT to ensure that the benefits of RT far outweigh the risks.

Considering that NRG/RTOG 1005 was the first breast cancer cooperative group RT trial requiring detailed delineation of the targets and OARs and restricted DV constraints for whole breast irradiation, the 20% to 35% failure rate in the first submission of the 2-step QA is not discouraging. The reduction of the failure rate in the subsequent submission of the QA process implies that additional training and instruction is important. Consequently, the rapid review QA process was eliminated in the late part of this trial as the trial investigators became familiar with the required delineation and DV constraints.

Recognizing that participating in this trial might be the first time for some institutions performing 3D treatment planning, the study chairs made a great deal of effort to ensure the communication to institutional investigators was sufficient. These efforts included (1) detailed guidelines and explanations in the protocol on contouring, treatment planning techniques, and DV constraints, (2) the protocol specific workshops/group discussions organized by the study cochairs at NRG/RTOG semiannual meetings before and during the trial opening period, (3) availability of study chairs to address questions and concerns, and (4) links to cooperative group QA websites for institutions to download relevant guidelines and sample cases. Such communications should be considered necessary for similar future trials.

Conclusions

The review of the benchmark and rapid review QA submissions to the NRG/RTOG 1005 protocol indicates that the benchmark QA failure appeared to be independent of plan modality; however, for 3DCRT and IMRT planning without SIB, acceptable variations or unacceptable deviations for the ipsilateral lung dose constraints were the most commonly observed cause of benchmark QA failure, whereas for SIB DV failures were seen most frequently for the heart. Rapid review QA failures tended to arise owing to unacceptable deviations in target contouring rather than OAR contouring. A rigorous QA process is necessary for high quality and homogeneity in radiation therapy in multi-institutional trials of breast cancer to ensure that the benefits of RT far outweigh the risks.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Input from Kristofer Kainz, PhD in data extraction is appreciated.

Sources of support: This project was supported by grants U24CA180803 Imaging and Radiation Oncology Core (IROC), U10CA180868 (NRG Oncology Operations), U10CA180822 (NRG Oncology Statistics and Data Management Center [SDMC]), UG1CA189867 (NRG Oncology NCI Community Oncology Research Program [NCORP]) from the National Cancer Institute (NCI).

Disclosures: Ms McNulty, Ms Moughan, Drs Galvin, Arthur, Bazan, Cheston, Currey, Fields, Freedman, Hijal, Kavadi, Kokeny, Lyons, Mitchell, White, and Xiao have nothing to disclose. Dr Anderson reports standard cooperative group grant support per patient enrolled from NRG Oncology/Radiation Therapy Oncology Group. Dr Li reports grants from Elekta, Siemens Healthineers, and Manteia Med Technologies outside the submitted work. Dr Lock discloses personal fees from Abbvie and Ferring outside the submitted work. Dr Vicini reports an advisory role with ImpediMed during the conduct of the study.

This work was presented at the 2017 American Society for Radiation Oncology Annual Meeting.

References

1. Clarke M, Collins R, Darby S, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: An overview of the randomised trials. Lancet. 2005;366:2087–2106. [PubMed: 16360786]
2. Agarwal S, Pappas L, Neumayer L, et al. Effect of breast conservation therapy vs mastectomy on disease-specific survival for earlystage breast cancer. JAMA Surg. 2014;149:267–274. [PubMed: 24429935]
3. Morrow M, White J, Moughan J, et al. Factors predicting the use of breast-conserving therapy in stage I and II breast carcinoma. J Clin Oncol. 2001;19:2254–2262. [PubMed: 11304779]
4. Polednak AP. Trends in, and predictors of, breast-conserving surgery and radiotherapy for breast cancer in Connecticut, 1988–1997. Int J Radiat Oncol Biol Phys. 2002;53:157–163. [PubMed: 12007955]
5. Holloway CL, Panet-Raymond V, Olivotto L. Hypofractionation should be the new ‘standard’ for radiation therapy. The Breast. 2010;19:163–167. [PubMed: 20511064]
6. Whelan T, MacKenzie R, Julian J, et al. Randomized trial of breast irradiation schedules after lumpectomy for women with lymph node-negative breast cancer. J Natl Cancer Inst. 2002;94:1143–1150. [PubMed: 12165639]
7. Yarnold J, Ashton A, Bliss J, et al. Fraction sensitivity and dose response of late adverse effects in the breast after radiotherapy for early breast cancer: Long-term results of a randomised trial. Radiother Oncol. 2005;75:9–17. [PubMed: 15878095]
8. Owen JR, Ashton A, Bliss JM, et al. Effect of radiotherapy fraction size on tumor control in patients with early-stage breast cancer after local tumour excision: Long-term results of a randomized trial. Lancet Oncol. 2006;7:467–471. [PubMed: 16750496]
9. Yarnold JR, et al. The START Trialists’ Group. The UK Standardisation of Breast Radiotherapy (START) Trial A of radiotherapy hypofractionation for treatment of early breast cancer: A randomised trial. Lancet Oncol. 2008;9:331–341. [PubMed: 18356109]
10. Yarnold JR, et al. The START Trialists’ Group. The UK Standardisation of Breast Radiotherapy (START) Trial B of radiotherapy hypofractionation for treatment of early breast cancer: A randomised trial. Lancet. 2008;371:1098–1107. [PubMed: 18355913]
11. Haviland JS, Owen JR, Dewar JA, et al. The START Trialists’ Group. The UK Standardisation of Breast Radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials. Lancet Oncol. 2013;14:1086–1094. [PubMed: 24055415]

12. Higher per daily treatment-dose radiation therapy or standard per daily treatment radiation therapy in treating patients with early-stage breast cancer that was removed by surgery. ClinicalTrials.gov identifier: NCT01349322. https://clinicaltrials.gov/ct2/show/NCT01349322. Updated August 28, 2018 Accessed January 9, 2020.
13. Li XA, Tai A, Arthur DW, et al. Variability of target and normal structure delineation for breast cancer radiotherapy: an RTOG multi-institutional and multiobserver study. Int J Radiat Oncol Biol Phys. 2009;73:944–951. [PubMed: 19215827]
14. Peters LJ, 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 02.02. J Clin Oncol. 2010;28:2996–3001. [PubMed: 20479390]
15. Hansen CR, Johansen J, Kristensen CA, et al. Quality assurance of radiation therapy for head and neck cancer patients treated in DAHANCA 10 randomized trial. Acta Oncol. 2015;54:1669–1673. [PubMed: 26206516]
16. Ohri N, Shen X, Dicker AP, et al. Radiotherapy protocol deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials. J Natl Cancer Inst. 2013;105:387–393. [PubMed: 23468460]
17. Taylor C, Correa C, Duane FK, et al. Early Breast Cancer Trialists' Collaborative Group. Estimating the risks of breast cancer radiotherapy: Evidence from modern radiation doses to the lungs and heart and from previous randomized trials. J Clin Oncol. 2017;35: 1641–1649. [PubMed: 28319436]

Table 1

Benchmark QA resubmission rates

No. of resubmissions	Sites (n = 352)	3DCRT (n = 205)	IMRT (n = 100)	SIB (n = 119)
0	278[*] (79.0%)	168[†] (82.0%)	83[†] (83.0%)	98[†] (82.4%)
>1	74 (21.0%)	37 (18.0%)	17 (17.0%)	21 (17.6%)
1	61 (17.3%)	30 (14.6%)	15 (15.0%)	17 (14.3%)
2	12 (3.4%)	7 (3.4%)	2[≠] (2.0%)	4 (3.4%)
3	1[§] (0.3%)	—	—	—

Abbreviations'. 3DCRT = 3-dimensional conformai radiation therapy; IMRT = intensity-modulated radiation therapy; SIB = simultaneous integrated boost.

[*]Two sites failed on first benchmark submission and never resubmitted for any modality.

[†]Three, 6, and 5 sites failed on first submission and did not resubmit for 3DCRT, IMRT, and SIB, respectively.

[≠]One of these 2 sites never passed the IMRT benchmark.

[§]One site had one 3DCRT resubmission and 2 SIB resubmissions.

Table 2

Acceptable dose volume variations and unacceptable deviations among institutions requiring one benchmark resubmission

Criterion failure mode	3DCRT (n = 30)	IMRT (n = 15)	SIB (n = 17)
≥1 unacceptable deviations	9 (30.0%)	2 (13.3%)	5 (29.4%)
≥2 acceptable variations	17 (56.7%)	8 (53.3%)	7 (41.2%)
Both ≥ 1 and ≥ 2	4 (13.3%)	5 (33.3%)	3 (17.6%)
Isolated hot spot			1 (5.9%)
Did not pass visual review			1 (5.9%)

Abbreviations: 3DCRT = 3-dimensionalconformalradiationtherapy; IMRT = intensity-modulated radiation therapy; SIB = simultaneous integrated boost.

Table 3

3DCRT benchmark plans

Region of interest	DV criterion[12]	Failure mode	n	%*
Ipsilateral lung	Any	Variation acceptable	21	70.0
Surgical cavity boost PTV	Any	Variation acceptable	10	33.3
Breast PTV eval	Max point dose	Variation acceptable	9	30.0
Contralateral breast	Max point dose	Variation acceptable	9	30.0
Whole breast PTV	Any	Variation acceptable	7	23.3
Whole breast boost PTV	Any	Variation acceptable	6	20.0
Heart dose constraint 1	V20 Gy %	Variation acceptable	3	10.0
Heart dose constraint 3	Mean dose	Variation acceptable	2	6.7
Thyroid (missing, n = 17)†	Max point dose	Variation acceptable	2	15.4
Ipsilateral lung dose constraint 1	V20 Gy %	Variation acceptable	1	3.3
Surgical cavity boost PTV	Max point dose	Variation acceptable	1	3.3
Contralateral breast	Max point dose	Deviation unacceptable	8	26.7
Breast PTV eval	Max point dose	Deviation unacceptable	3	10.0
Surgical cavity boost PTV	Any	Deviation unacceptable	3	10.0
Whole breast PTV	Any	Deviation unacceptable	2	6.7
Whole breast boost PTV	Any	Deviation unacceptable	2	6.7
Ipsilateral lung	Any	Deviation unacceptable	2	6.7
Heart dose constraint 1	V20 Gy %	Deviation unacceptable	1	3.3

Abbreviations: 3DCRT = 3-dimensional conformal radiation therapy; DV = dose volume; IMRT = intensity-modulated radiation therapy; PTV = planning target volume; SIB = simultaneous integrated boost.

Targets with dose volume variations or deviations at initial review, for institutions whose benchmark plan required one resubmission (n = 30). Note that a benchmark case requiring resubmission can exhibit acceptable variations (2 or more) and unacceptable deviations.

*Denominator is 30 unless there are missing regions of interest, in which case denominator = 30 – missing.

†Added at second protocol amendment.

Table 4

IMRT benchmark plans

Region of interest	DV criterion[12]	Failure mode	n	%*
Ipsilateral lung (missing, n = 1)	Any	Variation acceptable	12	85.7
Whole breast PTV	Any	Variation acceptable	6	40.0
Breast PTV eval	Max point dose	Variation acceptable	5	33.3
Heart dose constraint 3 (missing, n = 2)	Mean dose	Variation acceptable	5	38.56
Surgical cavity boost PTV	Any	Variation acceptable	4	26.7
Contralateral breast	Max point dose	Variation acceptable	4	26.7
Ipsilateral lung dose constraint 2 (missing, n = 2)	V10 Gy %	Variation acceptable	3	23.1
Whole breast boost PTV	Any	Variation acceptable	2	13.3
Breast PTV eval	Max point dose	Variation acceptable	2	13.3
Surgical cavity boost PTV (missing, n = 1)	Max point dose	Variation acceptable	1	7.1
Ipsilateral lung dose constraint 1 (missing, n = 2)	V20 Gy %	Variation acceptable	1	7.7
Heart dose constraint 1 (missing, n = 2)	V20 Gy %	Variation acceptable	1	7.7
Heart dose constraint 2 (missing, n = 2)	V10 Gy %	Variation acceptable	1	7.7
Whole breast boost PTV	Any	Deviation unacceptable	3	20.0
Whole breast PTV	Any	Deviation unacceptable	1	6.7
Breast PTV eval	Max point dose	Deviation unacceptable	1	6.7
Surgical cavity boost PTV	Any	Deviation unacceptable	1	6.7
Ipsilateral lung dose constraint 1 (missing, n = 2)	V20 Gy %	Deviation unacceptable	1	7.7
Ipsilateral lung dose constraint 2 (missing, n = 2)	V10 Gy %	Deviation unacceptable	1	7.7
Contralateral breast	Max point dose	Deviation unacceptable	1	6.7
Heart dose constraint 3 (missing, n = 2)	mean dose	Deviation unacceptable	1	7.7

Abbreviations: 3DCRT = 3-dimensional conformal radiation therapy; DV = dose volume; IMRT = intensity-modulated radiation therapy; PTV = planning target volume; SIB = simultaneous integrated boost.

Targets with dose volume variations or deviations at initial review for institutions whose benchmark plan required one resubmission (n = 15); note that a benchmark case requiring resubmission can exhibit acceptable variations (2 or more) as well as unacceptable deviations.

*Denominator is 15 unless there are missing values, in which case denominator = 15 – missing.

Table 5

SIB benchmark plans

Region of interest	DV criterion[12]	Failure mode	n	%*
Heart dose constraint 3	Mean dose	Variation acceptable	8	47.1
Ipsilateral lung	Any	Variation acceptable	5	29.4
Breast PTV eval	Max point dose	Variation acceptable	4	23.5
Surgical cavity boost PTV	Any	Variation acceptable	4	23.5
Contralateral breast	Max point dose	Variation acceptable	4	23.5
Ipsilateral lung	Any	Variation acceptable	4	23.5
Ipsilateral lung dose constraint 1	V20 Gy %	Variation acceptable	2	11.8
Whole breast PTV	Any	Variation acceptable	2	11.8
Ipsilateral lung	Any	Deviation unacceptable	3	17.6
Ipsilateral lung dose constraint 2	V10 Gy %	Deviation unacceptable	2	11.8
Whole breast boost PTV	Any	Deviation unacceptable	1	5.9
Breast PTV eval	Max point dose	Deviation unacceptable	1	5.9
Surgical cavity boost PTV	Any	Deviation unacceptable	1	5.9
Contralateral breast	Max point dose	Deviation unacceptable	1	5.9
Heart dose constraint 3	Mean dose	Deviation unacceptable	1	5.9
Thyroid (missing, n = 4)†	Max point dose	Deviation unacceptable	1	7.7

Abbreviations: 3DCRT = 3-dimensional conformal radiation therapy; DV = dose volume; IMRT = intensity-modulated radiation therapy; PTV = planning target volume; SIB = simultaneous integrated boost.

Targets with dose volume variations or deviations at initial review for institutions whose benchmark plan required one resubmission (n = 17); note that a benchmark case requiring resubmission can exhibit acceptable variations (2 or more) and unacceptable deviations.

*Denominator is 17 unless there are missing values, in which case denominator = 17 – missing.

†Added during second protocol amendment.

Table 6

Rapid review QA acceptance and resubmission rates

No. of resubmissions	Dose volume QA (n = 280)	Contour QA (n = 280)
0	181 * (64.6%)	182 * (65.0%)
≥1	99 (35.4%)	98 (35.0%)
1	84 (84.8%)	82 (83.7%)
2	14 (14.1%)	15 (15.3%)
3	0 (0.0%)	0 (0.0%)
4	1 (1.0%)	1 (1.0%)

Abbreviation: QA = quality assurance.

* Three patients failed on first submission and the institutions did not resubmit.

Table 7

Rapid review contour QA: ROIs with deviations at initial review for patients who required one resubmission (n = 80)

ROI	Failure mode	n	%*
PTV boost eval (missing, n = 2)	Deviation unacceptable	23	29.5
Breast PTV eval	Deviation unacceptable	20	25.0
CTV boost	Deviation unacceptable	17	21.3
PTV boost (missing, n = 2)	Deviation unacceptable	15	19.2
Breast CTV	Deviation unacceptable	15	18.8
Breast PTV (missing, n = 1)	Deviation unacceptable	10	12.7
Thyroid† (missing, n = 3)	Deviation unacceptable	8	10.4
Contralateral breast (missing, n = 1)	Deviation unacceptable	5	6.3
Surgical bed (missing, n = 2)	Deviation unacceptable	4	5.1
Heart (missing, n = 1)	Deviation unacceptable	3	3.8
Ipsilateral lung (missing, n = 1)	Deviation unacceptable	1	1.3

Abbreviations: PTV = planning target volume; ROI = region of interest; QA = quality assurance.

*Denominator is 80 unless there are missing values, in which case denominator = 80 – missing.

†Added during second protocol amendment.

Table 8

Correlations between benchmark QA resubmissions and rapid review contour QA resubmissions

All plan modalities			
Benchmark resubmitted	**Rapid review contour resubmitted**		**Total**
	No	**Yes**	
No	152 (66.4%)	77 (33.6%)	229
Yes	30 (58.8%)	21 (41.2%)	51
Total	182	98	280
3DCRT plans			
Benchmark resubmitted	**Rapid review contour resubmitted**		**Total**
	No	**Yes**	
No	102 (64.6%)	56 (35.4%)	158
Yes	25 (62.5%)	15 (37.5%)	40
Total	127	71	198

Abbreviations: 3DCRT = 3-dimensional conformal radiation therapy; QA = quality assurance.