Can the Choice of Radiotherapy Delivery Technique Influence Which Target Delineation Protocol to Use? A Plan-Quality-Based Analysis in Left Breast Cancer

Irfan Ahmad¹,²,³ Kundan Singh Chufal¹ Chandi Prasad Bhatt²,³,⁴ Alexis Andrew Miller³,⁴ Ram Bajpai⁴,⁵ Rahul Lal Chowdhary¹ Anjali Kakria Pahuja¹ Akanksha Chhabra¹ Munish Gairola¹

¹Department of Radiation Oncology, Rajiv Gandhi Cancer Institute and Research Centre, Rohini, New Delhi, India
²Department of Radiation Oncology, Sarvodaya Hospital and Research Centre, Faridabad, Haryana, India
³Department of Radiation Oncology, Illawarra Cancer Care Centre, Wollongong, New South Wales, Australia
⁴School of Medicine, Keele University, Staffordshire, United Kingdom

Address for correspondence Kundan Singh Chufal, MD, Department of Radiation Oncology, Rajiv Gandhi Cancer Institute and Research Centre, Sector 5, Rohini, New Delhi 110085, India (e-mail: kundan25@gmail.com).

Abstract

Introduction This study investigates the optimal target delineation protocol stratified by treatment planning technique in patients undergoing whole breast radiotherapy after breast conservation surgery.

Materials and Methods Target delineation using Tangent (RTOG 0413 Whole Breast Irradiation Protocol), European Society for Radiotherapy and Oncology (ESTRO), and Radiation Therapy & Oncology Group (RTOG) guidelines was performed on 10 randomly selected treatment planning computed tomography datasets of patients with left-sided breast cancer. An objective plan quality metric (PQM) scoring schema was defined and communicated to the medical physicist prior to commencement of treatment planning. Treatment planning was performed using field-in-field (FiF) intensity modulated radiotherapy technique (IMRT), inverse IMRT, and volumetric modulated arc therapy (VMAT), for each type of target. Two-way repeated measures, analysis of variance was utilized to compare the total PQM scores and dosimetric variables, stratified by treatment planning method.

Results Total PQM score of plans for FiF, IMRT, and VMAT revealed that Tangent and ESTRO delineations were equivalent regardless of planning technique (Tangent vs. ESTRO for FiF, \(p = 0.099\); Tangent vs. ESTRO for IMRT, \(p = 0.029\); Tangent vs. ESTRO for VMAT, \(p = 0.438\)). Both delineation protocols were significantly superior to RTOG for all treatment planning techniques.

Conclusion For all treatment planning techniques, ESTRO and Tangent delineation were equivalent and both achieved significantly higher scores than RTOG delineation.
Introduction

Treatment planning for breast irradiation has evolved from conventional fields based on marks placed around the breast to modern three-dimensional computed tomography (3D CT)-based planning. Recognition of the late effects associated with breast irradiation is manifested primarily as an increased risk of long-term cardiac mortality, specifically in left-sided breast cancer.1,2 Technologically advanced radiotherapy delivery techniques have allowed a reduction in whole heart mean dose, which has emerged as the strongest predictor of radiation-induced cardiac mortality.3,4 Paralleling technological advances, guidelines have been formulated for the delineation of the whole breast target volume, though issues of interobserver variation persist.5-8 These parallel developments in target delineation and treatment delivery techniques have aided the evolution from conventional to 3D CT-based planning.

While the dosimetric performance of modern radiotherapy delivery techniques has been compared, meaningful interpretation is hampered by, but not limited to issues of type of target delineation performed, treatment planning proficiency and the large number of dosimetric parameters evaluated. As radiotherapy techniques and delineation guidelines continue to evolve, disentangling the confounding effects of both on each other may be difficult and is not currently addressed in the literature.

The purpose of this research was to determine the most optimal target delineation protocol for a given type of treatment planning technique, using an objective plan quality assessment tool.

Materials and Methods

We have previously reported the methodology for this analysis (►Fig. 1).9 In brief, we selected 10 left-sided breast cancer patients from our institutional database who had undergone breast conservation surgery and adjuvant radiotherapy (46 Gy in 23Fx with an electron boost of 12.5 Gy in 5Fx to lumpectomy cavity). All patients had undergone a free-breathing CT scan on Siemens Somatom Sensation Open with a slice thickness of 2 mm. Their DICOM CT datasets was retrieved, and target delineation was performed based on the RTOG 0413 (Tangent) Whole Breast Irradiation protocol, the RTOG (Radiation Therapy Oncology Group) consensus and ESTRO (European Societies of Radiation Therapy) consensus guidelines.5,10 Organs-at-risk (OAR), the left anterior descending (LAD) artery, and left ventricle (LV) were delineated as described in the RTOG 1005 protocol (NCT01349322) and a cardiac atlas, respectively.11

Subsequently, treatment planning was performed using field-in-field (FiF) intensity modulated radiotherapy technique (IMRT), inverse optimized IMRT, and volumetric modulated arc therapy (VMAT).12-19 To minimize interobserver variation, one radiation oncologist performed delineation of all structures on a single treatment planning system (Varian Eclipse v13.5, Varian Medical Systems) and a single medical physicist performed treatment planning on a single TPS (Varian Eclipse v13.5; AAA algorithm, Varian Medical Systems, Palo Alto, CA, United States) and delivery platform (Varian TrueBeam v2.5; Millennium 120 MLC, Varian Medical Systems).

Plan quality assessment was performed using an objective plan quality metric (PQM) scale.5,20 Dosimetric data from each plan for each structure was extracted from exported DVH files using DVHmetrics v0.3.5 (www.rdocumentation.org/packages/DVHmetrics), per level 2 reporting recommendations by the ICRU Report 83.21

Statistical comparison was performed using a two-way repeated measure, analysis of variance (ANOVA), after correcting for any possible interaction between target delineation protocol and planning technique in each ANOVA model.22 A Bonferroni correction was applied, and the significance level was set at <0.005 (0.05/9).23 All analyses were performed in Stata 14.2 SE (Statacorp Inc, College Station, TX, United States) statistical software.

Results

Ninety plans were selected from a pool of 450 generated plans on the basis of PQM score and minimum acceptance criteria.24 Descriptive data associated with PQM and dosimetric comparisons is provided in ►Supplementary Material (available in the online version).

PQM-Based Comparison of Delineation Guidelines

For plans based on FiF IMRT, the combined PQM scores for Tangent delineation were equivalent to ESTRO delineation (Tangent vs. ESTRO, p = 0.099) and both were significantly superior to RTOG delineation (Tangent vs. RTOG, p < 0.001; ESTRO vs. RTOG, p < 0.001) (►Fig. 2). The subscore analysis revealed that the planning target volume (PTV) score for ESTRO delineation was higher than both RTOG and Tangent delineations (ESTRO vs. RTOG, p < 0.001; ESTRO vs. Tangent, p = 0.001), and that Tangent delineation was superior to RTOG delineation (Tangent vs. RTOG, p = 0.001). The comparison of subscores for left lung, heart, and right breast revealed that ESTRO and Tangent delineations were equivalent and that both were superior to RTOG delineation.

For plans based on inverse optimized IMRT, the combined PQM scores for Tangent delineation were equivalent to ESTRO delineation (Tangent vs. ESTRO, p = 0.029) and both were significantly superior to RTOG delineation (Tangent vs. RTOG, p < 0.001; ESTRO vs. RTOG, p < 0.001) (►Fig. 3). The subscore analysis for PTV score revealed that all delineation protocols were equivalent. The comparison of subscores for heart and right breast revealed that Tangent and ESTRO delineations were equivalent and both were superior to RTOG delineation. When comparing subscores for left lung, ESTRO delineation was superior to both RTOG and Tangent delineation protocols (ESTRO vs. RTOG, p < 0.001; ESTRO vs. Tangent, p = 0.003), and Tangent delineation was superior to RTOG delineation (Tangent vs. RTOG, p = 0.001).

For plans based on VMAT planning technique (►Fig. 4), the combined PQM scores for Tangent delineation were equivalent to ESTRO delineation (Tangent vs. ESTRO, p = 0.438) and
TPS (Varian Eclipse v13.5; AAA algorithm, Varian Medical Systems, Palo Alto, CA, United States) and delivery platform (Varian TrueBeam v2.5; Millennium 120 MLC, Varian Medical Systems).

Plan quality assessment was performed using an objective plan quality metric (PQM) scale.

Dosimetric data from each plan for each structure was extracted from exported DVH files using DVHmetrics v0.3.5 (www.rdocumentation.org/packages/DVHmetrics), per level 2 reporting recommendations by the ICRU Report 83.

Statistical comparison was performed using a two-way repeated measure, analysis of variance (ANOVA), after correcting for any possible interaction between target delineation protocol and planning technique in each ANOVA model. A Bonferroni correction was applied, and the significance level was set at <0.005 (0.05/9).

All analyses were performed in Stata 14.2 SE (Statacorp Inc, College Station, TX, United States) statistical software.

Results

Ninety plans were selected from a pool of 450 generated plans on the basis of PQM score and minimum acceptance criteria.

Descriptive data associated with PQM and dosimetric comparisons is provided in ►Supplementary Material (available in the online version).

PQM-Based Comparison of Delineation Guidelines

For plans based on FiF IMRT, the combined PQM scores for Tangent delineation were equivalent to ESTRO delineation (Tangent vs. ESTRO, \(p = 0.099\)) and both were significantly superior to RTOG delineation (Tangent vs. RTOG, \(p < 0.001\); ESTRO vs. RTOG, \(p < 0.001\)) (►Fig. 2). The subscore analysis revealed that the planning target volume (PTV) score for ESTRO delineation was higher than both RTOG and Tangent delineations (ESTRO vs. RTOG, \(p < 0.001\); ESTRO vs. Tangent, \(p = 0.001\)), and that Tangent delineation was superior to RTOG delineation (Tangent vs. RTOG, \(p = 0.001\)). The comparison of subscores for left lung, heart, and right breast revealed that ESTRO and Tangent delineations were equivalent and that both were superior to RTOG delineation.

For plans based on inverse optimized IMRT, the combined PQM scores for Tangent delineation were equivalent to ESTRO delineation (Tangent vs. ESTRO, \(p = 0.029\)) and both were significantly superior to RTOG delineation (Tangent vs. RTOG, \(p < 0.001\); ESTRO vs. RTOG, \(p < 0.001\)) (►Fig. 3). The subscore analysis for PTV score revealed that all delineation protocols were equivalent. The comparison of subscores for heart and right breast revealed that Tangent and ESTRO delineations were equivalent and both were superior to RTOG delineation. When comparing subscores for left lung, ESTRO delineation was superior to both RTOG and Tangent delineation protocols (ESTRO vs. RTOG, \(p < 0.001\); ESTRO vs. Tangent, \(p = 0.003\)), and Tangent delineation was superior to RTOG delineation (Tangent vs. RTOG, \(p = 0.001\)).

For plans based on VMAT planning technique (►Fig. 4), the combined PQM scores for Tangent delineation were equivalent to ESTRO delineation (Tangent vs. ESTRO, \(p = 0.438\)) and...

Fig. 1 Study design.\(^5\) ESTRO, European Society for Radiotherapy and Oncology; IMRT, intensity modulated radiotherapy technique; PQM, plan quality metric; RTOG, Radiation Therapy & Oncology Group; VMAT, volumetric modulated arc therapy.
both were significantly superior to RTOG delineation (Tangent vs. RTOG, $p = 0.001$; ESTRO vs. RTOG, $p < 0.001$) (Fig. 4). The subscore analysis for PTV score and right breast revealed that all delineation protocols were equivalent. When comparing
subscores for the heart, ESTRO and Tangent delineation protocols were equivalent (ESTRO vs. Tangent, \( p = 0.353 \)) and RTOG was inferior to ESTRO (RTOG vs. ESTRO, \( p = 0.001 \)) but equivalent to Tangent delineation protocol (RTOG vs. Tangent,
For the left lung, ESTRO was equivalent to Tangent delineation protocol (ESTRO vs. Tangent, $p = 0.907$) and both were superior to RTOG delineation protocol (ESTRO vs. RTOG, $p < 0.001$; ESTRO vs. Tangent, $p < 0.001$).

$p = 0.009$). For the left lung, ESTRO was equivalent to Tangent delineation protocol (ESTRO vs. Tangent, $p = 0.907$) and both were superior to RTOG delineation protocol (ESTRO vs. RTOG, $p < 0.001$; ESTRO vs. Tangent, $p < 0.001$).
Dosimetric Comparison of Delineation Guidelines

►Fig. 5 provides a summarized visual representation of the results of dosimetric comparison for all structures and dosimetric parameters analyzed (full descriptive data with associated p-values are provided in Supplemental Material, available in the online version). The data visualization technique chosen for comprehensively depicting results of all dosimetric comparisons is a modified version of the sunburst chart.24 Each analyzed structure and its associated dosimetric parameters form the origin of hierarchy, and three concentric circles represent the rank order, with the innermost circle being most inferior. Each dosimetric parameter has three radial arms representing planning technique, upon which the comparators (target delineation guidelines) are ranked. This method of qualitative data visualization was chosen because the values being compared were not on a uniform scale.

1. PTV parameters (►Fig. 5A)

Across all planning techniques, RTOG delineation protocol was either dosimetrically inferior to ESTRO delineation protocol (12/24 comparisons) or the overall comparison between all delineation protocols was not statistically significant (11/24 comparisons). The comparisons for tangent delineation protocol were more heterogeneous. Tangent delineation was in nearly equal measures dosimetrically superior to (7/24 comparisons) or equivalent to RTOG delineation (6/24 comparisons).

2. Heart, LAD artery, and LV parameters (►Fig. 5B–D)

For the whole heart contour, across almost all dosimetric parameters and planning techniques, RTOG delineation was dosimetrically inferior to both ESTRO and Tangent delineation (23/30 comparisons), except $D_{2\%}$ for IMRT planning where it was dosimetrically equivalent to Tangent delineation. Except for the comparisons where there was no statistically significant difference amongst all delineations (6/30 comparisons), ESTRO delineation was dosimetrically equivalent to Tangent delineation across all planning techniques and dosimetric parameters (23/30 comparisons).

The dosimetric comparisons for LAD revealed that RTOG delineation was dosimetrically inferior to ESTRO delineation (12/21 comparisons) and Tangent delineation was dosimetrically equivalent to ESTRO delineation (9/21 comparisons). Overall in 9/21 comparisons, there was no statistically significant difference among all delineations.

Similarly, almost all dosimetric comparisons for the LV demonstrated that RTOG delineation was dosimetrically inferior to ESTRO delineation (16/21 comparisons), irrespective of planning technique. Tangent delineation was dosimetrically equivalent to ESTRO delineation for most comparisons (15/21 comparisons), except $D_{2\%}$ with IMRT technique where it was dosimetrically inferior.

3. Left lung (►Fig. 5E)

For the left lung, across all planning techniques and almost all dosimetric parameters, RTOG delineation protocol was dosimetrically inferior to ESTRO delineation protocol (23/24 comparisons). Tangent delineation protocol was dosimetrically equivalent to ESTRO delineation protocol for almost all comparisons (20/24 comparisons) and consequently dosimetrically superior to RTOG delineation for almost all comparisons (22/24 comparisons).

4. Right breast (►Fig. 5F)

Except for two comparisons ($V_3$ and $V_5$ for VMAT planning) where ESTRO delineation was dosimetrically equivalent to RTOG delineation, across all planning techniques and almost all comparisons for the right breast, either RTOG delineation was dosimetrically inferior to both ESTRO and Tangent delineations (16/24 comparisons), or there was no significant difference between delineation protocols (6/24 comparisons).

Discussion

We believe that this study is the first to analyze the performance of contouring protocols when viewed from the perspective of the medical physicist (which is the most optimal target delineation technique for a particular planning technique?). Our results suggest that irrespective of the planning technique employed, total PQM scores for ESTRO and Tangent delineation were equivalent and both achieved significantly higher scores than RTOG delineation.

Our results on the performance of delineation guidelines can be explained by the shared similarities of Tangent and ESTRO delineation guidelines. The principle behind the ESTRO guideline is that target delineation should not result in volumes larger than “conventional” simulator-based RT volumes.7 A notable difference of the ESTRO guideline is to restrict the dorsal border at the pectoralis major muscle and gradually move ventrally near the caudal edge to exclude subcutaneous abdominal fat. However, despite this modification in the ESTRO guidelines, the heart subscores between Tangent and ESTRO delineations were not significantly different (despite the Tangent delineation including the chest wall). The known issue of variation in the medial–lateral direction influencing OAR dose can explain the performance of the RTOG delineation, though delineation was in strict adherence to the published guideline.8 Furthermore, our study did not permit any contour modification after the plans were obtained, in contrast to common clinical practice.8

Our analysis can be criticized as being antithetical to modern radiotherapy practice with limited applicability, as planning technique is determined by target delineation and not vice-versa. We want to emphasize that our objective was purely exploratory and the intent of which was to empirically compare delineation protocols, controlling for as many variables as possible. Obviously, the gold standard for comparing delineation protocols would be clinical data on local failures. Data on regional nodal failures stratified by delineation protocols is now emerging, and while our analysis was not designed to assess the quality of regional nodal irradiation, this is an important avenue for research and our group will be addressing it in the near future.25

Our analysis shows that if a particular planning technique is preferred, some thought should be given to the delineation protocol best suited for it. At this time, in the absence of a local control analysis, it is impossible to decide whether ESTRO or RTOG delineation guidelines should be followed, and until such a report appears, radiation oncologists can
Fig. 5 Modified sunburst charts of dosimetric comparisons, depicting the ordinal performance of delineation guidelines stratified by planning technique (please refer to Supplementary Material, available in the online version, for descriptive data). (A) Planning target volume (PTV), (B) heart, (C) left anterior descending (LAD) artery, (D) left ventricle (LV), (E) left lung, and (F) right breast.
choose either. The results of this analysis and a companion analysis\(^6\) support our conclusion that either FiF or IMRT when coupled with Tangent or ESTRO delineation protocol yields the most optimal combination of target coverage and OAR sparing. We therefore advise against the routine use of RTOG delineation or VMAT planning technique purely on plan quality and dosimetric grounds and await the results of clinical trials utilizing the RTOG delineation guidelines. Just as planning techniques are sympathetic to delineation protocols, delineation protocols can also be sympathetic to planning techniques.

**Conflict of Interest**

None declared.

**References**
