The Breast 72 (2023) 103578

Contents lists available at ScienceDirect

The Breast

journal homepage: www.journals.elsevier.com/the-breast

Contouring variation affects estimates of normal tissue complication probability for breast fibrosis after radiotherapy

Tanwiwat Jaikuna [a,b], Eliana Vasquez Osorio [a], David Azria [c], Jenny Chang-Claude [d,e], Maria Carmen De Santis [f], Sara Gutiérrez-Enríquez [g], Marcel van Herk [a], Peter Hoskin [a], Maarten Lambrecht [h], Zoe Lingard [a], Petra Seibold [d], Alejandro Seoane [i], Elena Sperk [j], R Paul Symonds [k], Christopher J. Talbot [k], Tiziana Rancati [l], Tim Rattay [k], Victoria Reyes [m], Barry S. Rosenstein [n], Dirk de Ruysscher [o], Ana Vega [p,q,r], Liv Veldeman [s], Adam Webb [t], Catharine M.L. West [a], Marianne C. Aznar [a,*]

[a] Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Christie NHS Foundation Trust Hospital, Manchester, United Kingdom
[b] Division of Radiation Oncology, Department of Radiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[c] Department of Radiation Oncology, Montpellier Cancer Institute, Université Montpellier, Inserm, U1194, France
[d] Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[e] University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Germany
[f] Radiation Oncology, Fondazione IRCCS Isituto Nazionale dei Tumori, Milan, Italy
[g] Hereditary Cancer Genetics Group, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Hospital Campus, Barcelona, Spain
[h] KU Leuven, Department of Radiation Oncology, Leuven, Belgium
[i] Medical Physics Department, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain
[j] Department of Radiation Oncology, Mannheim Cancer Center, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany
[k] Leicester Cancer Research Centre, University of Leicester, United Kingdom
[l] Data Science Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[m] Radiation Oncology Department, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain
[n] Department of Radiation Oncology, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA
[o] Maastricht University Medical Center, Department of Radiation Oncology (Maastro Clinic), GROW School for Oncology and Developmental Biology, Maastricht, the Netherlands
[p] Fundación Pública Galega de Medicina Xenómica, Grupo de Medicina Xenómica (USC), Santiago de Compostela, Spain
[q] Instituto de Investigación Sanitaria de, Santiago de Compostela, Spain
[r] Biomedical Network on Rare Diseases (CIBERER), Spain
[s] Ghent University Hospital, Department of Radiation Oncology, Ghent, Belgium
[t] Department of Genetics and Genome Biology, University of Leicester, United Kingdom

ARTICLE INFO

Keywords:
Radiotherapy
Inter-observer variation
Late effects
Fibrosis
Breast cancer
NTCP modelling

ABSTRACT

Background: Normal tissue complication probability (NTCP) models can be useful to estimate the risk of fibrosis after breast-conserving surgery (BCS) and radiotherapy (RT) to the breast. However, they are subject to uncertainties. We present the impact of contouring variation on the prediction of fibrosis.

Materials and methods: 280 breast cancer patients treated BCS-RT were included. Nine Clinical Target Volume (CTV) contours were created for each patient: i) CTV_crop (reference), cropped 5 mm from the skin and ii) CTV_skin, uncropped and including the skin, iii) segmenting the 95% isodose ($Iso_{95\%}$) and iv) 3 different auto-contouring atlases generating uncropped and cropped contours (Atlas_skin/Atlas_crop). To illustrate the impact of contour variation on NTCP estimates, we applied two equations predicting fibrosis grade ≥ 2 at 5 years, based on Lyman-Kutcher-Burman (LKB) and Relative Seriality (RS) models, respectively, to each contour. Differences were evaluated using repeated-measures ANOVA. For completeness, the association between observed fibrosis events and NTCP estimates was also evaluated using logistic regression.

* Corresponding author. Radiotherapy Related Research, Department 58, The Christie NHS foundation trust, Wilmslow road, Manchester, M20 4BX, United Kingdom.
E-mail address: Marianne.aznar@manchester.ac.uk (M.C. Aznar).

https://doi.org/10.1016/j.breast.2023.103578
Received 30 June 2023; Received in revised form 25 August 2023; Accepted 8 September 2023
Available online 11 September 2023

Results: There were minimal differences between contours when the same contouring approach was followed (cropped and uncropped). CTV_skin and Atlas_skin contours had lower NTCP estimates (−3.92%, IQR 4.00, $p < 0.05$) compared to CTV_crop. No significant difference was observed for Atlas_crop and $Iso_{95\%}$ contours compared to CTV_crop. For the whole cohort, NTCP estimates varied between 5.3% and 49.5% (LKB) or 2.2% and 49.6% (RS) depending on the choice of contours. NTCP estimates for individual patients varied by up to a factor of 4. Estimates from "skin" contours showed higher agreement with observed events.
Conclusion: Contour variations can lead to significantly different NTCP estimates for breast fibrosis, highlighting the importance of standardising breast contours before developing and/or applying NTCP models.

1. Introduction

Radiotherapy plays a vital role in the management of breast cancer [1]. Especially for early-stage breast cancer, radiotherapy reduces the risk of local recurrence by half and contributes to a long-term reduction in breast cancer mortality [2–4]. However, several studies show that radiation can cause side-effects for healthy tissues such as skin, breast tissue, heart and lungs, which can negatively impact the patient's quality of life [5–10]. A common side-effect is the change in breast tissue structure after receiving radiation, this would encompass atrophy, nipple distortion, and also leading to breast fibrosis both in the tumour bed and outside the tumour bed [11]. Many studies attempt to establish models to predict the risk of radiation-induced breast fibrosis, often based on clinical parameters and/or dosimetric parameters for example dose volume histograms and dose surface histograms [12–17].

The robustness of the prediction model relies on many parameters, for instance, biological parameters and derived dosimetric parameters, and uncertainties associated with defining those parameters have been acknowledged by several studies. The robustness of the dosimetric parameters can be influenced by the dose calculation algorithm [18–20] and differences between planned and delivered dose including intra- and interfractional differences [21]. However, another recognized source of uncertainty in radiotherapy is inter-observer contouring variation, and its influence has rarely been estimated in modelling studies. The impact of contour variation on tumour control probability (TCP) and normal tissue complication probability (NTCP) estimation has been reported in the pelvis region [22], but not investigated in breast radiotherapy.

Inter-observer variations have been estimated to be large for breast contours in radiotherapy before the release of international guidelines such as European Society for Radiotherapy and Oncology (ESTRO) and Radiation Therapy Oncology Group (RTOG) [23,24]. For example, Li XA et al. (2009) found up to 42.5% deviation for breast contouring volumes between centres and individual observers. This inter-observer variation was reduced to around 10% after training and providing breast contouring guidelines [25,26]. Moreover, variations in contouring were also observed in organs at risk relevant to breast radiotherapy, such as the heart and lungs [27].

Additionally, it is also important to note a discrepancy in international guidelines. A key difference in recommendations for contouring breast clinical target volumes (CTVs) between RTOG and ESTRO is the inclusion of the skin. ESTRO guidelines recommend cropping the CTV 5 mm from the skin while in RTOG guidelines, the CTV goes up to the skin though the cropping 5 mm from the skin often happens when constructing the planning target volume (PTV).

In addition, even after guideline implementation, some variation remains, either from different interpretations of the guidelines, or from different workflows and contouring conventions between centres. Therefore, variations in CTV contouring between centres and observers are expected, especially in large retrospective cohorts [28,29] which can span several decades for data collection during which period clinical practice has evolved. This uncertainty in CTV definition could directly affect the accuracy of modelling estimation. The aim of this study is to investigate the effect of breast contour variation on the NTCP estimation of breast fibrosis (grade $\geq$2 at 5 years).

2. Material and methods

REQUITE (www.requite.eu) is a large prospective multi-national study aimed at identifying risk factors associated with radiation toxicity and providing the highest standardisation in terms of prospectively collected toxicity outcomes and clinical data (including planning CT, contours and 3D planned dose distribution). REQUITE recruited 2059 breast cancer patients treated with breast-conserving surgery and radiotherapy between 2014 and 2019 from 26 centres in 7 countries [30,31].

For this study, we used a subset of the REQUITE dataset consisting of 280 female patients, all treated with external beam radiotherapy in the supine position to the whole breast radiotherapy, with or without boost (i.e., excluding patients treated in prone position). Further selection criteria included: 1) the presence of a manually contoured CTV, and 2) no supraclavicular irradiation as it directly affected the $Iso_{95\%}$ contour generation (see section "contours"). No additional selection criteria were applied in this study. Patient and treatment characteristics are shown in Table 1.

Fibrosis was evaluated by a doctor or research nurse using questions derived from Common Terminology Criteria for Adverse Events version 4.0 (CTCAEv4.0) (*see Appendix A*). It was scored immediately after radiotherapy, at post-RT, 1-year post-RT, and 2-years post-RT and categorised into four levels: no induration, mild induration (able to slide and pinch up skin), moderate induration (able to slide skin but unable to pinch skin), and severe induration (unable to slide or pinch skin). Though fibrosis was scored separately for the tumour and outside the tumour bed, those scores were combined for the purpose of this analysis and the grade of fibrosis in the breast was defined as the highest value reported.

2.1. Contours

All patients had manually contoured breast CTVs: for some patients, the CTV was contoured up to the skin and the PTV was cropped by 5 mm. In others, the CTV itself was cropped 5 mm from the skin. To obtain consistent contours for the ipsilateral breast over the whole patient dataset, we retrospectively created two contours from the existing manual contour: i) CTV_crop, which was defined as the CTV cropped 5 mm from the skin and ii) CT_skin, which was uncropped and including the skin. To create CTV_skin in patients where the manual contour was already cropped, CTV_crop was automatically expanded laterally and anteriorly by 5 mm.

To simulate the impact of inter-observer variation, we generated additional breast contours for each patient using atlas-based auto-segmentation (RayStation v6.99), and an atlas created using contours from an independent observer (experienced breast clinical oncologist) in 20 patients ("template patients"). All template patients had both breasts contoured including skin (but otherwise following the ESTRO guidelines [23]) and were not part of the studied cohort. Three atlases were created varying the number of template patients (5, 10 and 20), and we refer to the generated ipsilateral breast contours as Atlas_skin1, Atlas_skin2 and Atlas_skin3. These contours were then cropped 0.5 cm from the skin to create Atlas_crop1, Atlas_crop2, and Atlas_crop3.

Finally, the last contour was created by segmenting the 95% isodose

Table 1
Patient and treatment characteristics of studied cohort, including 280 breast patients from the REQUITE data set.

Demographic	Mean (SD)
Age (year)	59.3 (±10.1)
Breast treatment site	143 Left breast (51%), 137 Right Breast (49%)
Breast cup size*	Cup size 1: 3 (1%), Cup size 2: 29 (10%), Cup size 3: 106 (38%), Cup size 4: 83 (30%), Cup size 5: 50 (18%), Cup size 6: 8 (2.8%), and NA: 1 (0.2%)
Breast band**	Band 2: 1 (0.2%), Band 3: 31 (11%), Band 4: 71 (25.3%), Band 5: 90 (32.1%), Band 6: 41 (15%), Band 7: 23 (8.2%), Band 8: 13 (5%), Band 9: 6 (2%), Band 10: 3 (1%), NA: 1 (0.2%)
Body Mass Index	25.4 (±4.5)
Staging	
T	Tis: 70 (25%), T1: 172 (61.4%), T2: 31 (11.4%), T3: 1 (0.2%), NA: 6 (2%)
N	N0 or NX: 271 (96.8%), N1: 5 (1.8%), NA: 4 (1.4%)
M	M0 or MX: 271 (96.8%), NA: 9 (3.2%)
Fibrosis events (CTCAEv4.0)	
None	94 (33.6%)
Grade 1	124 (44.3%)
Grade 2	54 (19.3%)
Grade 3	8 (2.8%)
Radiotherapy	
Technique	3D-CRT, Field in field or IMRT: 262 (93.6%), VMAT: 18 (6.4%)
Dose per fraction	Conventional fractionation (1.8–2.0Gy): 258 (92%), Hypofractionation (>2Gy): 22 (8%)
Boost status	Boost: 198 (71%), and No Boost: 82 (29%)

*Breast cup size was defined as 1 = AA; 2 = A; 3 = B; 4 = C; 5 = D; 6 = E, DD, and NA = not available.
**Breast band was defined as 1 = 28 (UK); 2 = 30 (UK); 3 = 32 (UK), 70 (EU), 85 (FR), 1 (IT); 4 = 34 (UK), 75 (EU), 90 (FR), 2 (IT); 5 = 36 (UK), 80 (EU), 95 (FR), 3 (IT); 6 = 38 (UK), 85 (EU), 100 (FR), 4 (IT); 7 = 40 (UK), 90 (EU), 105 (FR), 5 (IT); 8 = 42 (UK), 95 (EU), 110 (FR), 6 (IT); 9 = 44 (UK), 100 (EU), 115 (FR), 7 (IT); 10>above.

Table 2
Radiobiological parameters of NTCP estimate based on LKB and RS model for grade 1+ breast fibrosis at 5 years after radiotherapy [7].

Parameters	$EDU_3(50)$	α/β (Gy)	*m*	*n*	*γ*	*s*
LKB model	62.4	3	0.27	0.78	–	–
RS model	62.4	3	–	–	1.47	0.12

level of the prescribed dose ($Iso_{95\%}$). This was done to simulate an alternative approach for breast planning in clinical practice in several institutions, where a "PTV_eval" is generated from the placement of the tangential beams [32]. For the patient who was planned by simultaneous integrated boost (SIB) technique, the prescription dose for the whole breast irradiation will be normalised and used to be 100% of the prescription dose. The $Iso_{95\%}$ contour will be generated at 95% of the whole breast prescription.

In total, 9 contours were then available for each patient. Though those contours do not strictly speaking represent "inter-observer variation", they were used to provide an estimate of the impact of two types of contour variation: 1) systematic variation (different definitions e.g., skin vs crop vs $Iso_{95\%}$) and 2) random variation (e.g., between different atlases using the same definition, e.g., Atlas_crop1, Atlas_crop2, and Atlas_crop3). All contours were visually reviewed. Geometrical variations were investigated using the Dice similarity coefficient (DSC), mean distance to agreement (MDA), Hausdorff Distance (HD), and the difference of contour volume (ΔV), using CTV_crop as a reference.

2.2. Estimation of NTCP for grade 2 plus breast fibrosis

Differential dose-volume histogram (dDVH) of all breast contours were extracted from RayStation treatment planning system version 6.99. Published literature was reviewed for NTCP models for breast fibrosis. We identified only one such published model, modelling grade ≥2 toxicity at 5 years [15]. Using this model and its parameters, we calculated NTCP for each contour based on Lyman-Kutcher-Burman (LKB) and relative seriality (RS) models using pyRadiobiology [33, 34]. For completeness, we present these parameters in Table 2.

2.3. Statistics

The differences in breast volume and NTCP estimates was assessed using one-way repeated measure ANOVA. The association between the observed event of grade ≥1 breast fibrosis and NTCP estimates was tested using binary logistic regression with a 95% confidence interval in SPSS version 28 (a p-value of $p < 0.05$ was considered statistically significant).

3. Results

3.1. Contour evaluation

Fig. 1 shows an example of all contours generated for one example

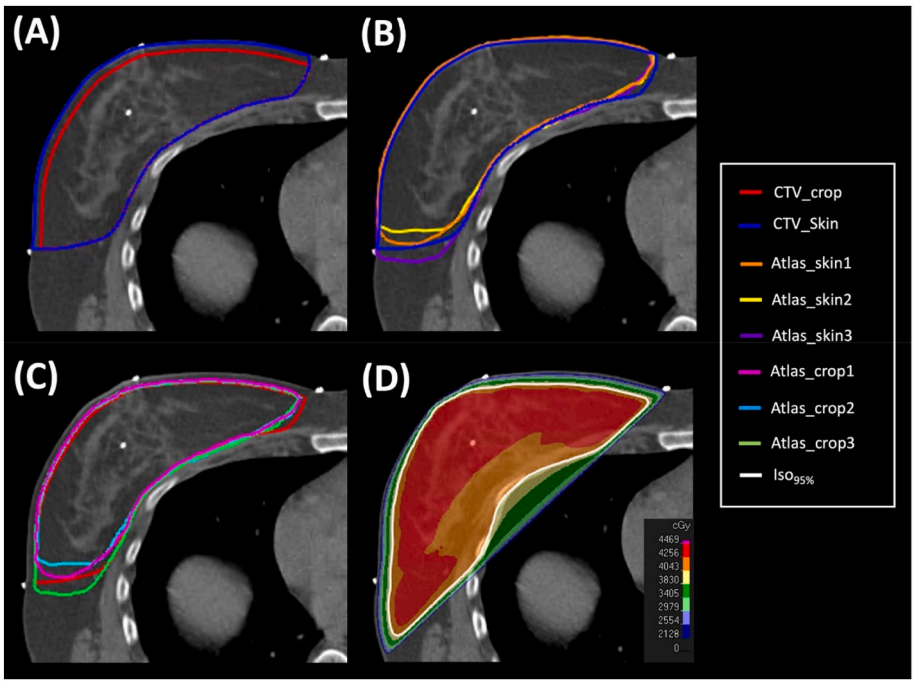

Fig. 1. Example of all breast contours used in this study. (A) CTV_crop and CTV_skin, (B) atlas generated contours including skin and (C) excluding skin, and (D) 95% isodose level of the prescribed dose ($Iso_{95\%}$).

patient (see the volume of difference contouring approach of the whole cohort in Appendix B). In atlas-generated contours, more variation was observed at the lateral border (Fig. 1B and C). CTV_crop volumes across the cohort ranged from 63.73 to 1913.96 cm^3 (Fig. 2A). The quantitative analysis of variations between the different breast contours in terms of volumes is shown in Fig. 2B, and volume similarity measures in Fig. 2C–E.

As expected, "crop" contours (CTV_crop and all Atlas_crop) are systematically smaller than "skin" contours (CTV_skin and all Atlas_skin) (Fig. 2A). The median volume difference between CTV_crop and CTV_skin was 165.18 (IQR 78.35) cm^3. The MDA for the contours including skin was 0.43 (IQR 0.21) cm compared to CTV_crop; confirming that the crop was correctly applied. A large systematic difference was observed in $Iso_{95\%}$ compared with CTV_crop, 256.47 (IQR 237.07) cm^3 of median volume difference and 0.72 (IQR 0.36) cm of MDA. We also found a difference between $Iso_{95\%}$ contours when generated from 3D-CRT and VMAT technique shown in Fig. 3A and B. The DSC and MDA of $Iso_{95\%}$ contour generated from 3D-CRT were 0.71 ± 0.10 and 0.79 ± 0.28 cm while the DSC was increased to 0.80 ± 0.09 and the MDA decreased to 0.49 ± 0.20 cm for VMAT. The 3D dose distribution for contour generation is shown in Fig. 3C-D for 3DCRT and VMAT planning technique.

Contour differences between CTV_crop and all Atlas_crop contours were small, with a median of breast volume difference of 94.4 (IQR 221.8) cm^3, MDA of 0.24 (IQR 0.14) cm and DSC of 0.82 ± 0.49. Similarly, the differences between CTV_skin and all Atlas_skin were small, with a median breast volume difference of 140.4 (IQR 260.9) cm^3, MDA of 0.27 (IQR 0.09) cm and DSC of 0.89 (IQR 0.03).

3.2. Normal tissue complication probability estimates

The NTCP estimates from LKB and RS model for breast fibrosis grade ≥2 at 5 years for all breast contours are shown in Fig. 4A and B. The "skin" contours yield a significant lower absolute NTCP estimates in both models (20.2%, IQR 6.2% and 16.3%, IQR 7.5% for LKB and RS model) than the "crop" contours (25.2%, IQR 7.3% and 22.1%, IQR 8.7% for LKB and RS model) ($p < 0.05$). The average absolute difference of the NTCP estimate between the "skin" contours and the "crop" for both models was 3.9% (IQR 4.2%). Nonetheless, the NTCP estimate from $Iso_{95\%}$ contour was 23.9% (IQR 4.0%) and 20.3% (IQR 4.9%) for LKB and RS models. This study found a slightly different NTCP estimate between $Iso_{95\%}$ and CTV_crop contour of around 0.1%, IQR 3.7% and −0.2%, IQR 4.2% for LKB and RS models.

The median NTCP from the LKB model of patients who experienced grade ≥1 breast fibrosis was 25.5% (IQR 6.8%) for the "crop" contours and 20.4% (IQR 6.1%) for the "skin" contours. For the RS model, these values were 22.4% (IQR 7.9%) and 16.6% (IQR 7.2%) for "crop" and "skin" contours. Both models estimated higher NTCP values in patients with an observed high fibrosis grade (*see*Appendix C). In addition, the median NTCP for patients without breast fibrosis was 24.5% (IQR 8.1%) for the "crop" contours and 19.8% (IQR 6.3%) for the "skin" contours from the LKB model, and 21.2% (IQR 9.9%) and 15.7% (IQR 7.6%) from the RS model, respectively (Fig. 6). For "skin" contours, both models showed a statistically significant associations between NTCP estimates and the observed event of breast fibrosis (Table 3). This was not observed for "crop" contours.

Averaged over the whole cohort, random contour variation had a modest impact on absolute NTCP estimates: the differences between the 3 Atlas_Crop contours and CTV_crop was 1.4% (IQR 4.0). However, for individual patients, the relative difference in NTCP estimates due to contour variation, both systematic and random, could be up to 80% (Fig. 4C and D). The correlation of NTCP estimated between the "skin" contours and the "crop" contours was plotted, shown in Fig. 5. Higher NTCP estimates from both models were obtained when applying the model to contours cropped from the skin. The absolute difference in NTCP between LKB and RS models derived from all contours was 3.5% (IQR 1.4).

4. Discussion

To the best of our knowledge, this is the first time the impact of contouring variation on NTCP estimation of breast fibrosis has been

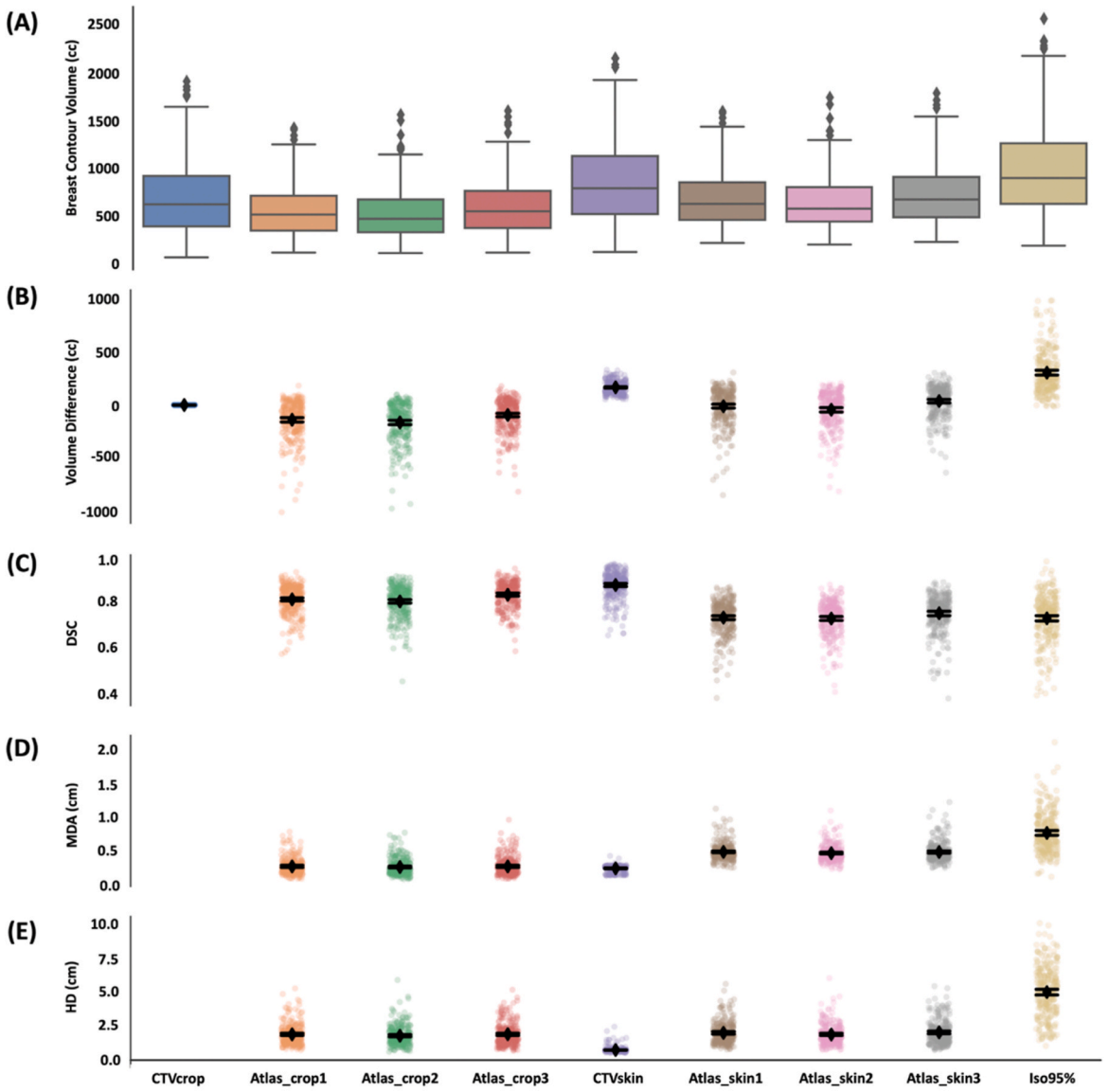

Fig. 2. Variation in contours across the cohort of 280 patients. (A) Breast volume of each different contour generation method extracted from Raystation, (B) The volume difference of simulated contours compared with CTV_crop (cm^3), (C) the DSC, (D) the MDA (in cm), and (E) the HD (in cm) for simulated contours compared with CTV_crop.

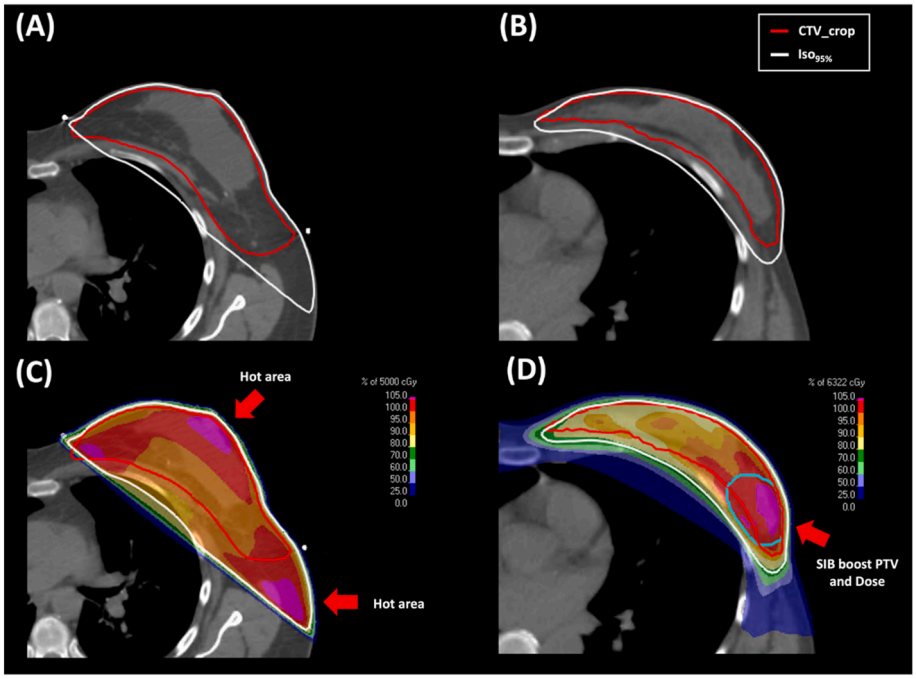

Fig. 3. An example of contour similarity between CTV_crop and $Iso_{95\%}$ when generated from (A) 3D-CRT and (B) Volumetric Arc Radiotherapy, and 3D dose distribution from (C) 3DCRT and (D) Volumetric Arc Radiotherapy planning technique.

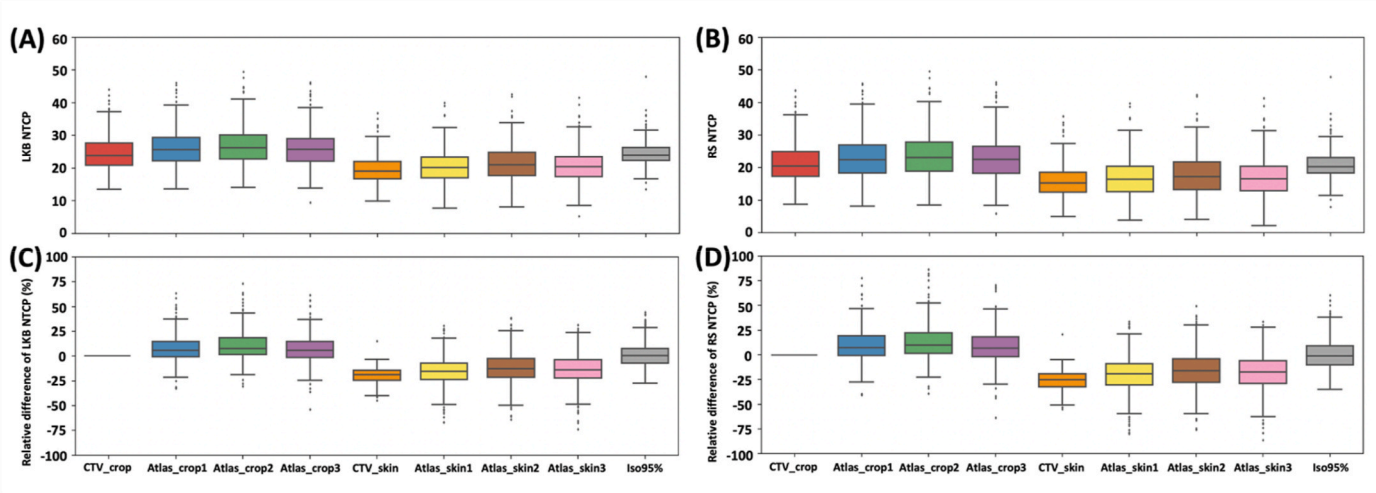

Fig. 4. The estimation of NTCP for grade 1+ breast fibrosis at 5 years after radiotherapy from (A) LKB model and (B) RS model. (C) The difference percentage of NTCP estimation of simulated contours compared with CTV_crop for LKB model, and (D) for RS model. The relative difference was calculated by $\frac{NTCP_{simulated\ contour} - NTCP_{CTV\_crop}}{NTCP_{CTV-crop}} \times 100$. (A) and (B) give an overview of the range of NTCP over the whole patient cohort, while (C) and (D) represents the range of the impact of contour variation on NTCP estimates for individual patients.

demonstrated. This influence was observed in two different NTCP models with fixed radiobiological parameters. It is well-recognized that NTCP models are subject to considerable uncertainties, e.g., could be up to 5–10% from dose shifting when using different dose calculation algorithms [20]. It should be noted that the accuracy of dose algorithms is limited at the surface region of the breast, including the skin, therefore NTCP estimates should be interpreted with caution. Large differences in relative NTCP estimates could be observed depending on biological parameters, for example the NTCP for heart toxicity is increased from 0.14% to 0.88% when applying around 5% uncertainty in D_{50} and 0.2% in γ *and s* into the NTCP estimation [35]. Compared to other accuracy requirements in radiotherapy, these remain small [36]. Though the impact of contouring uncertainty over the whole cohort might be small compared to these other sources, it is arguably easier to address, and therefore should be considered. Of note, published papers presenting a NTCP model may not always describe their contouring strategy [15]. In the case of our chosen model from Alexander et al. this data might not have been available, since the authors only mention the use of dose volume histograms in their analysis and their study pre-dates the publication of international contouring guidelines. However, moving forward, standardised contours, either

Table 3
The association between the observed event of breast fibrosis and NTCP estimates from difference contouring approach.

NTCP Variable	Odd ratio	*p* value
LKB-Skin	1.025	0.047*
LKB-Crop	1.017	0.138
RS-Skin	1.021	0.047*
RS-Crop	1.015	0.126

*Statistically significant at $p < 0.05$.

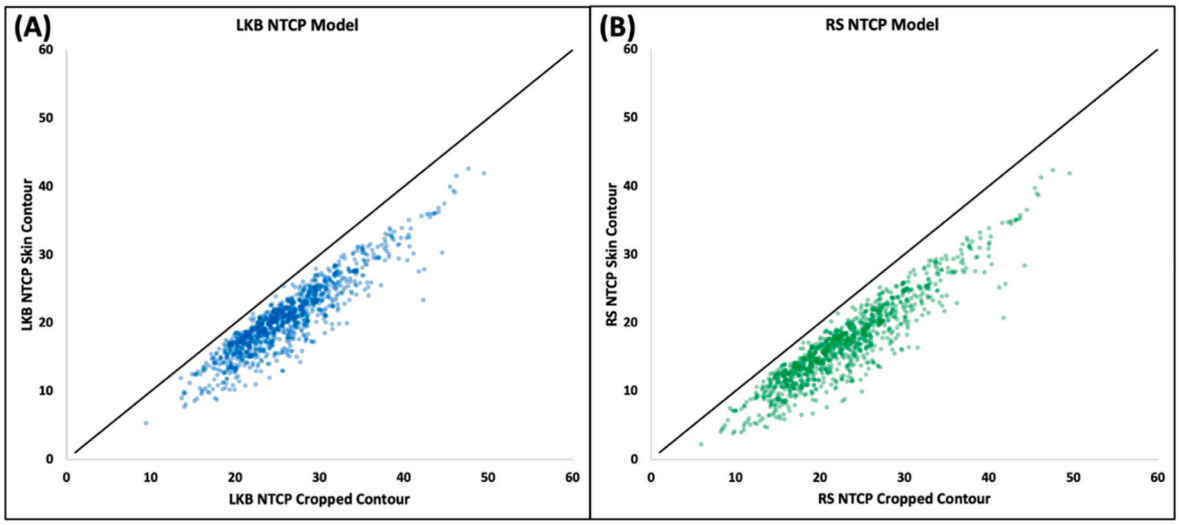

Fig. 5. The correlation of NTCP between contours including skin (CTV skin and all Atlas skin contours) and excluding skin (CTV crop and all Atlas crop contours) of (A) LKB model and (B) RS model.

manually created using guidelines or automatically generated following those guidelines, should be widely available to modellers and their use should be encouraged. Though our cohort size (280 patients) would be considered too small to develop or validate and NTCP model, our goal in this study was to raise awareness of the impact of contouring uncertainty. We hope model developers will provide information about the contours they used, and model users will reflect on whether a published NTCP model used comparable contours to their studied population.

Target delineation is an essential process for treatment planning purposes. While the main objective of the CTV is to provide the basis for safe and effective radiotherapy delivery, it should be acknowledged that contours will also be used in retrospective evaluations. As clinical practices evolve, large retrospective studies include patients treated over many years, and using different contouring conventions. Hence systematic variations in CTV contouring (for example cropping the CTV or not) are likely to be present in multi-centre studies, and also in studies from a single institution spanning many years. We have demonstrated that including or excluding skin led to considerable variations in NTCP estimates, especially in 3DCRT when the hot spot is often situated in the beam entrance and the breast dome's ventral region (Fig. 3C). In addition, variations between contours often occur at the entry points of medial and lateral tangential beams. Excluding this high dose region could impact NTCP estimates for breast fibrosis because the hot spot volume and location are linked to chronic lymphoedema as well as to moist desquamation, which are linked to fibrosis development [37–39].

Auto-contouring is an attractive solution for clinical practice: it can potentially save time and reduce inter-observer variation [40–42]. It is also a logical solution for large retrospective studies, where re-generating contours pre-analysis could provide a form of standardisation. However, auto-contouring solutions, either atlas-based or using machine learning (ML), rely on accurate manual delineations for training and validation. It is worth noting that it is challenging to train an auto-contouring solution on cropped contours: the skin is an easily identifiable boundary, while "5 mm below the skin" is not. As a result, auto-contouring solutions may have more success training on CTV_skin contours. It is worth noting that, although re-creating a "CTV_skin" from a cropped contour is possible, it is not straightforward. As such, it might be preferable to save both CTV_skin and CTV_crop types of contours for the purpose of future data analysis.

Our study has several limitations. Though we simulated inter-observer variations using different atlases, the dose distribution was not altered to reflect the new generated contours. Hence, this might underestimate the full impact of inter-observer variation in clinical practice. Also, we used NTCP parameters from a single and older study. We are also limited by the choice of toxicity grade ($\geq$1) modelled in this study: though it may not be the most clinically relevant endpoint, it ensures a sufficient number of events in our dataset. A strength of this study is the use of a prospectively collected dataset from several institutions and a range of clinical practices.

In conclusion, the difference in breast contour definition and inter-observer variation influences the NTCP estimation. This also suggests that building NTCP models from inconsistent datasets (with mixed contour definitions or non-adherence to guidelines) could reduce model robustness and lead to considerable variation in the predicted risk of

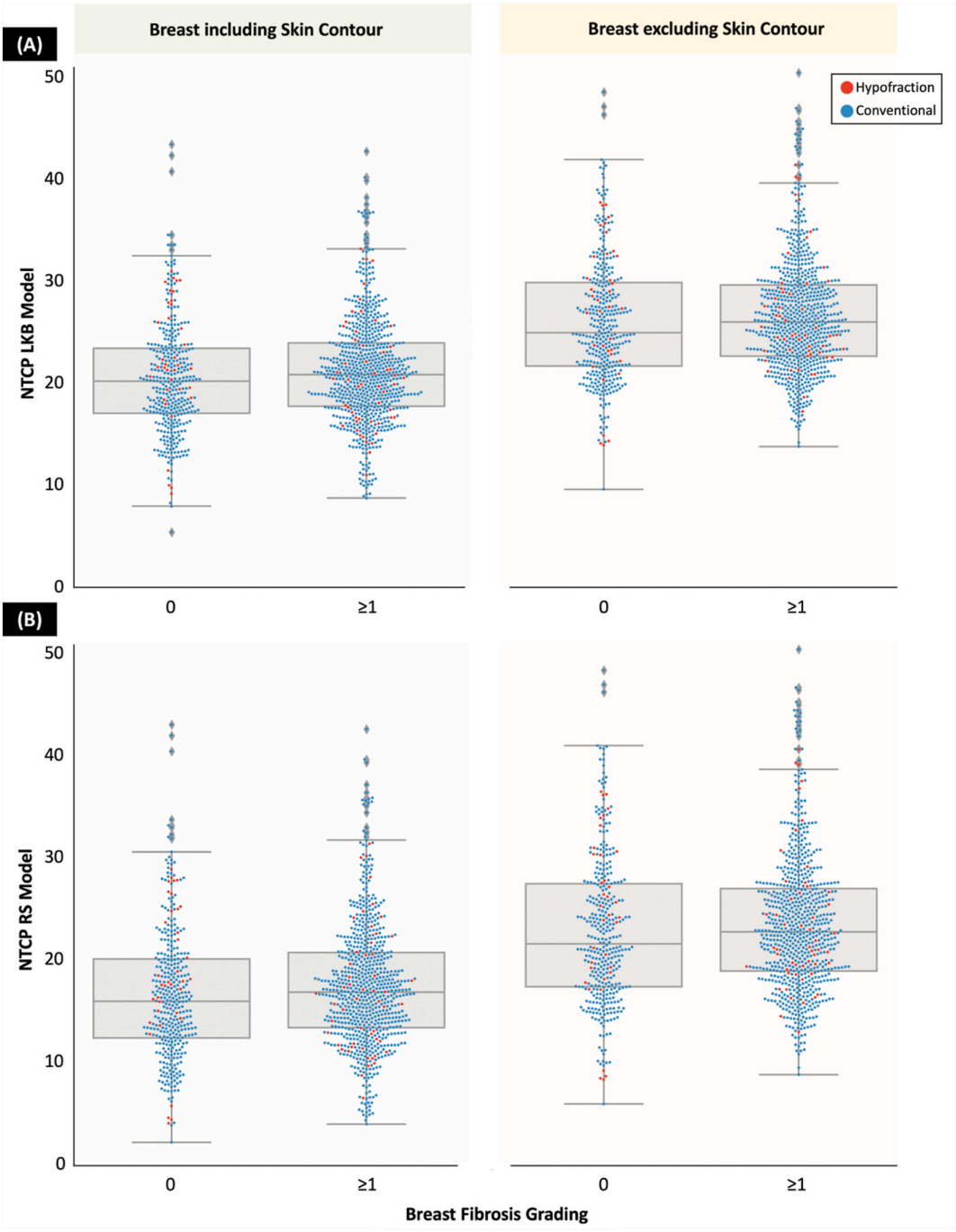

Fig. 6. The NTCP estimates of difference fractionation scheme from LKB (A) and RS (B) models for including (skin) and excluding skin (Crop) contours of non-fibrosis cohort and fibrosis grade ≥1 cohort (*see also*Appendix B).

breast fibrosis. Therefore, a clear breast contour definition is essential and should be determined before generating and applying the model in the clinic.

Acknowledgements

REQUITE received funding from the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no. 601826.

We thank all patients who participated in the REQUITE study and all study personnel involved in the REQUITE project.

Marianne Aznar acknowledges the support of the Engineering and Physical Sciences Research Council (Grant number EP/T028017/1)

This work was supported by Cancer Research UK RadNet Manchester [C1994/A28701] and the NIHR Manchester Biomedical Research Centre (NIHR203308).

The researchers at DKFZ also thank Anusha Müller, Irmgard Helmbold, Thomas Heger, Sabine Behrens, Juan Camilo Rosas. Petra Seibold was supported by ERA PerMed 2018 funding (BMBF #01KU1912) and BfS funding (#3619S42261).

S. Gutiérrez-Enríquez is supported by the Government of Catalonia 2021SGR01112.

The VHIO authors acknowledge the Cellex Foundation for providing research equipment and facilities and thank CERCA Program/Generalitat de Catalunya for institutional support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.breast.2023.103578.

References

[1] Swanick CW, Smith BD. Indications for adjuvant radiation therapy in breast cancer: a review of the evidence and recommendations for clinical practice. Chin Clin Oncol 2016 Jun;5(3):38.
[2] Bijker N, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma-in-situ: ten-year results of European organisation for research and treatment of cancer randomized phase III trial 10853–a study by the EORTC breast cancer cooperative Group and EORTC radiotherapy Group. J Clin Oncol 2006;24 (21):3381–7.
[3] Joshi SCKF, Pant I, Shukla AN. Role of radiotherapy in early breast cancer: an overview. Int J Health Sci 2007;1(2):259–64.
[4] Giannakeas V, Sopik V, Narod SA. Association of radiotherapy with survival in women treated for ductal carcinoma in situ with lumpectomy or mastectomy. JAMA Netw Open 2018 Aug 3;1(4):e181100.
[5] Haviland JS, Owen JR, Dewar JA, et al. 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(11):1086–94.
[6] Murray Brunt A, Haviland JS, Wheatley DA, et al. Hypofractionated breast radiotherapy for 1 week versus 3 weeks (FAST-Forward): 5-year efficacy and late normal tissue effects results from a multicentre, non-inferiority, randomised, phase 3 trial. Lancet 2020;395(10237):1613–26.
[7] Whelan TJ, Pignol JP, Levine MN, et al. Long-term results of hypofractionated radiation therapy for breast cancer. N Engl J Med 2010;362(6):513–20.
[8] Offersen BV, Alsner J, Nielsen HM, et al. Hypofractionated versus standard fractionated radiotherapy in patients with early breast cancer or ductal carcinoma in situ in a randomized phase III trial: the DBCG HYPO trial. J Clin Oncol 2020;38 (31):3615–25.
[9] Yarnold J, Ashton A, Bliss J, et al. Fractionation 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(1):9–17.
[10] Darby SC, Ewertz M, McGale P, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med 2013;368(11):987–98.
[11] Yi A, Kim HH, Shin HJ, Huh MO, Ahn SD, Seo BK. Radiation-induced complications after breast cancer radiation therapy: a pictorial review of multimodality imaging findings. Korean J Radiol 2009 Sep-Oct;10(5):496–507.
[12] Hammer C, Maduro JH, Bantema-Joppe EJ, et al. Radiation-induced fibrosis in the boost area after three-dimensional conformal radiotherapy with a simultaneous integrated boost technique for early-stage breast cancer: a multivariable prediction model. Radiother Oncol 2017 Jan;122(1):45–9.
[13] Mukesh MB, Harris E, Collette S, et al. Normal tissue complication probability (NTCP) parameters for breast fibrosis: pooled results from two randomised trials. Radiother Oncol 2013 Aug;108(2):293–8.
[14] Avanzo M, Stancanello J, Trovo M, et al. Complication probability model for subcutaneous fibrosis based on published data of partial and whole breast irradiation. Phys Med 2012 Oct;28(4):296–306.
[15] Alexander MA, Brooks WA, Blake SW. Normal tissue complication probability modelling of tissue fibrosis following breast radiotherapy. Phys Med Biol 2007 Apr 7;52(7):1831–43.
[16] Jothy Basu KS, Amit Bahl, Subramani V, et al. Normal tissue complication probability of fibrosis in radiotherapy of breast cancer: accelerated partial breast irradiation vs conventional external-beam radiotherapy. J Cancer Res Therapeut 2008;4(3):126–30.
[17] Palma G, Cella L. A new formalism of Dose Surface Histograms for robust modeling of skin toxicity in radiation therapy. Phys Med 2019 Mar;59:75–8.
[18] Wiklund K, Toma-Dasu I, Lind BK. The influence of dose heterogeneity on tumour control probability in fractionated radiation therapy. Phys Med Biol 2011 Dec 7;56 (23):7585–600.
[19] Wiklund K, Toma-Dasu I, Lind BK. Impact of dose and sensitivity heterogeneity on TCP. Comput Math Methods Med 2014;2014:182935.
[20] Chaikh A, Ojala J, Khamphan C, et al. Dosimetrical and radiobiological approach to manage the dosimetric shift in the transition of dose calculation algorithm in radiation oncology: how to improve high quality treatment and avoid unexpected outcomes? Radiat Oncol 2018 Apr 3;13(1):60.
[21] Buranaporn P, Dankulchai P, Jaikuna T, et al. Relation between DIR recalculated dose based CBCT and GI and GU toxicity in postoperative prostate cancer patients treated with VMAT. Radiother Oncol 2021 Apr;157:8–14.
[22] Fiorino C. Uncertainties in volume definition: impact on NTCP/TCP estimations. Phys Med 2001 Jan;17:45–50.
[23] 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 Jan;114(1):3–10.
[24] White J, Tai A, Arthur D. Breast cancer atlas for radiation therapy planning: consensus definitions. In: Book breast cancer atlas for radiation Therapy planning. Consensus Definitions. RTOG; 2009.
[25] van Pelt VWJ, Gerrets S, Simoes R, et al. Evaluation of delineating the target volume by radiation therapists in breast cancer patients. Technical Innovations & Patient Support in Radiation Oncology 2021 Mar;17:78–81.
[26] Mathew T, Chao M, Lapuz C, et al. Consistency of ESTRO and RTOG contouring guidelines for target volume delineation in early stage breast cancer. Int J Real Ther 2020;7(5):133–40.
[27] 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 Mar 1;73(3):944–51.
[28] Kim KS, Cheong KH, Kim K, et al. Interobserver variability in clinical target volume delineation in anal squamous cell carcinoma. Sci Rep 2021 Feb 2;11(1):2785.
[29] Joskowicz L, Cohen D, Caplan N, et al. Inter-observer variability of manual contour delineation of structures in CT. Eur Radiol 2019 Mar;29(3):1391–9.
[30] West C, Azria D, Chang-Claude J, et al. The REQUITE project: validating predictive models and biomarkers of radiotherapy toxicity to reduce side-effects and improve quality of life in cancer survivors. Clin Oncol 2014 Dec;26(12):739–42.
[31] Seibold P, Webb A, Aguado-Barrera ME, et al. REQUITE: a prospective multicentre cohort study of patients undergoing radiotherapy for breast, lung or prostate cancer. Radiother Oncol 2019 Sep;138:59–67.
[32] Msika R, Tkatchenko N, Robilliard M, et al. Evaluation of a software for automatic delineation of the mammary gland and organs at risk in patients treated for breast cancer in lateral position. Cancer Radiother 2020 Dec;24(8):799–804.
[33] Tsougos I, Grout I, Theodorou K, et al. A free software for the evaluation and comparison of dose response models in clinical radiotherapy (DORES). Int J Radiat Biol 2009 Mar;85(3):227–37.
[34] Germany: sachpazidis I. https://www.sachpazidis.com. https://www.sachpazidis.com/software-projects/pyradiobiology/ [internet].
[35] Sachpazidis I, González LB, Ferrando JR, et al. Breast cancer normal tissue complication modelling and parameter uncertainties. Advances in Smart Systems Research 2012;6(1):27–34.
[36] Van der Merwe D, Van Dyk J, Healy B, et al. Accuracy requirements and uncertainties in radiotherapy: a report of the international atomic energy agency. Acta Oncol 2017 Jan;56(1):1–6.
[37] Ashforth K, Morgner S, L V. A new treatment for soft tissue fibrosis in the breast. Journal of Lymphoedema 2011;6(2):5.
[38] Pignol JP, Truong P, Rakovitch E, et al. Ten years results of the Canadian breast intensity modulated radiation therapy (IMRT) randomized controlled trial. Radiother Oncol 2016 Dec;121(3):414–9.
[39] Borm KJ, Loos M, Oechsner M, et al. Acute radiodermatitis in modern adjuvant 3 D conformal radiotherapy for breast cancer - the impact of dose distribution and patient related factors. Radiat Oncol 2018 Nov 7;13(1):218.
[40] Reed VK, Woodward WA, Zhang L, et al. Automatic segmentation of whole breast using atlas approach and deformable image registration. Int J Radiat Oncol Biol Phys 2009 Apr 1;73(5):1493–500.
[41] Eldesoky AR, Yates ES, Nyeng TB, et al. Internal and external validation of an ESTRO delineation guideline - dependent automated segmentation tool for loco-regional radiation therapy of early breast cancer. Radiother Oncol 2016 Dec;121 (3):424–30.
[42] Lim JY, Leech M. Use of auto-segmentation in the delineation of target volumes and organs at risk in head and neck. Acta Oncol 2016 Jul;55(7):799–806.