

Consistency of ESTRO and RTOG contouring guidelines for target volume delineation in early stage breast cancer

Abstract

Introduction: Adjuvant radiotherapy (RT), including regional lymph node (RLN) irradiation is an important treatment in early stage breast cancer patients after breast-conserving surgery both the Radiation Therapy Oncology Group (RTOG) and European Society for Radiotherapy and Oncology (ESTRO). Our primary aim was to quantitatively compare target volumes delineated by ROs, to assess if either guideline has superior contouring reproducibility.

Methods: Three ROs contoured breast clinical target volumes (CTVs), axillary lymph node levels 1-3, supraclavicular and internal mammary nodal (CTVn_IMN) volumes for 8 post-operative patients using the ESTRO and RTOG guidelines for each patient. The inter-observer variability in contouring was measured by the generalized Dice Similarity Coefficient (DSC). Mean DSC values were calculated to compare the common volume accuracy between observers.ss

Results: The breast CTV contours showed the highest level of agreement between the ROs within each guideline (ESTRO DSC mean 0.92 ± 0.04 and RTOG DSC mean 0.90 ± 0.03). Due to the small standard deviation in DSC values, this difference of 0.02 was found to be statistically significant ($p=0.031$). The mean DSC for CTVn_IMN was 0.63 (± 0.10) for ESTRO and 0.62 (± 0.20) for RTOG guidelines, showing the least level of agreement for both guidelines. Within the axillary nodal volumes, level 1 showed the greatest agreement among the ROs for both ESTRO and RTOG guidelines with a mean DSC of 0.81 (± 0.08) & 0.82 (± 0.06) respectively. The range of mean DSC values for ESTRO and RTOG guidelines was 0.66-0.92 and 0.62-0.90 respectively.

Conclusion: Overall, neither guideline showed a notable improvement in reproducibility between ROs.

Keywords: breast cancer, regional lymph node, ESTRO guidelines, adjuvant radiotherapy

Volume 7 Issue 4 - 2020

Tony Mathew,¹ Michael Chao,¹ Carminia Lapuz,¹ Jonathan Tomaszewski,² Bob Zhang,¹ Megan Hall,¹ Mark Tacey,¹ Farshad Foroudi¹

¹Department of Radiation Oncology, Olivia Newton-John Cancer Wellness and Research Centre, Australia

²Ballarat Austin Radiation Oncology Centre, Australia

Correspondence: Tony Mathew, Department of Radiation Oncology, Olivia Newton-John Cancer Wellness and Research Centre, Austin Health, Heidelberg, Victoria 3084, Australia, Tel +447551976558, Email tmathew@nhs.net

Received: August 03, 2020 | **Published:** August 24, 2020

Introduction

In early stage breast cancer (ESBC), breast-conserving surgery followed by adjuvant radiotherapy (RT) is an alternative to mastectomy alone.¹ Regional lymph node (RLN) irradiation has an established role in locally advanced breast cancer and recent trials have suggested a benefit to patients with earlier stage disease.² There has been an increase use of intensity modulated radiation therapy (IMRT) or volumetric modulated radiation therapy (VMAT). This has made it possible to deliver precise dose distributions to the planning target volume (PTV) and significantly spare OARs thereby reducing potential toxicities.³ However, safe and effective delivery hinges on accurate and consistent delineation of target volumes and organs at risk. There can be large variations in target volumes when contouring the breast, post-operative tumour bed, chest wall and RLNs in ESBC.⁴⁻⁸ A multi-institutional study by Radiation Therapy Oncology Group (RTOG) focused on interobserver differences in breast cancer radiotherapy delineation. This study found, that in three breast cancer cases contoured by nine radiation oncologists specializing in breast RT, there were volume variations with standard deviations up to 60% and structure overlaps as low as 10%. With the highest variations within the delineation of regional lymph nodes.⁹ Studies focusing on breast target volume delineation emphasize the importance of detailed and unambiguous protocols for delineation in order to minimize

variation.¹⁰ The RTOG Breast Cancer Atlas¹¹ was published in 2009 and has been the foundation of a number of subsequent breast cancer trials.¹²⁻¹⁴ This contouring atlas was based on multicentre expert consensus to minimize the variability between different physicians. This and other similar guidelines,^{12,15} are largely based on the fixed bony and muscular landmarks used for simulator-based treatment set up, which has the potential to lead to larger treatment areas than needed after addition of margins.¹⁶ In 2015, the European Society for Radiotherapy and Oncology (ESTRO) consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer detailed instructions for the transition to CT based volumes.¹⁷

ESTRO has aimed to reduce the CTV volume while still ensuring coverage of the lymphatic drainage system by reducing the margins around the veins, given the proximity of the nodes to these vessels. The differences between the guidelines are more focused on the dorsal, cranial and caudal boundaries of the lymphatic nodal levels. For example, the cranial border of the SCF is noted to be more superior in the RTOG compared to ESTRO guidelines resulting in a larger SCF volume using the RTOG atlas. Nevertheless, to our knowledge there has been limited published research to date comparing the consistency produced between the ESTRO and RTOG guidelines. Gee et al has compared differing breast contouring guidelines including those focused on in this study but discussed possible patterns of failure.¹⁸

The aim of this study was to assess inter-observer variability for the ESTRO and RTOG guidelines for ESBC with regional nodal volumes to determine which guidelines produce results that are more consistent.

Methods

Ethics approval was obtained via Austin Health Human Research Ethics Committee (LNR/16/Austin/95). This prospective study has identified eight patients (four right-sided and four left-sided) with ESBC that had planning CT scans completed between February 2016 and April 2016. Three radiation oncologists (ROs) specialising in breast RT used the ESTRO guidelines while delineating the target volumes for each patient and then, again used the RTOG atlas while they contoured the same volumes for the same patients without any prior formalised training using either guideline. Contoured structures included CTVp_Breast (breast clinical target volume), CTVn_SCF (supraclavicular nodal clinical target volume, ESTRO guidelines have defined it as CTVn_L4), CTVn_L1 (axillary level 1 nodal clinical target volume), CTVn_L2 (axillary level 2 nodal clinical target volume, including interpectoral lymph nodes), CTVn_L3 (axillary level 3 nodal clinical target volume), CTVn_IMN (internal mammary nodal clinical target volume). This will create a total of 16 data sets for each RO (total of 48 datasets).

The data recorded were the target volumes for each structure set in cm³ and the overlap for the same volumes between the different ROs using the same guidelines. Using the Monaco (Elekta, Stockholm) planning system, the volumes (CTVp_Breast, CTVn_SCF, CTVn_IMN, CTVn_L1-L3) produced by each specific guideline have been compared between all the ROs, with the Dice Similarity Coefficient (DSC)⁶ used to measure the level of agreement between the three observers. The DSC was calculated as:

$$DSC = 2 \frac{|A \cap B|}{|A| + |B|}$$

where A and B are the target volumes for observer A and B for the same patient.

Mean DSC values were calculated according to the method indicated in Carillo¹⁹ to compare the common volume accuracy between each pair of observers when using either the ESTRO or RTOG guidelines. We have used a DSC value >0.8 to be considered as very good concordance and a DSC value >0.6 as good concordance.²⁰ Results will be presented as mean and standard deviation (SD), upon confirmation that all target volumes and DSC values were approximately normally distributed using the Shapiro-Wilk test.

Student's t-test was test for differences between RTOG and ESTRO mean DSC values. In addition to the estimated of DSC values, inter-observer variability was also assessed using intra-class correlation (ICC) coefficients estimated as absolute agreement, with fixed effects assumed as the observers were not selected at random. Therefore, the assessment of ICC values does not examine the degree of common volume shared by each observer, but degree of consistency in overall volumes measured by each observer pair. Stata version 15.1 (StataCorp, College Station, Texas, USA) was used to conduct the statistical analysis, with a p-value of less than 0.05 considered to indicate statistical significance.

Results

The mean volumes for each contouring target volume are presented in Table 1. There is evidence of larger target volumes contoured by the observers using the RTOG guidelines except for the CTVn_IMN contour. The estimated intra-class correlation coefficient indicates slightly higher absolute agreement for the CTVn_L1, CTVn_L2, and CTVn_L3 target volumes according to the RTOG guidelines, although the 95% confidence intervals for these coefficients of correlation are wide. The range of mean DSC values when contouring volumes based on the ESTRO guidelines was 0.66-0.92 and the range for the RTOG guidelines was 0.62-0.90. Within each guideline, the CTVp_Breast showed the highest level of agreement between the ROs (Figure 2). CTVp_Breast was also the only volume to show a statistically significant difference in the DSC mean value (p=0.031), with 0.92 (0.04) for ESTRO guidelines compared with 0.90 (0.03) for the RTOG guidelines. The mean DSC for CTVn_IMN was 0.63 (0.10) for ESTRO and 0.62 (0.20) for RTOG guidelines thus showing the least level of agreement within both guidelines but there was no significant difference between the mean DSC values for each guideline (p=0.84). Within the axillary nodal and SCF target volumes, there was similar agreement amongst the observers for both ESTRO & RTOG guidelines. The mean DSC values for the CTVn_L1, CTVn_L2, CTVn_L3 and SCFn ranged from 0.71-0.78 for the ESTRO guidelines and 0.70-0.80 for the RTOG guidelines. When assessing the volumes by each RO that produced the highest DSC value (CTVp_Breast), we can see clear evidence of good consistency by the close distances between the ROs volumes (Figure 3A). Whereas, the volumes that formed at the lowest DSC value (CTVn_IMN), we can see a greater degree of variation in the volumes (Figure 3B), particularly for the ESTRO guideline with one observer consistently delineating a larger volume for overall CTVn_IMN contour. This is supported by the ICC_{AA} for the CTVn_IMN being quite low at 0.06 for the ESTRO guideline, and only slightly higher at 0.08 for the RTOG guidelines.

Table 1 Overall Target volumes by guideline and Intra-Class Correlation, n=24 observer pairs for each guideline

Target volume	Guideline	Mean cm ³ (standard deviation)	ICCAA
CTVp Breast	ESTRO	888.7 (233.8)	0.85 (0.28–0.97)
	RTOG	947.2 (243.8)	0.82 (0.22–0.96)
CTVn IMN	ESTRO	12.02 (5.76)	0.06 (0–0.33)
	RTOG	4.79 (2.85)	0.08 (0-0.39)
CTVn LI	ESTRO	60.44 (18.80)	0.29 (0-0.75)
	RTOG	76.15 (14.49)	0.39 (0-0.79)

Table Continued

Target volume	Guideline	Mean cm ³ (standard deviation)	ICCAA
CTVn L2	ESTRO	13.97 (4.17)	0.27 (0-0.70)
	RTOG	16.31 (5.75)	0.46 (0.01-0.84)
CTVn L3	ESTRO	18.37 (7.35)	0.39 (0.02-0.79)
	RTOG	20.50 (8.36)	0.45 (0.04-0.83)
SCFn	ESTRO	16.42 (5.18)	0.55 (0.15-0.87)
	RTOG	24.89 (8.28)	0.40 (0-0.80)

ICC_{AA}, intra-class correlation coefficient-absolute agreement; CI, confidence interval

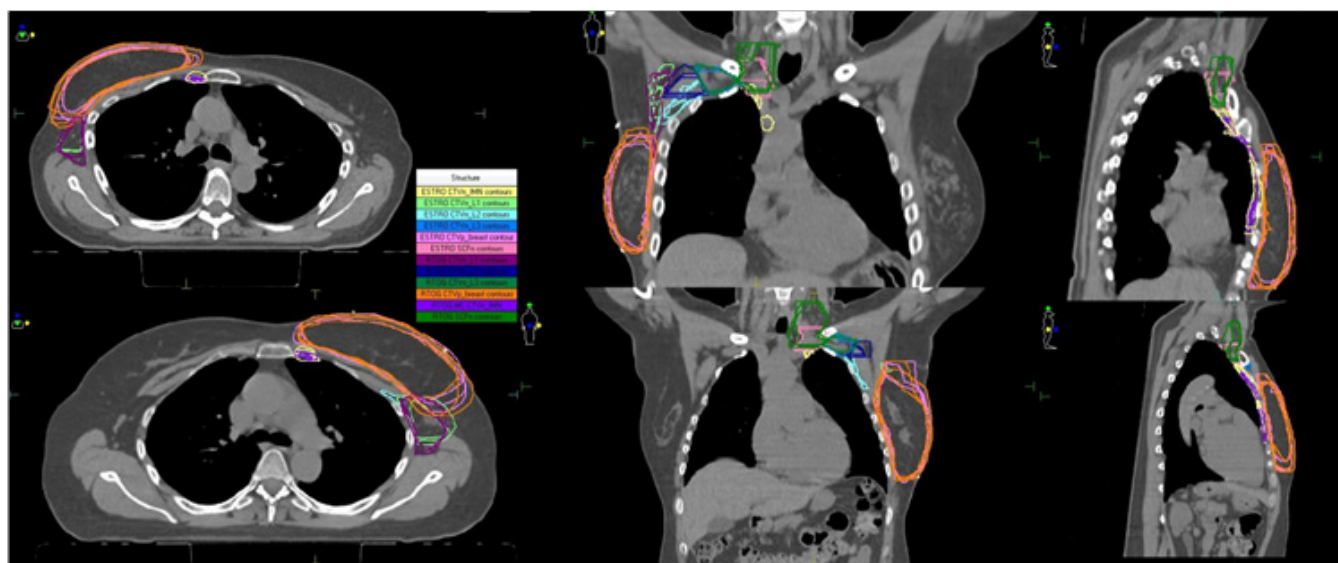


Figure 1 The target volumes produced using the ESTRO versus the RTOG guidelines in a right and left ESBC patient.

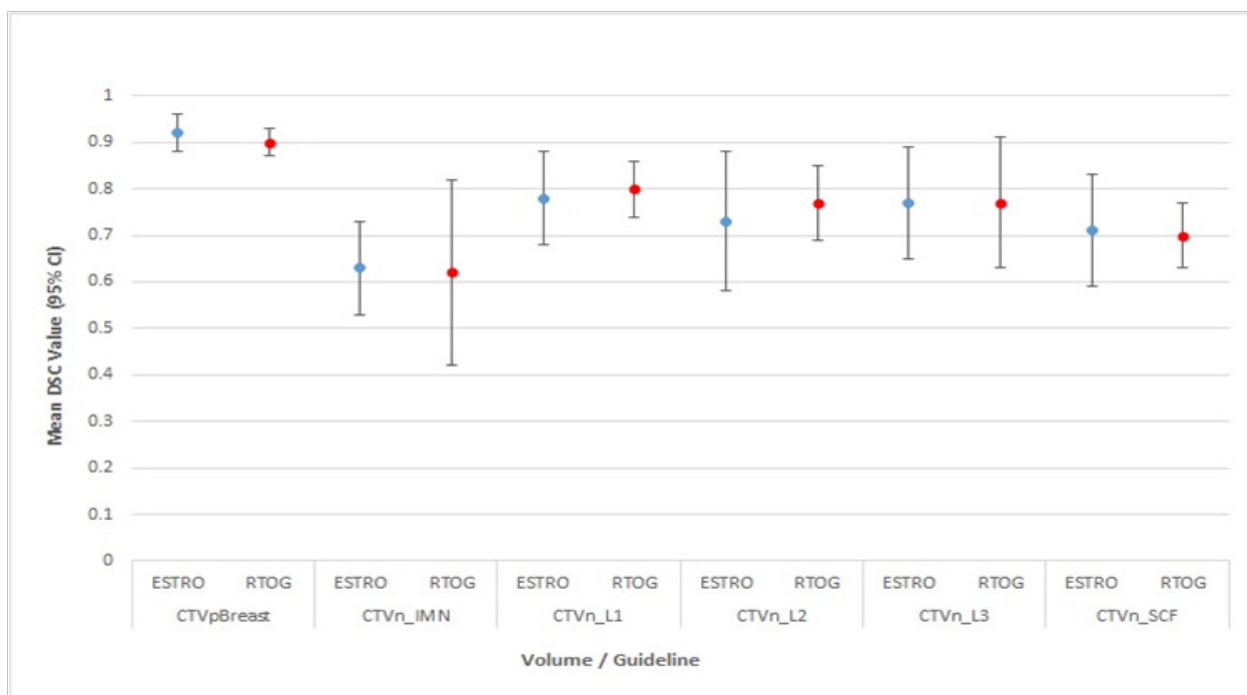


Figure 2 Mean DSC and 95% Confidence Intervals (CI) for contours created using the ESTRO and the RTOG guidelines.

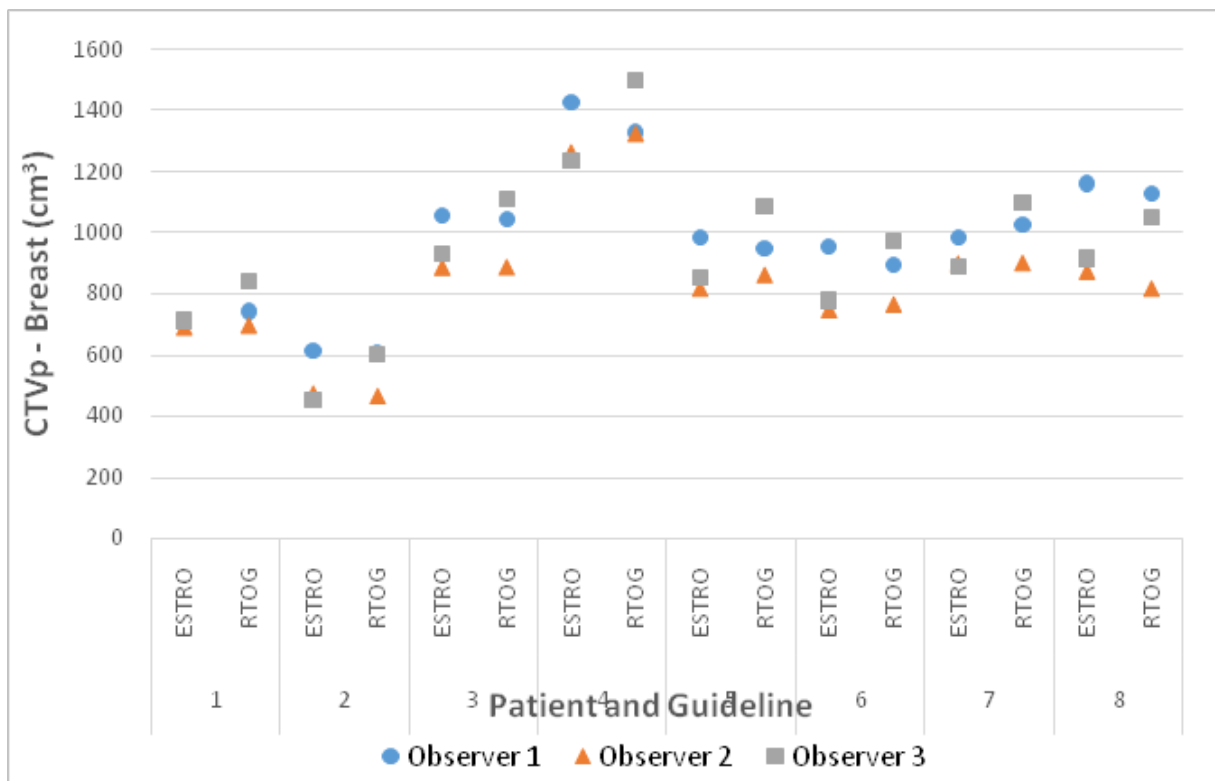


Figure 3A Total volumes for CTVp_Breast contoured by each observer for the respective guidelines.

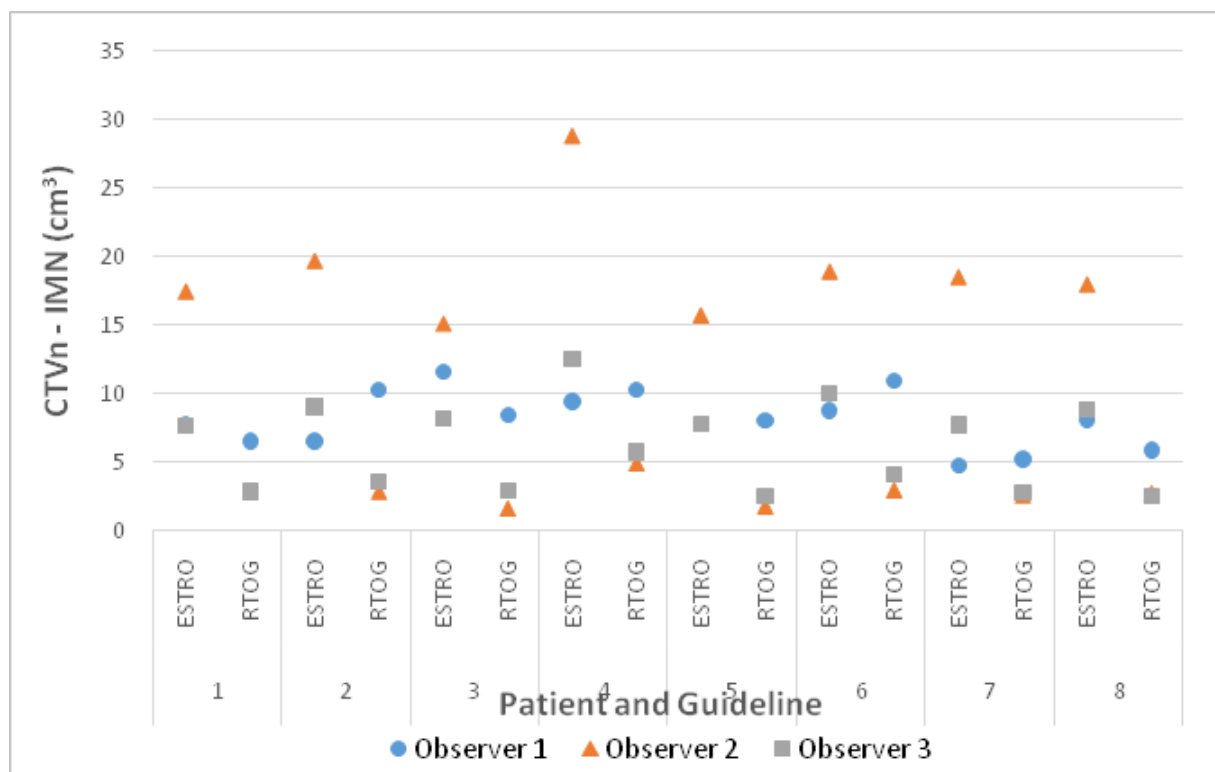


Figure 3B Total volumes for CTVn_IMN contoured by each observer for the respective guidelines.

Discussion

Despite the guidance available regarding breast radiotherapy,^{11,12,15,17} there is still variation in target volume delineation using these contouring guidelines.⁸ With improvement in accurate radiotherapy delivery such as IMRT and VMAT, target volume contouring variability is an important factor to address to reduce the risk of geometric inaccuracy.²¹ There have been previous studies comparing differing breast contouring guidelines.¹⁸ However, our study was purely aimed at the two major international guidelines to quantify if either guideline can better ensure consistency of contouring in ESBC. Even though the CTVp_Breast volume showed that ESTRO guidelines delivered more consistency than the RTOG guidelines, the difference was only small and unlikely to be of any clinical impact. A study of whole breast contouring consistency involving those trained in the RTOG guidelines showed a mean Jaccard coefficient of 86.2% (5.9%) between eight observers.⁷ Although this study by Yang considered a group consensus (or common volume) approach to estimating the Jaccard coefficient rather than an average of pairs approach. Therefore, it cannot be directly compared to our RTOG DSC of 0.90 (0.03) with three observers, as indicated in Kouwenhoven.²² Our results represent very good consistency within both guidelines for breast delineation. The anterior border in the RTOG guidelines is defined as skin but the atlas suggests a deduction from the skin that has not been defined, unlike the ESTRO guidelines which state ‘5mm under skin surface’, this could potentially lead to a small discrepancy in contouring within the RTOG contours. Overall, we can confidently say both guidelines led to consistent contouring of the CTVp_Breast volume. The Danish Breast Cancer Cooperative Group set consensus guidelines to be used for contouring of target volumes and OARs.¹⁵ This study assessed contours prior to the consensus guidelines and after the established guidance and used the DSC to assess the difference. When reviewing the difference for the axillary volumes prior to the guidelines and after, we can definitely see an improvement with the use of the guidelines. The mean DSC range prior to use of the guidelines was 0.56-0.65 and after using the guidelines, the mean DSC range was 0.70-0.76. The

range of mean DSC values for the axillary nodal levels in our study was 0.73-0.78 for the ESTRO guidelines and 0.77-0.80 for the RTOG guidelines. These results all suggest good concordance when using the above guidance when contouring the axillary volumes.

CTVn_IMN was noted as the least consistent contour using either guideline. This could be due to the use of the DSC to assess small contours. DSC is a less validated absolute measure for contour similarity for smaller volumes. No overlap in a few slices can result in a lower DSC value even though the exact position may not differ significantly.¹⁵ An interobserver variability study in head and neck cancer also calculated the DSC value to compare contours.²³ It noted that the clinical application of the DSC might be difficult given that a definite difference between a good and poor value is unknown and that the same absolute errors in larger volumes are more forgiving than for smaller volumes. Additional observers delineating the volumes would have provided additional statistical power to support our conclusions. There is no distinctly recognized technique by which to analyze radiotherapy contouring, and this study only focused on volume and DSC similarly to other studies addressing interobserver variability.^{15,22} A combination of other metrics may have been used such as centre of volume, dimension metrics or shape/surface variations. We felt this was unnecessary to answer our research question.

Conclusion

Ultimately, our aim of this study was to understand how best we could improve our quality of radiotherapy by ensuring consistency in target volume contouring given our advancing ability to in precise delivery techniques. Only the CTVp_Breast volume suggested greater RO consistency with the ESTRO guidelines, although the discrepancy was very small and unlikely to be of clinical significance. Overall, there was no evidence to suggest that using either the ESTRO or RTOG delineating guideline was superior in regards to greater consistency of delineating breast and nodal target volumes in ESBC patients (Appendix 1).

Appendix I Table of ESTRO and RTOG target contouring guideline

	RTOG guidelines	ESTRO guidelines
CTVp_Breast		
Cranial	Clinical reference + second rib insertion	Upper border of palpable/visible breast tissue; maximally up to the inferior edge of the sternoclavicular joint
Caudal	Clinical reference + loss of CT apparent breast	Most caudal slice with visible breast
Anterior	Skin	5 mm under skin surface
Posterior	Excludes pectoralis muscle, Chest wall muscles, ribs	Pectoralis major or costal and intercostal muscles where no muscle
Lateral	Clinical reference and mid-axillary line, typically excludes latissimus dorsi muscle	Lateral breast fold; anterior to the lateral thoracic artery
Medial	Sternal-rib junction	Lateral to the medial perforating mammary vessels; maximally to the edge of the sternal bone
CTVn_IMN		
Cranial	Superior aspect of the medial 1st rib.	Caudal limit of SCF

Table Continued

	RTOG guidelines	ESTRO guidelines
Caudal	Cranial aspect of the 4th rib	Cranial side of the 4th rib (in selected cases 5th rib on certain cases)
Anterior	encompass the internal mammary/ thoracic vessels	Anterior limit of the vascular area
Posterior	encompass the internal mammary/ thoracic vessels	Pleura
Lateral	encompass the internal mammary/ thoracic vessels	5 mm from the internal mammary vein and artery
Medial	encompass the internal mammary/ thoracic vessels	5 mm from the internal mammary vein and artery
SCFn		
Cranial	Caudal to cricoid cartilage	Includes the cranial extent of the subclavian artery (i.e. 5 mm cranial of subclavian vein)
Caudal	Junction of brachiocephalic-axillary veins/caudal edge clavicle head (superior border of breast field)	Includes the subclavian vein with 5 mm margin, thus connecting to the cranial border of CTVn_IMN
Anterior	Sternocleidomastoid muscle (SCM)	Sternocleidomastoid muscle, dorsal edge of the clavicle
Posterior	Anterior aspect scalene muscle	Pleura
Lateral	Cranial: lateral edge of SCM; caudal: junction first rib-clavicle	Includes the anterior scalene muscles and connects to the medial border of CTVn_L3
Medial	Exclude thyroid and trachea	Including the jugular vein without margin; excluding the thyroid gland and the common carotid artery
CTVn_L3		
Cranial	Axillary vessels cross lateral edge of pectoralis minor muscle	Cranial extent of the subclavian artery (i.e. 5 mm cranial of subclavian vein)
Caudal	Pectoralis major insert onto ribs	5 mm caudal to subclavian vein (or top ALND)
Anterior	Plane defined by anterior surface of pectoralis major and lat dorsi	Pectoralis major
Posterior	Anterior surface subscapularis muscle	Up to 5 mm dorsal of subclavian vein or to costae and intercostal muscles
Lateral	Medial border of latissimus dorsi	Medial side of the minor pectoral muscle
Medial	Lateral border pectoralis minor	Junction of subclavian and internal jugular veins → level 4
CTVn_L2		
Cranial	Axillary vessels cross medial edge of pectoralis minor muscle	Cranial extent of the axillary artery (i.e. 5 mm cranial of axillary vein)
Caudal	Axillary vessels cross lateral edge of pectoralis minor muscle	Caudal border of pectoralis minor muscle, (or top of ALND)
Anterior	Anterior surface pectoralis minor	Pectoralis minor
Posterior	Ribs and intercostal muscle	Up to 5 mm dorsal of axillary vein or to costae and intercostal muscles
Lateral	Lateral border pectoralis minor	Lateral border pectoralis minor
Medial	Medial border pectoralis minor	Medial border pectoralis minor
CTVn_L1		

Table Continued

	RTOG guidelines	ESTRO guidelines
Cranial	Axillary vessels cross lateral edge of pectoralis minor muscle	Medial: 5 mm cranial to the axillary vein Lateral: max up to 1 cm below the edge of the humeral head, 5 mm around the axillary vein
Caudal	Pectoralis major insert onto ribs; clinically at the base of the anterior axillary line	To the level of rib 4 and 5, taking also into account the visible effects of the sentinel lymph node biopsy
Anterior	Plane defined by anterior surface of pectoralis major and lat dorsi	Pectoralis major and minor muscles
Posterior	Anterior surface subscapularis muscle	Cranially up to the thoraco-dorsal vessels, and more caudally up to an imaginary line between the anterior edge of the latissimus dorsi muscle and the intercostal muscles
Lateral	Medial border of latissimus dorsi muscle	Cranially up to an imaginary line between the major pectoral and deltoid muscles, and further caudal up to a line between the major pectoral and latissimus dorsi muscles
Medial	Lateral border pectoralis minor muscle	Axillary level II, the interpectoral level and the thoracic wall

Ethical statement section

Ethics approval was obtained via Austin Health Human Research Ethics Committee (LNR/16/Austin/95)

- (i) We have observed the principles laid down in the most recent NHMRC Statement on Ethical Conduct in Human Research and The Declaration of Helsinki
- (ii) This Study's protocol was ethically reviewed and approved by the Office for Research, Austin Health, 145 Studley Road, Heidelberg, VIC 3084.
- (iii) This study did not involve patients. All details pertaining to patients have been anonymised.
- (iv) Human and animal testing was unnecessary in this study.

Acknowledgments

None.

Funding

None.

Conflicts of interest

Authir declares that there is no conflict of interest.

References

1. Van Maaren MC, de Munck L, de Bock GH, et al. 10 year survival after breast-conserving surgery plus radiotherapy compared with mastectomy in early breast cancer in the Netherlands: a population-based study. *Lancet Oncol.* 2016;17(8):1158–1170.
2. Budach W, Bolke E, Kammers K, et al. Adjuvant radiation therapy of regional lymph nodes in breast cancer - a meta-analysis of randomized trials- an update. *Radiat Oncol.* 2015;10:258.
3. Popescu CC, Olivotto IA, Beckham WA, et al. Volumetric modulated arc therapy improves dosimetry and reduces treatment time compared to conventional intensity-modulated radiotherapy for locoregional radiotherapy of left-sided breast cancer and internal mammary nodes. *Int J Radiat Oncol Bio Phys.* 2010;76(1):187–195.
4. Hurkmans CW, Borger JH, Pieters BR, et al. Variability in target volume delineation on CT scans of breast. *Int J Radiat Oncol Bio Phys.* 2001;50(5):366–372.
5. Struikmans H, Warlam-Rodenhuis C, Stam T, et al. Interobserver variability of clinical target volume delineation of glandular breast tissue and of boost volume in tangential breast irradiation. *Radiother Oncol.* 2005;76(3):293–299.
6. Dijkema IM, Hofman P, Raaijmakers CP, et al. Loco-regional conformal radiotherapy of the breast: delineation of the regional lymph node clinical target volumes in treatment position. *Radiother Oncol.* 2004;71(3):287–295.
7. Yang J, Woodward WA, Reed VK, et al. Statistical modelling approach to quantitative analysis of interobserver variability in breast contouring. *Int J Radiat Oncol Biol Phys.* 2014;89(1):214–221.
8. Van Mourik AM, Elkhuizen PH, Minkema D, et al. Multiinstitutional study on target volume delineation variation in breast radiotherapy in the presence of guidelines. *Radiother Oncol.* 2010;94(3):286–291.
9. 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(3):944–951.
10. Pitkänen MA, Holli KA, Ojala AT, et al. Quality Assurance in Radiotherapy of Breast Cancer. Variability in Planning Target Volume Delineation. *Acta Oncol.* 2001;40(1):50–55.
11. Radiation Therapy Oncology Group. *Breast cancer atlas for radiation therapy planning: consensus definitions.* 2013.
12. Gentile MS, Usman AA, Neuschler EI, et al. Contouring Guidelines for the Axillary Lymph Nodes for the Delivery of Radiation Therapy in Breast Cancer: Evaluation of the RTOG Breast Cancer Atlas. *Int J Radiat Oncol Biol Phys.* 2015;93:257–265.
13. Brown LC, Diehn FE, Boughey JC, et al. Delineation of Supraclavicular Target Volumes in Breast Cancer Radiation Therapy. *Int J Radiat Oncol Biol Phys.* 2015;92(3):642–649.
14. Jing H, Wang SL, Li J, et al. Mapping Patterns of Ipsilateral Supraclavicular Nodal Metastases in Breast Cancer: Rethinking the Clinical Target Volume for High-risk Patients. *Int J Radiat Oncol Biol Phys.* 2015;93(2):268–276.

15. Nielsen MH, Berg M, Pedersen AN, et al. Delineation of target volumes and organs at risk in adjuvant radiotherapy of early breast cancer: National guidelines and contouring atlas by the Danish Breast Cancer Cooperative Group. *Acta Oncol.* 2013;52(4):703–710.
16. Borm KJ, Voppichler J, Düsberg M, et al. FDG/PET-CT - Based Lymph Node Atlas in Breast Cancer Patients. *Int J Radiat Oncol Biol Phys.* 2019;103(3):574–582.
17. Offersen BV, Boersma LJ, Kirkove C, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. *Radiother Oncol.* 2015;114(1):3–10.
18. Gee HE, Moses L, Stuart K, et al. Contouring consensus guidelines in breast cancer radiotherapy: Comparison and systematic review of patterns of failure. *J Med Imaging Radiat Oncol.* 2019;63(1):102–115.
19. Carillo V, Cozzarini C, Perna L, et al. Contouring Variability Assessment Using a Generalized Concordance Index. *Int J Radiation Oncol Biol Phys.* 2012;84(3):841–846.
20. Bland M. *An introduction to medical statistics.* 2015.
21. Weiss E, Hess CF. The impact of gross tumour volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. *Strahlenther Onkol.* 2003;179(1):21–30.
22. Kouwenhoven E, Giezen M, Struikmans H. Measuring the similarity of target volume delineations independent of the number of observers. *Phys Med Biol.* 2009;54(9):2863–2873.
23. van der Veen J, Gulyban A, Nuyts S, et al. Interobserver variability in delineation of target volumes in head and neck cancer. *Radiother Oncol.* 2019;137:9–15.