Physics and Imaging in Radiation Oncology 32 (2024) 100635

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

Optimising inter-patient image registration for image-based data mining in breast radiotherapy

Tanwiwat Jaikuna [a,b], Fiona Wilson [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], Lea Kotzki [h], Maarten Lambrecht [i], Zoe Lingard [a], Petra Seibold [d], Alejandro Seoane [j], Elena Sperk [k], R Paul Symonds [l], Christopher J. Talbot [l], Tiziana Rancati [m], Tim Rattay [l], Victoria Reyes [n], Barry S. Rosenstein [o], Dirk de Ruysscher [p], Ana Vega [q,r,s], Liv Veldeman [t], Adam Webb [u], Catharine ML West [a], Marianne C Aznar [a,1], Eliana Vasquez Osorio [a,1,*]

[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] *University Federation of Radiation Oncology of Mediterranean Occitanie, Montpellier Cancer Institute ICM, Université Montpellier, INSERM 1194 IRCM, Montpellier, 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] *University Federation of Radiation Oncology of Mediterranean Occitanie, Gard Cancer Institute ICG, CHU Caremeau, Nimes, France*
[i] *KU Leuven, Department of Radiation Oncology, Leuven, Belgium*
[j] *Medical Physics Department, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*
[k] *Department of Radiation Oncology, Mannheim Cancer Center, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany*
[l] *Leicester Cancer Research Centre, University of Leicester, United Kingdom*
[m] *Data Science Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*
[n] *Radiation Oncology Department, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*
[o] *Department of Radiation Oncology, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, NY, USA*
[p] *Maastricht University Medical Center, Department of Radiation Oncology (Maastro Clinic), GROW School for Oncology and Developmental Biology, Maastricht, the Netherlands*
[q] *Fundación Pública Galega de Medicina Xenómica, Grupo de Medicina Xenómica (USC), Santiago de Compostela, Spain*
[r] *Instituto de Investigación Sanitaria de, Santiago de Compostela, Spain*
[s] *Biomedical Network on Rare Diseases (CIBERER), Spain*
[t] *Ghent University Hospital, Department of Radiation Oncology, Ghent, Belgium*
[u] *Department of Genetics and Genome Biology, University of Leicester, United Kingdom*

ARTICLE INFO

Keywords:
Breast radiotherapy
Image registration
Image-based data mining
Spatial normalisation

ABSTRACT

Background and purpose: Image-based data mining (IBDM) requires spatial normalisation to reference anatomy, which is challenging in breast radiotherapy due to variations in the treatment position, breast shape and volume. We aim to optimise spatial normalisation for breast IBDM.
Materials and methods: Data from 996 patients treated with radiotherapy for early-stage breast cancer, recruited in the REQUITE study, were included. Patients were treated supine (n = 811), with either bilateral or ipsilateral arm(s) raised (551/260, respectively) or in prone position (n = 185). Four deformable image registration (DIR) configurations for extrathoracic spatial normalisation were tested. We selected the best-performing DIR configuration and further investigated two pathways: *i*) registering prone/supine cohorts independently and *ii*) registering all patients to a supine reference. The impact of arm positioning in the supine cohort was quantified.

* Corresponding author at: Radiotherapy Related Research, Paterson Building, The Christie NHS Foundation Trust, Wilmslow Road, Manchester M20 4BX, United Kingdom.
E-mail address: eliana.vasquezosorio@manchester.ac.uk (E. Vasquez Osorio).
[1] Shared senior position.

https://doi.org/10.1016/j.phro.2024.100635
Received 20 July 2024; Received in revised form 19 August 2024; Accepted 20 August 2024
Available online 7 September 2024
2405-6316/© 2024 Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

DIR accuracy was estimated using Normalised Cross Correlation (NCC), Dice Similarity Coefficient (DSC), mean Distance to Agreement (MDA), 95 % Hausdorff Distance (95 %HD), and inter-patient landmark registration uncertainty (ILRU).

Results: DIR using B-spline and normalised mutual information (NMI) performed the best across all evaluation metrics. Supine-supine registrations yielded highest accuracy (0.98 ± 0.01, 0.91 ± 0.04, 0.23 ± 0.19 cm, 1.17 ± 1.18 cm, 0.51 ± 0.26 cm for NCC, DSC, MDA, 95 %HD, and ILRU), followed by prone-prone and supine-prone registrations. Arm positioning had no significant impact on registration performance. For the best DIR strategy, uncertainty of 0.44 and 0.81 cm in the breast and shoulder regions was found.

Conclusions: B-spline algorithm using NMI and registered supine and prone cohorts independently provides the most optimal spatial normalisation strategy for breast IBDM.

1. Introduction

Radiotherapy is an important treatment modality for breast cancer treatments. As the survival rate improves [1], long-term side effects become increasingly important to address. Adverse effects from radiotherapy, such as skin toxicity, breast fibrosis, lymphoedema, and shoulder stiffness, can cause treatment delays and negatively impact patients' quality of life [2,3]. Finding dose-outcome relationships can help clinicians prevent unwanted toxicity following breast radiotherapy.

Dose-outcome models in radiotherapy traditionally use simplified representations of dose distributions derived from dose-volume histograms (DVHs) or dose surface histograms (DSHs). However, this approach has some limitations, e.g., it requires a predefined hypothesis about structures or organs at risk (OARs) related to the investigated toxicity, where contour delineation uncertainties can bias prediction outcomes [4]. Generating a toxicity model from DVHs or DSHs [5–7] does not account for the heterogeneous dose distribution within OARs, as it summarises the complete 3D dose distribution to a single value, i.e., mean dose. Image-based data mining (IBDM), a voxel-based analysis technique, can help overcome these limitations [8]. IBDM involves analysing all voxels of the dose distribution and generating a spatial map to describe the dose–response relationship, including other confounding variables directly or through post-hoc analysis [9]. This is particularly relevant in breast cancer, where multimodal treatment such as surgery, chemotherapy, and hormonal therapy, as well as patient factors, e.g., comorbidities and body mass index (BMI), play a role in toxicity [10,11].

IBDM relies on spatial normalisation of the dose distribution to a common coordinate system (CCS), typically a reference template using Deformable Image Registration (DIR) before voxel-wise statistical analysis is performed [12]. While IBDM has been successfully applied to various tumour sites, including the lung, oesophagus, head and neck, and prostate cancer [13–17], its application in breast cancer remains largely unexplored. The dose–response relationships for many side effects of breast radiotherapy, such as lymphoedema and shoulder mobility, are poorly understood and may involve multiple structures or sub-structures. Therefore, implementing IBDM has the potential to advance our knowledge in this field.

However, spatial normalisation for breast radiotherapy presents significant challenges. There are considerable variations in breast size and shape between patients. Additionally, patients are treated in different positions: either prone or supine, the latter with either both arms raised or only the ipsilateral arm raised. These variations could affect the accuracy of spatial normalisation and the identification of sensitive sub-regions. This study aims to evaluate the accuracy of the spatial normalisation process and establish an optimal DIR strategy.

2. Methods

2.1. Patient dataset

Data from 996 early-stage breast cancer patients treated with radiotherapy after breast-conserving surgery, with/without systemic therapy [18], were included in this study, Table 1. These patients were recruited prospectively in the REQUITE project (www.requite.eu) from 9 institutes in 5 countries. REQUITE established a resource for multinational validation of models and biomarkers that predict the risk of late toxicity following radiotherapy [18]. REQUITE was approved by local ethics committees in participating countries (UK NRES Approval 14/NW/0035) and registered at https://www.controlled-trials.com (ISRCTN98496463) [19].

Due to the lack of anatomical landmarks and consistent contours in the full cohort, a smaller "representative cohort" (160 patients) was identified (*see section Identifying the Optimal Image Registration Pathway*). The purpose of the representative cohort is two-fold: 1) it allowed an in-depth exploration of the image registration performance, and 2) it

Table 1
The characteristics of 996 breast cancer patients recruited from the REQUITE dataset.

Characteristics	Supine	Prone
Number of patients	811	185
Age (years)	59 (23–90)	59 (35–79)
Treatment site (%)	Left = 52.6 %; Right = 47.4 %	Left = 48.6 %; Right = 51.4 %
Breast volume (cm^3)	670 (40–5450)	664 (107–1869)
Breast cup size (%)*	1 = 6.0 %; 2 = 8.6 %; 3 = 38.3 %; 4 = 27.2 %; 5 = 18.7 %; 6 = 5.4 %; 7 = 1.3 %; 8 = 0.0 %; 9 = 0.0 %; ≥ 10 = 0.0 %	1 = 0.6 %; 2 = 5.0 %; 3 = 29.3 %; 4 = 30.4 %; 5 = 26.5 %; 6 = 6.6 %; 7 = 0.6 %; 8 = 0.6 %; 9 = 0.6 %; ≥ 10 = 0.0 %
Breast band (%)**	1 = 0.0 %; 2 = 0.1 %; 3 = 8.2 %; 4 = 19.9 %; 5 = 26.0 %; 6 = 18.1 %; 7 = 9.8 %; 8 = 9.9 %; 9 = 3.5 %; 10 = 4.4 %	1 = 0.6 %; 2 = 0.0 %; 3 = 1.7 %; 4 = 12.2 %; 5 = 10.5 %; 6 = 17.1 %; 7 = 11.6 %; 8 = 23.8 %; 9 = 15.5 %; 10 = 7.2 %
BMI (kg/m^2)	25.0 (16.9–38.1)	25.2 (17.4–36.5)
Surgery (%)		
Breast	Yes = 98.4 %	Yes = 100.0 %
Axillary	NA=1.6 % Yes = 89.5 % No = 8.9 % NA=1.6 %	NA=0.0 % Yes = 94.5 % No = 5.5 % NA=0.0 %

* Breast cup size was defined as 1 = AA; 2 = A; 3 = B; 4 = C; 5 = D; 6 = E/DD; 7 = F(E in Italy); 8 = G(F in Italy); 9 = H(FF in Italy); ≥ 10 = J(G in Italy); 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 where EU = European Union; FR = France; IT = Italy; UK = United Kingdom.

facilitated the estimation of the residual uncertainty covering different patient characteristics in the full cohort. Multiple manual landmarks and breast contours were created on each patient's planning computed tomography (CT) scan in this cohort.

2.2. Selection of common coordinate system

To provide a suitable CCS for our patient subgroups, one patient was selected as a "*reference patient*" for 1) supine, 2) left-prone, and 3) right-prone positions. Note the prone patients were divided into left and right due to significant anatomical differences in breast shape and volume. To select the 'most-average' patient, we used variables collected in REQUITE: breast cup representing breast size, breast band representing chest diameter, and body mass index (BMI). Note that breast cup and breast band were mostly reported by the patient. For each subgroup, these patient characteristics were plotted as a 3D point cloud, and the patient at or nearest to the centroid was selected as CCSs.

2.3. Optimisation of image registration pathway

The performance of multiple DIR algorithms was investigated for each registration approach, Fig. 1. We compared the B-spline algorithm from NiftyReg [20] using either NMI (B-spline_NiftyReg-NMI) or sum of square difference (B-spline_NiftyReg-SSD) as a cost function, as well as B-spline and Demon algorithms in simple ITK [21] (B-spline_sITK-NMI and Demon_sITK-HistogramMatch, respectively. These algorithms were chosen to represent various solutions accessible to the research community. To reduce the bias related to DIR initialisation, all deformable registrations were performed after the same affine registration, which was performed using Aladin (Affine_Aladin, part of NiftyReg).

2.3.1. Identifying the optimal image registration pathway

To identify the representative cohorts, patients in the full cohort were clustered according to breast cup size, breast band, and BMI using K-means: 100 clusters for the supine subgroup and 30 for each left-prone and right-prone subgroup. The selected representative patients were those at or nearest to the centroid of each cluster, *see supplementary Figure E.1*.

1) Defining registration pathway for supine and prone set-up: We investigated two pathways of DIR for spatial normalisation for breast IBDM, Fig. 2. *Pathway 1* was a two-step process: In the first step, all patients in a similar set-up position were registered to their corresponding CCS, i.e., all supine patients were registered to the supine CCS, all left-prone patients to the left-prone CCS and all right-prone patients to the right-prone CCS. In the second step, the left- and right-prone CCS were registered to the supine CCS. In contrast, *Pathway 2* consisted in registering all patients in the representative supine, left- and right-prone subgroup directly to the supine CCS.

2) Evaluating the impact of arm position in the supine sub-group: We compared the registration performance between patients treated with bilateral arms raised (70/100) and ipsilateral arms raised (30/100) to the supine CCS patient (treated with bilateral arms raised) in the representative supine subgroup using Mann-Whitney test in SPSS version 29.

2.3.2. Evaluation of image registration performance

1) Evaluation of Image Registration for the Representative Cohort: Image registration performance was evaluated qualitatively using visual

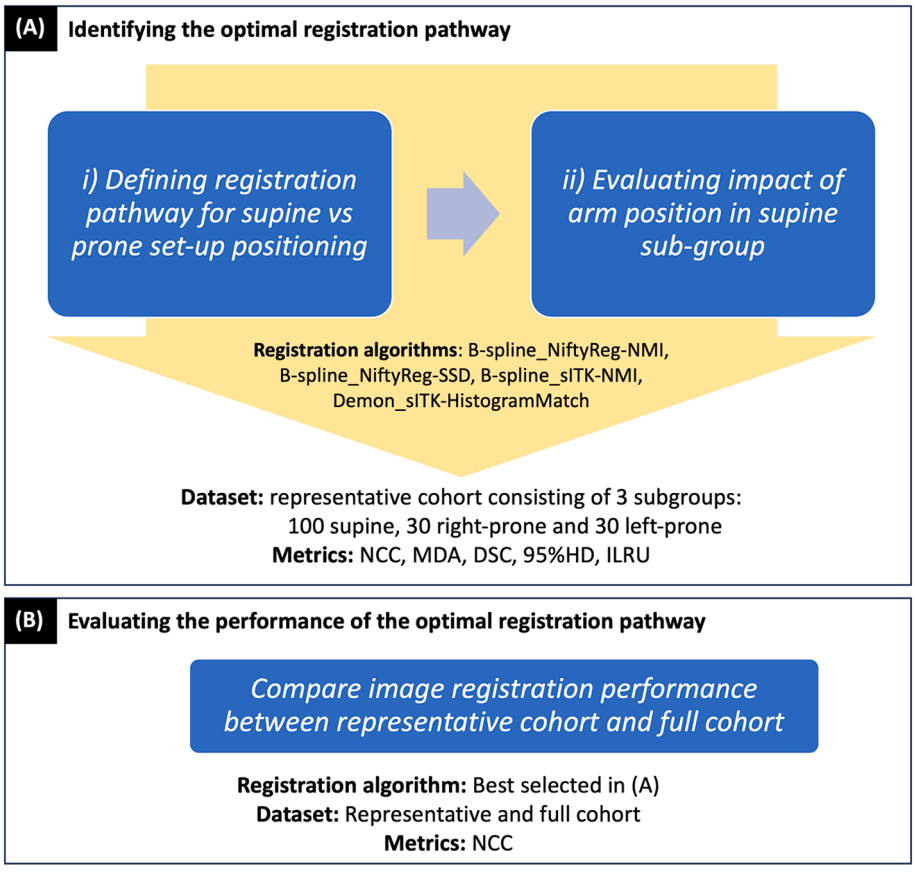

Fig. 1. Overview workflow for selecting the optimal spatial normalisation pathway for breast IBDM by considering (A) the representative cohort using difference image registration algorithm and evaluation metrics and (B) the full cohort. **NMI stands for normalised mutual information, SSD stands for sum of square difference, NCC stands for normalised correlation coefficient, MDA stands for mean distance to agreement, DSC stands for Dice similarity coefficient, HD stands for Hausdorff distance, and ILRU stands for interpatient landmark registration uncertainty.*

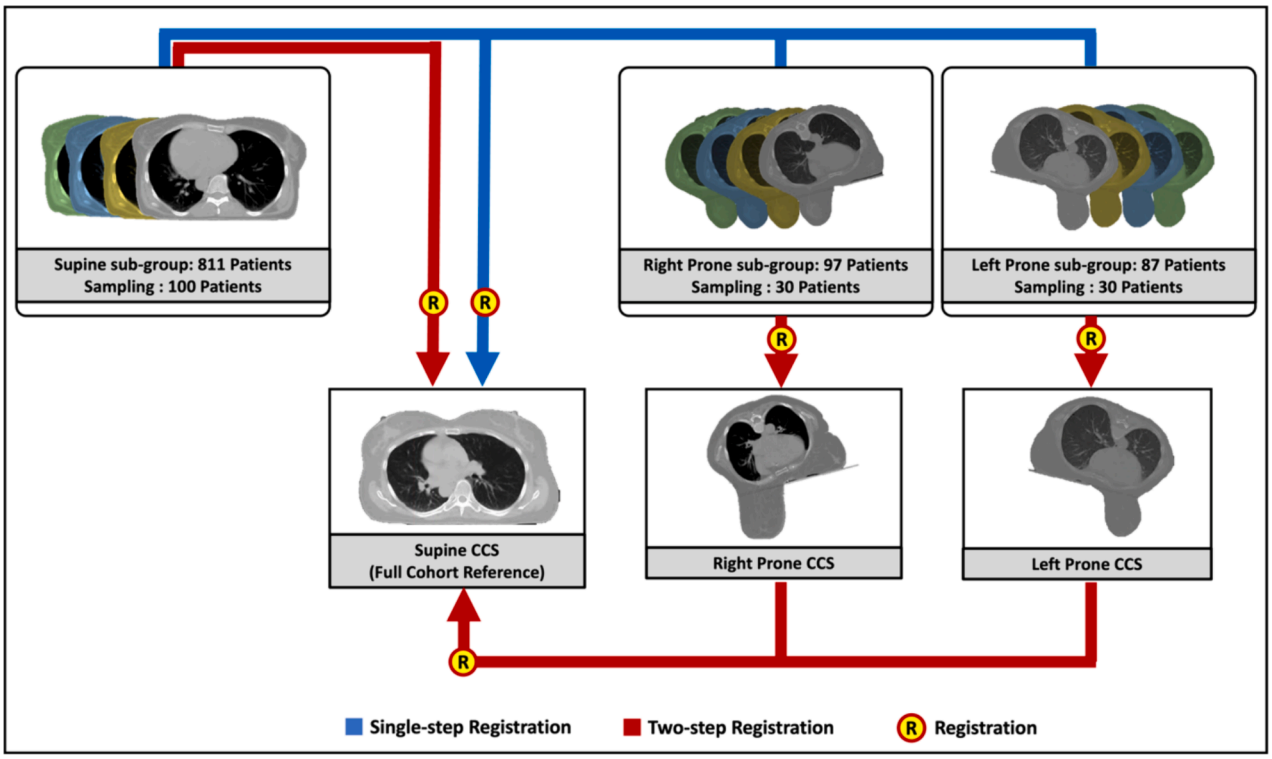

Fig. 2. Schematic of the two image registration pathways in our study. The red pathway represents the two-step registration (*Pathway 1*). First, the registration is performed within subgroups (using the CCS in the same set-up positioning), then the prone CCSs are registered to the supine CCSs and propagated this registration to the subgroup. The blue pathway represents a single-step registration (*Pathway 2*) where all patients were directly registered to the supine CCS, regardless of set-up positioning. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inspection as well as quantitatively using five different metrics: 1) Normalised Cross Correlation (NCC), 2) Dice similarity coefficient (DSC), 3) mean distance to agreement (MDA), 4) 95 % Hausdorff distance (95 %HD) and 5) interpatient landmark registration uncertainty (ILRU).

NCC is an image-based metric used to estimate registration accuracy. We used a region of interest for NCC assessment on the CCSs covering 5 cm beyond the superior and inferior border of the breast (Fig. 3A). For DSC, MDA and 95 % HD, we used ipsilateral and contralateral breast contours generated by atlas-based auto segmentation (Fig. 3B). For ILRU, bony and soft tissue landmarks were identified on the 3D CT images of the representative cohort by an expert clinician to estimate interpatient registration uncertainty, point-based evaluation similar to the transformation registration error (TRE) [22], (Fig. 3C). Definition and further information on these landmarks are detailed in *Supplementary Table E.1*. Note that some landmarks were only identifiable for

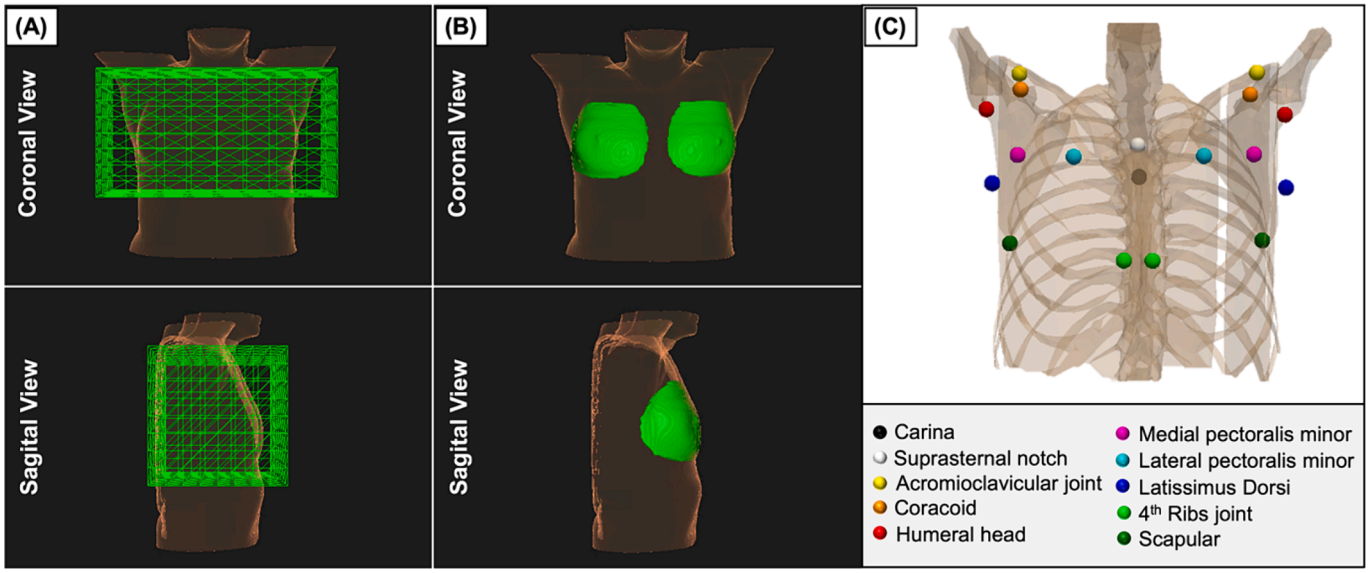

Fig. 3. The regions of interest for evaluating inter-patient image registration performance: (A) box over the chest region for NCC assessment in the representative and full cohort, (B) breast contour for DSC, MDA, and 95 %HD analysis, and (C) anatomical landmarks chosen to be relevant for breast radiotherapy for interpatient landmark registration uncertainty in the representative cohort.

raised arms due to large anatomical differences. The average uncertainty of image registration in each pathway was defined by the average standard deviation from ILRU as it provides the quantitative vector unit.

***2) Evaluation of Image Registration for the Full Cohort*:** we evaluated the registration performance of the full cohort by using the image-based metric NCC. Differences in NCC between the representative cohort and full cohort were analysed using the Mann-Whitney test in SPSS version 29.

***3) Consistency of Image Registration for Additional CCS*:** To investigate the reproducibility of image registration in multiple CCS, we randomly selected two patients around the centroid of the supine subgroup to be the new CCS (2nd and 3rd CCSs). DIR was performed for the supine subgroup to the first full cohort reference and two new CCSs. The performance of DIR among each CCS was reported in terms of NCC to ensure the consistency and repeatability of the spatial normalisation.

3. Results

3.1. Definition of image registration approach

Registration performance for all registration pathways and algorithms is shown in Fig. 4. The highest registration accuracy was obtained with B-spline_NiftyReg-NMI, achieving the lowest MDA (0.23 ± 0.19 cm, 0.38 ± 0.26 cm, and 1.19 ± 1.08 cm), 95 %HD (1.17 ± 1.18 cm, 1.55 ± 0.90 cm, and 4.18 ± 2.45 cm), and ILRU (0.51 cm, 0.53 cm, 0.90 cm) in supine-supine, prone-prone, and supine-prone registration, respectively. Likewise, it achieved a high NCC (0.98 ± 0.01, 0.96 ± 0.03, 0.95 ± 0.06) and high DSC for each sub-cohort (0.91 ± 0.04, 0.86 ± 0.09, and 0.61 ± 0.21). Registration pathway 1 (two-step) performed better than pathway 2. The best registration performance was observed for the supine-supine registrations ($\overline{sd}$ = 0.06, 0.11, 0.36 cm, and 1.49 cm for NCC, DSC, MDA, and 95 %HD), common to both pathway 1 and pathway 2. Performance for the prone registration was better for pathway 1 (each prone patient registered to each prone specific CCS before registering to the supine CCS, $\overline{sd}$=0.07, 0.16, 0.55 cm, and 1.57 cm for NCC, DSC, MDA, and 95 %HD), than for pathway 2 (all prone patients registered to the supine CCS, $\overline{sd}$=0.09, 0.19, 1.19 cm, and 2.45 cm for NCC, DSC, MDA, and 95 %HD). Note that further dividing subgroups and using multiple CCSs did not improve registration performance, *see supplementary Figure E.2 and Table E.2*.

Similar findings were observed when examining ILRU values presented in Table 2. ILRU for all landmarks were lower for supine-supine and prone-prone registrations compared to prone-supine CCS. We observed high variation in the shoulder region around 0.94 ± 0.22 cm and a low variation in the breast and chest region (0.52 ± 0.27 cm) when considering the ILRU of the B-spline_NiftyReg-NMI (Table 2). The average ILRU for Antero-posterior (AP), Superior-inferior (SI), and Right-left (RL) directions in the shoulder region were 0.16 cm, 0.62 cm, and 0.86 cm, and 0.21 cm, 0.43 cm, and 0.49 cm for chest and breast regions. The lowest mean variation was observed for the carina landmark (0.11 cm for similar position registration and 0.51 cm for difference position registration when considering at B-spline_NiftyReg-NMI), while the largest mean variation was found for the humeral head landmark, around 1 cm. For the best and worst registrations using B-spline_NiftyReg-NMI, see supplementary Figure E.3.

In the representative supine subgroup, no difference in DIR accuracy was observed between ipsilateral and bilateral arm(s) raised *(see supplementary Figure E.4)*. The NCC, DSC, MDA, 95 %HD, and ILRU for B-spline_NiftyReg-NMI registration from patients treated with bilateral arms raised were 0.98 ± 0.02, 0.92 ± 0.03, 0.21 ± 0.16 cm, 1.12 ± 1.11 cm, and 0.34 ± 0.15 cm, while these values were 0.98 ± 0.01, 0.90 ± 0.04, 0.27 ± 0.23 cm, 1.28 ± 1.33 cm, and 0.42 ± 0.18 cm for ipsilateral arm raised registration.

3.2. Evaluation of image registration for the full cohort

The median NCC for the full cohorts was 0.98 (IQR 0.01) when applying pathway 1, showing no significant difference from the representative cohort, $p = 0.534$ (*supplementary Figure E.5A*). By considering the performance of image registration in additional CCS, the median NCC for the first, second, and third CCS was 0.98 (IQR 0.01), 0.98 (IQR 0.01), and 0.98 (IQR 0.03), respectively (*supplementary Figure E.5B*).

4. Discussion

The feasibility of a spatial normalisation pathway for breast IBDM was demonstrated in our in-depth analysis. We identified a two-step registration approach using Nifty registration with NMI that is suitable for breast spatial normalisation. This approach optimised registration metrics in both the breast and the shoulder regions. To quantify the spatial normalisation uncertainty for both these regions, we propose and describe a set of landmarks other researchers can adopt. Quantifying the accuracy of spatial normalisation is key, as it affects dose mapping to the CCS and identifying any dose-sensitive regions [23,24].

We identified many challenges in applying spatial normalisation to breast patients. Our study shows the feasibility of using the patient at the centroid or nearby the centroid as a reference for entire sub-cohorts, divided by set-up position (supine, left-prone, right-prone). In previous IBDM work considered in other anatomical sites, the reference patient for IBDM was randomly selected from the entire cohort [16] or was represented by an anatomical phantom [25]. Our registration result from the selected approach for breast spatial normalisation complies with the criteria suggested by AAPM task group 132 when considering DSC (above 0.8) and MDA (0.2–0.3 cm) for the representative cohort and was in line with the recommendations in the "roadmap to clinical translation" proposed by McWilliam et al. [22,24]. Although the right-prone cohort registration showed a lower agreement of registration performance, it is still in the acceptable range of the recommendation.

We found that increasing the number of reference patients did not improve the registration accuracy, *see supplementary Figure E.2 and Table E.2*; therefore, the use of a single reference patient is proposed for each set-up position. This result was also observed in the study by Vasquez Osorio et al. (2018) on the robustness of reference patient selection in lung cancer, which showed that the choice of reference patient did not change the identified region of IBDM [26]. However, a recent IBDM study on head and neck showed that in a small cohort of about 100 patients, different sub-regions could be identified when applying alternative reference patients [17]. Therefore, it will be important to use multiple CCS to assess the consistency of identified sub-regions in breast IBDM, even if the impact on registration accuracy is limited.

This work was performed with a cohort of about 1000 patients from 9 centres in 5 countries. As such, they represent a range of clinical practices, e.g. regarding CT image resolution, positioning approaches (prone vs supine, ranges of arm positions in the supine position), etc. as well as a range of BMI and breast size. However, 94 % of the cohort were European ancestry, and this lack of diversity constitutes a limitation. We investigated four different registration algorithms and showed that the choice of image registration algorithm did impact registration accuracy in the breast and shoulder area, depending on breast shape and volume. This is in agreement with published studies in lung cancer [27]. Therefore, consideration should be given to the selection of the best registration algorithm and optimisation parameters for breast IBDM. Because a gross misregistration will result in patients being excluded from the subsequent voxel-based analysis. Eliminating some patients could introduce biases and limit the applicability of the results to patients outside the cohort studied, as well as decrease the size or toxicity event of the cohort.

Evaluating image registration accuracy in breast cancer also requires assessment of the axilla and shoulder area in order to enable analysis of toxicity endpoints such as lymphoedema and shoulder morbidities.

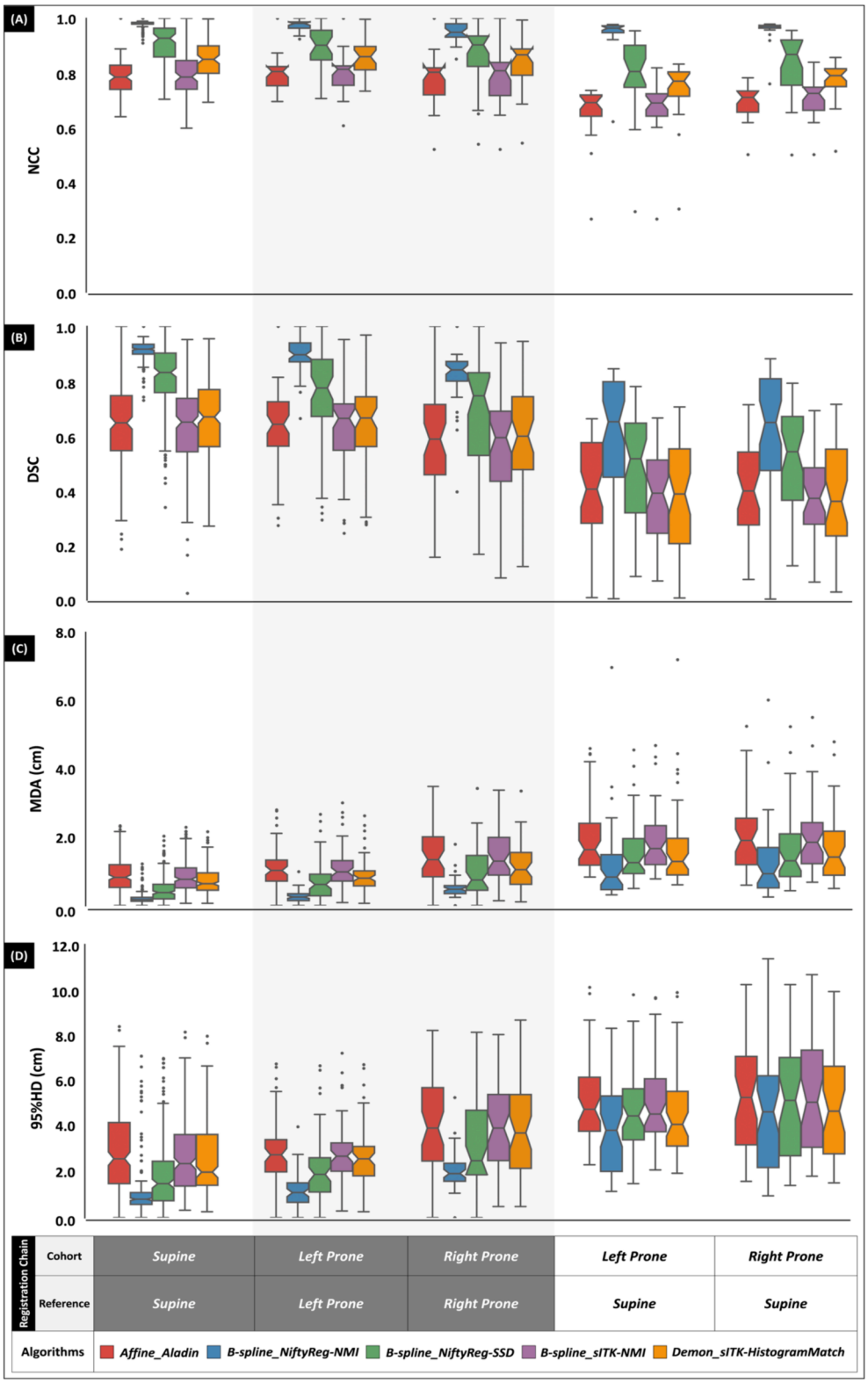

Fig. 4. The image registration accuracy for all registrations considering (A) NCC, (B) DSC, (C) MDA, and (D) 95% HD. For NCC and DSC, the ideal value is one, while for MDA and 95% HD, zero is ideal. Note that supine-supine CCS registrations are common for pathway 1 and pathway 2. Registrations for prone-prone CCS were part of pathway 1 (*highlighted in grey*), while registrations for prone-supine CCS were part of pathway 2. The performance of Affine_Aladin is presented as it is the preliminary step for all DIR pathways.

Table 2
Average standard deviation vector for evaluation of the interpatient landmark registration uncertainty (ILRU). Results are presented as combined for ipsi- and contralateral sides, and for ipsilateral side only. The landmarks representing the breast region include the 4thRibs Joint, Scapular, Suprasternal notch, Carina, Medial and Lateral pectoralis minor, and Latissimus dorsi, while the Acromioclavicular joint, Coracoid process, and Humeral head landmarks represent the shoulder region.

	Landmark	The average vector of standard deviation, $\overline{SD}$ (cm)					
		Approach 1 and 2: Supine – Supine Registration		Approach 1 Prone – Prone Registration		Approach 2 Supine – Prone Registration	
		Affine_Aladin	B-spline_NiftyReg-NMI	Affine_Aladin	B-spline_NiftyReg-NMI	Affine_Aladin	B-spline_NiftyReg-NMI
Bony landmarks	**Acromioclavicular joint**	0.97	0.85	0.89	0.79	1.06	1.08
	Ipsilateral treatment side	0.93	0.63	0.97	0.94	0.99	0.74
	Coracoid process	1.35	0.73	0.74	0.67	0.95	0.95
	Ipsilateral treatment side	1.03	0.54	0.86	0.85	0.86	0.72
	Humeral head	0.94	0.86	1.24	1.15	1.59	1.36
	Ipsilateral treatment side	0.88	0.72	1.79	1.64	1.95	1.23
Soft tissue	**4thRibs Joint**	0.43	0.28	0.60	0.59	1.08	0.99
	Ipsilateral treatment side	0.32	0.23	0.76	0.80	0.94	0.51
	Scapular	0.55	0.56	0.42	0.44	0.86	0.82
	Ipsilateral treatment side	0.45	0.39	0.42	0.47	0.71	0.74
	Suprasternal notch	0.23	0.25	0.31	0.19	0.76	0.64
	Carina	0.24	0.14	0.23	0.08	0.59	0.51
	Lateral pectoralis minor	0.64	0.52	0.75	0.68	1.00	0.96
	Ipsilateral treatment side	0.62	0.48	0.74	0.76	0.76	0.60
	Latissimus dorsi	0.59	0.59	0.51	0.45	0.86	0.82
	Ipsilateral treatment side	0.47	0.43	0.59	0.56	0.62	0.58
	Medial pectoralis minor	0.31	0.30	0.34	0.28	0.98	0.90
	Ipsilateral treatment side	0.28	0.24	0.36	0.26	0.96	0.87

*Highlight is the best registration result among each pathway of the landmark investigated.

Image-based metric such as NCC may not be suitable for analysing the shoulder region due to the differences in the set-up of arm position. The shoulder and axillary anatomy are complexly linked bony, soft tissue and vascular structures which are not routinely delineated in breast radiotherapy. Our study focuses on breast registration. However, registration of OARs would also be of interest, for example, to investigate late cardiac effects. Though this is out of scope for this work, the pipeline could easily be modified to allow this investigation in the future.

In this study, we only evaluated the performance of image registration based on the breast contour and anatomical landmarks within the external thorax region. The inter-observer variation of both contour (e.g., breast and heart) and landmarks can reduce the validity of the geometrical evaluation of DIR [28,29]. Hence, delineating more relevant contours for breast radiotherapy, such as the humeral head, major- and minor-pectoralis muscle, and individual lymph node level, could increase the robustness of registration analysis. However, there is currently a lack of a universally accepted benchmark for evaluating image registration in the breast, shoulder, and axilla region, whether using contour-based or point-based methods. It is the fact that it is difficult to define soft tissue landmarks directly in the axillary region. Therefore, we propose a set of anatomical points, *Supplementary Table E.1*, based on the feasibility of consistent, reproducible identification on CT scan and coverage of the relevant anatomy for the expected radiotherapy dose distribution (on and off target) in patients treated with whole breast and/or regional lymph node irradiation.

Even though our study found spatial normalisation for IBDM in breast RT was feasible, we still faced some limitations when using the best-performing algorithm of our study (B-spline_NiftyReg-NMI) in the very large and pendulous breast where the breast could fall by the side of the body. Thus, investigating skin toxicity for breast radiotherapy may not be possible when applying our registration algorithm. Perhaps an additional landmark for point-point registration [30], surface-based registration, or using artificial intelligence for breast registration may further improve the performance of DIR [31]. However, those methods cannot yet achieve the performance observed in our study [32]. In addition, AI-based DIRs will require a large training model dataset to achieve robust results.

Our study includes only free-breathing patients. However, many breast cancer treatments are now delivered in deep inspiration breath-hold patients. As DIBH images are clearer and contain fewer artifacts, we expect that the performance of our pipeline would only improve, but this hypothesis will require confirmation.

Only one experienced clinician generated landmarks in our study; therefore, it might generate a subjective error, and multiple observers may be required. In future research, using our proposed landmarks in a point-based registration may be beneficial. In addition, dose metric evaluation is not performed in our spatial normalisation study. Still, it would benefit the dose mapping variation assessment, the next step of the voxel-based analysis [24]. Finally, though we have demonstrated the feasibility of DIR performance in breast RT, it remains important to test DIR performance in each cohort where IBDM is applied, especially when gathering large cohort data from multiple centres.

In conclusion, spatial normalisation for breast IBDM is feasible. Our findings will be useful for enabling IBDM investigations and integrating the registration uncertainty in the spatial normalisation processes for breast IBDM.

CRediT authorship contribution statement

Tanwiwat Jaikuna: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Fiona Wilson:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **David Azria:** Resources, Data curation, Writing – review & editing. **Jenny Chang-Claude:** Resources, Data curation, Writing – review & editing. **Maria Carmen De Santis:** Resources, Data curation, Writing – review & editing. **Sara Gutiérrez-Enríquez:** Resources, Data curation, Writing – review & editing. **Marcel van Herk:** Writing – review & editing, Supervision. **Peter Hoskin:** Writing – review & editing, Supervision. **Lea Kotzki:** Resources, Data curation, Writing – review & editing. **Maarten Lambrecht:** Resources, Data curation, Writing – review & editing. **Zoe Lingard:** Resources, Data curation, Writing – review & editing. **Petra Seibold:** Resources, Data curation, Writing – review & editing. **Alejandro Seoane:** Resources, Data curation, Writing – review & editing. **Elena Sperk:** Resources, Data curation, Writing – review & editing. **R Paul Symonds:** Resources, Data curation, Writing – review & editing. **Christopher J. Talbot:** Resources, Data curation, Writing – review & editing. **Tiziana Rancati:** Resources, Data curation, Writing – review & editing. **Tim Rattay:** . **Victoria Reyes:** Resources, Data curation, Writing – review & editing. **Barry S. Rosenstein:** Resources, Data curation, Writing – review & editing. **Dirk de Ruysscher:** Resources, Data curation, Writing – review & editing. **Ana Vega:** Resources, Data curation, Writing – review & editing. **Liv Veldeman:** Resources, Data curation, Writing – review & editing. **Adam Webb:** Resources, Data curation, Writing – review & editing. **Catharine ML West:** Resources, Data curation, Writing – review & editing, Supervision. **Marianne C Aznar:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Eliana Vasquez Osorio:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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.

This work was supported by Cancer Research UK RadNet Manchester [C1994/A28701], the Cancer Research UK Cancer Research Manchester Centre (C147/A25254) and the NIHR Manchester Biomedical Research Centre (NIHR203308).

Marianne Aznar acknowledges the support of the Engineering and Physical Sciences Research Council (Grant number EP/T028017/1) and the NIHR Manchester Biomedical Research Centre (NIHR203308).

Peter Hoskin was supported by the NIHR Manchester Biomedical Research Centre (NIHR203308).

Catharine M L West was supported by the NIHR Manchester Biomedical Research Centre (NIHR203308).

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

Sara Gutiérrez-Enríquez was supported by ERAPerMed JTC2018 funding (ERAPERMED2018-244 and SLT011/18/00005) and currently by the Government of Catalonia (2021SGR01112).

The VHIO acknowledge the Cellex Foundation for providing research facilities, the CERCA Programme/Generalitat de Catalunya for institutional support, and the Agencia Estatal de Investigación for their financial support as a Center of Excellence Severo Ochoa (CEX2020-001024-S/AEI/10.13039/501100011033).

Tim Rattay is supported by the NIHR Leicester Biomedical Research Centre. He was previously an NIHR Clinical Lecturer and was also funded by an NIHR Doctoral Research Fellowship. This publication presents independent research funded by the NIHR. The views expressed are

those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

This work uses data that has been provided by patients and collected by the NHS as part of their care and support. The data are collated, maintained and quality assured by the National Disease Registration Service, which is part of NHS England.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2024.100635.

References

[1] Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, et al. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022;72:409–36.
[2] Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Darby S, McGale P, Correa C, Taylor C, Arriagada R, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Lancet 2011;378:1707–16.
[3] Offersen BV, Alsner J, Nielsen HM, Jakobsen EH, Nielsen MH, Krause M, 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;38:3615–25.
[4] Jaikuna T, Osorio EV, Azria D, Chang-Claude J, De Santis MC, Gutiérrez-Enríquez S, et al. Contouring variation affects estimates of normal tissue complication probability for breast fibrosis after radiotherapy. Breast 2023;72:103578.
[5] Lauridsen MC, Overgaard M, Overgaard J, Hessov IB, Cristiansen P. Shoulder disability and late symptoms following surgery for early breast cancer. Acta Oncol 2008;47:569–75.
[6] Chen MF, Chen WC, Lai CH, Hung CH, Liu KC, Cheng YH. Predictive factors of radiation-induced skin toxicity in breast cancer patients. BMC Cancer 2010;10:508. Published 2010 Sep 23.
[7] Córdoba EE, Lacunza E, Güerci AM. Clinical factors affecting the determination of radiotherapy-induced skin toxicity in breast cancer. Radiat Oncol J 2021;39:315–23.
[8] Ebert MA, Gulliford S, Acosta O, de Crevoisier R, McNutt T, Heemsbergen WD, et al. Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations. Phys Med Biol 2021;66. https://doi.org/10.1088/1361-6560/ac0681. Published 2021 Jun 17.
[9] Green A, Vasquez Osorio E, Aznar MC, McWilliam A, van Herk M. Image based data mining using per-voxel cox regression. Front Oncol 2020;10:1178. Published 2020 Jul 21.
[10] Pasquier D, Le Tinier F, Bennadji R, Jouin A, Horn S, Escande A, et al. Intensity-modulated radiation therapy with simultaneous integrated boost for locally advanced breast cancer: a prospective study on toxicity and quality of life. Sci Rep 2019;9:2759. Published 2019 Feb 26.
[11] De Langhe S, Mulliez T, Veldeman L, Remouchamps V, van Greveling A, Gilsoul M, et al. Factors modifying the risk for developing acute skin toxicity after whole-breast intensity modulated radiotherapy. BMC Cancer 2014;14:711. Published 2014 Sep 25.
[12] Palma G, Monti S, Cella L. Voxel-based analysis in radiation oncology: A methodological cookbook. Phys Med 2020;69:192–204.
[13] Palma G, Monti S, Xu T, Scifoni E, Yang P, Hahn SM, et al. Spatial dose patterns associated with radiation pneumonitis in a randomized trial comparing intensity-modulated photon therapy with passive scattering proton therapy for locally advanced non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2019;104:1124–32.
[14] McWilliam A, Kennedy J, Hodgson C, Vasquez Osorio E, Faivre-Finn C, van Herk M. Radiation dose to heart base linked with poorer survival in lung cancer patients. Eur J Cancer 2017;85:106–13.
[15] Monti S, Xu T, Mohan R, Liao Z, Palma G, Cella L. Radiation-induced esophagitis in non-small-cell lung cancer patients: voxel-based analysis and NTCP modeling. Cancers (Basel) 2022;14. Published 2022 Apr 5.
[16] Beasley W, Thor M, McWilliam A, Green A, Mackay R, Slevin N, et al. Image-based data mining to probe dosimetric correlates of radiation-induced trismus. Int J Radiat Oncol Biol Phys 2018;102:1330–8.
[17] Vasquez Osorio E, Abravan A, Green A, van Herk M, Lee LW, Ganderton D, et al. Dysphagia at 1 year is associated with mean dose to the inferior section of the brain stem. Int J Radiat Oncol Biol Phys 2023;117:903–13.
[18] West C, Azria D, Chang-Claude J, Davidson S, Lambin P, Rosenstein B, 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 (R Coll Radiol) 2014;26:739–42.
[19] Seibold P, Webb A, Aguado-Barrera ME, Azria D, Bourgier C, Brengues M, et al. REQUITE: A prospective multicentre cohort study of patients undergoing radiotherapy for breast, lung or prostate cancer. Radiother Oncol 2019;138:59–67.
[20] http://sourceforge.net/projects/niftyreg/.
[21] Yaniv Z, Lowekamp BC, Johnson HJ, Beare R. SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research [published correction appears in J Digit Imaging. 2019 Sep 4]. J Digit Imaging 2018;31:290–303.
[22] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. Med Phys 2017;44:e43–76.
[23] Murr M, Brock KK, Fusella M, Hardcastle N, Hussein M, Jameson MG, et al. Applicability and usage of dose mapping/accumulation in radiotherapy. Radiother Oncol 2023;182:109527.
[24] McWilliam A, Palma G, Abravan A, Acosta O, Appelt A, Aznar M, et al. Voxel-based analysis: roadmap for clinical translation. Radiother Oncol 2023;188:109868.
[25] Palma G, Monti S, Pacelli R, Liao Z, Deasy JO, Mohan R, et al. Radiation pneumonitis in thoracic cancer patients: multi-center voxel-based analysis. Cancers (Basel) 2021;13:3553. Published 2021 Jul 15.
[26] Vasquez Osorio E. Estimating inter-patient registration uncertainty for image-based data mining. Abstract from Biomedical Image Registration, Leiden, Netherlands. 2018.
[27] Monti S, Pacelli R, Cella L, Palma G. Inter-patient image registration algorithms to disentangle regional dose bioeffects. Sci Rep 2018;8:4915. Published 2018 Mar 20.
[28] Li XA, Tai A, Arthur DW, Buchholz TA, Macdonald S, Marks LB, 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;73:944–51.
[29] Mathew T, Chao M, Lapuz C, Tomaszewski T, Zhang B, Hall M, et al. Consistency of ESTRO and RTOG contouring guidelines for target volume delineation in early stage breast cancer. Int J Radiol Radiat Ther 2020;7:133–40.
[30] Xie X, Song Y, Ye F, Yan H, Wang S, Zhao X, et al. Improving deformable image registration with point metric and masking technique for postoperative breast cancer radiotherapy. Quant Imaging Med Surg 2021;11:1196–208.
[31] Ouyang X, Liang X, Xie Y. Preliminary feasibility study of imaging registration between supine and prone breast CT in breast cancer radiotherapy using residual recursive cascaded networks. IEEE Access 2021;9:3315–25.
[32] Henderson EGA, Van Herk M, Green AF, Vasquez Osorio E. An anatomically-informed correspondence initialisation method to improve learning-based registration for radiotherapy. In: XXth international conference on the use of computers in radiation therapy; 2024. p. 100–3.