Efficacy and Mechanism Evaluation Programme

Evaluation of Image Guided Radiotherapy for more accurate Partial Breast Intensity-Modulated Radiotherapy: comparison with standard imaging technique.

EJ Harris\textsuperscript{1a}, M Mukesh\textsuperscript{2a}, R Jena\textsuperscript{2}, A Baker\textsuperscript{3}, H Bartelink\textsuperscript{4}, C Brooks\textsuperscript{1}, J Dean\textsuperscript{2}, EM Donovan\textsuperscript{1}, S. Collette\textsuperscript{5}, S Eagle\textsuperscript{6}, JD Fenwick\textsuperscript{7}, PH Graham\textsuperscript{8}, JS Haviland\textsuperscript{9}, AM Kirby\textsuperscript{10}, H Mayles\textsuperscript{3}, RA Mitchell\textsuperscript{1}, R Perry\textsuperscript{11}, P Poortmans\textsuperscript{12}, A Poynter\textsuperscript{13}, G Shentall\textsuperscript{14}, J Titley\textsuperscript{9}, A Thompson\textsuperscript{15}, JR Yarnold\textsuperscript{10}, CE Coles\textsuperscript{2b}, PM Evans\textsuperscript{1b*}

On behalf of the IMPORT Trials Management Group

\textsuperscript{1}Joint Department of Physics at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, London, UK
\textsuperscript{2}Oncology Centre, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
\textsuperscript{3}The Clatterbridge Cancer Centre NHS Foundation Trust, Wirral, UK
\textsuperscript{4}Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, Netherlands
\textsuperscript{5}Statistics Department, EORTC Headquarters, Brussels, Belgium.
\textsuperscript{6}Department of Radiotherapy, Royal Marsden Hospital NHS Foundation Trust, London, UK
\textsuperscript{7}Department of Oncology, University of Oxford, Oxford, UK
\textsuperscript{8}Cancer Care Centre, St George Hospital, Kogarah, Sydney, Australia
\textsuperscript{9}ICR-CTSU, Institute of Cancer Research, London, UK
\textsuperscript{10}Breast Unit, Royal Marsden NHS Foundation Trust, London, UK
\textsuperscript{11}Ipswich Hospitals NHS Trust, Ipswich, UK
\textsuperscript{12}Department of Radiation Oncology, Dr Bernard Verbeeten Instituut, Tilburg, Netherlands
\textsuperscript{13}Radiotherapy Department, Peterborough City Hospital, UK
\textsuperscript{14}Rosemere Cancer Centre, Lancashire Teaching Hospitals NHS Trust, Preston, UK
\textsuperscript{15}School of Medicine, University of Dundee, Dundee, UK

\textsuperscript{a}Joint first authors
\textsuperscript{b}Joint principal investigators
* Corresponding author: Professor Philip Evans, Centre for Vision Speech and Signal Processing, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK, GU2 7XH

Competing interests: None declared

All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare (1) no financial support for the submitted work from anyone other than their employer; (2) no financial relationships with commercial entities that might have an interest in the submitted work; (3) no spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; and (4) no non-financial interests that may be relevant to the submitted work.
Abstract

**Background:** Whole breast radiotherapy (RT) is the standard treatment for breast cancer following breast-conserving surgery. Evidence shows that tumour recurrences occur near the original cancer: the tumour bed. New treatment developments include increasing dose to the tumour bed during whole breast RT (synchronous integrated boost) and only irradiating the region around the tumour bed, for patients at high and low risk of tumour recurrence respectively. Currently, standard imaging uses bony anatomy to ensure accurate delivery of whole breast RT. It is debatable whether more targeted treatments such as synchronous integrated boost and partial breast RT require image-guided radiotherapy focussing on implanted tumour bed clips (clip-based IGRT).

**Objectives:** Primary - to compare accuracy of patient set-up using standard imaging versus clip-based IGRT. Secondary objectives – comparison of standard imaging with clip-based IGRT for (i) adequate RT safety margins around the tumour bed to avoid geographical miss (ii) volume of breast tissue irradiated around tumour bed, (iii) estimated breast toxicity following development of a normal tissue control probability model, and (iv) time taken for each imaging method.

**Design:** Multi-centre observational study embedded within a national randomised trial: IMPORT (Intensity Modulated Partial Organ Radiotherapy) HIGH testing synchronous integrated boost and using clip-based IGRT.

**Setting:** Five radiotherapy departments, participating in IMPORT HIGH.

**Participants:** 218 patients receiving breast radiotherapy within IMPORT HIGH

**Interventions:** There was no direct intervention in patients’ treatment. Experimental and control intervention were clip-based IGRT and standard imaging, respectively. IMPORT HIGH patients received clip-based IGRT as routine; standard imaging data was obtained from IGRT images.

**Main outcome measures:** As per objectives.
Results: The primary outcome of overall mean difference in clip-based IGRT and standard imaging using daily set-up errors was 2mm to 2.6 mm (p<0.001). Heterogeneity testing between centres found a statistically significant difference in set-up errors at one centre. For 4 centres (179 patients), clip-based IGRT gave a mean decrease in the systematic set-up error of between 1 and 2 mm compared to standard imaging. Secondary outcomes were as follows: clip-based IGRT and standard imaging safety margins were less than 5mm, and 8 mm, respectively. Using clip-based IGRT, the median volume of tissue receiving 95% of prescribed boost dose was decreased by 29 cm$^3$ (range 11-193 cm$^3$) when compared with standard imaging. Difference in median time required to perform clip-based IGRT compared to standard imaging was X-ray imaging technique dependent (range 8 s to 76 s). It was not possible to estimate differences in breast toxicity as the normal tissue control probability model indicated that for breast fibrosis, maximum radiotherapy dose is more important than volume of tissue irradiated.

Conclusions and implications for clinical practice: Margins less than 8 mm cannot be used safely without clip-based IGRT for patients receiving concomitant tumour bed boost as there is a risk of geographical miss of the tumour bed being treated within the high dose region. In principle, smaller but accurately placed margins may influence local control and toxicity rates, but this needs to be evaluated from mature clinical trial data in the future.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of tables and figures</td>
<td>7</td>
</tr>
<tr>
<td>List of abbreviations</td>
<td>11</td>
</tr>
<tr>
<td>Scientific summary</td>
<td>13</td>
</tr>
<tr>
<td>Plain English summary</td>
<td>20</td>
</tr>
<tr>
<td><strong>Chapter 1 – Introduction</strong></td>
<td>21</td>
</tr>
<tr>
<td>Structure of this report</td>
<td>21</td>
</tr>
<tr>
<td>Background</td>
<td>22</td>
</tr>
<tr>
<td>Risks and benefits in breast radiotherapy</td>
<td>26</td>
</tr>
<tr>
<td>Research hypotheses</td>
<td>28</td>
</tr>
<tr>
<td>Research objectives</td>
<td>28</td>
</tr>
<tr>
<td><strong>Chapter 2 – A critical review of the relationship between irradiated breast volume and late breast tissue complications</strong></td>
<td>30</td>
</tr>
<tr>
<td>Introduction</td>
<td>30</td>
</tr>
<tr>
<td>Methods</td>
<td>31</td>
</tr>
<tr>
<td>Impact of boost volume on breast tissue complications</td>
<td>31</td>
</tr>
<tr>
<td>Partial breast irradiation studies</td>
<td>39</td>
</tr>
<tr>
<td>Breast fractionation studies</td>
<td>47</td>
</tr>
<tr>
<td>Dose modulating effect on the breast</td>
<td>47</td>
</tr>
<tr>
<td>Limitations of this review</td>
<td>48</td>
</tr>
<tr>
<td>Conclusions</td>
<td>49</td>
</tr>
<tr>
<td>Future work</td>
<td>50</td>
</tr>
<tr>
<td><strong>Chapter 3 – Normal tissue complication modelling for breast tissue</strong></td>
<td>55</td>
</tr>
<tr>
<td>Introduction</td>
<td>55</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>56</td>
</tr>
<tr>
<td>Results</td>
<td>62</td>
</tr>
<tr>
<td>Discussion</td>
<td>70</td>
</tr>
<tr>
<td>Conclusion</td>
<td>72</td>
</tr>
</tbody>
</table>
## List of Tables and Figures

<table>
<thead>
<tr>
<th>Table</th>
<th>Caption</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.</td>
<td>Effect of brachytherapy boost volume on NTCP.</td>
<td>38</td>
</tr>
<tr>
<td>Table 2.</td>
<td>Phase II-III randomised controlled trials comparing WBI versus PBI.</td>
<td>46</td>
</tr>
<tr>
<td>Table 3.</td>
<td>Dose-volume characteristics from the Cambridge and the EORTC dataset used for the NTCP model.</td>
<td>63</td>
</tr>
<tr>
<td>Table 4.</td>
<td>Summarised results of the best-fit NTCP parameters for moderate-severe breast fibrosis.</td>
<td>70</td>
</tr>
<tr>
<td>Table 5.</td>
<td>Mean and Variance of $S_{\text{DIFF}}$ for all 112 patients and individual centres. (Highlighted data were used in the sample size calculation).</td>
<td>92</td>
</tr>
<tr>
<td>Table 6.</td>
<td>Summary of imaging technique and patient accrual for each centre.</td>
<td>94</td>
</tr>
<tr>
<td>Table 7.</td>
<td>Systematic and random errors using bony anatomy and clip verification for each centre and for all centres combined.</td>
<td>97</td>
</tr>
<tr>
<td>Table 8.</td>
<td>Delta errors (difference between bony anatomy and clips, $S_{\text{DIFF}}$) in the LR, SI and AP directions and the magnitude of their 3D vector. Time required for image matching with both techniques has also been summarised.</td>
<td>99</td>
</tr>
<tr>
<td>Table 9.</td>
<td>Test of homogeneity of variances of delta errors from all centres, those using techniques 2D-KV technique and after excluding centre using MV-CT.</td>
<td>100</td>
</tr>
<tr>
<td>Table 10.</td>
<td>Mean, Random and Systematic errors using different image verification protocols.</td>
<td>102</td>
</tr>
</tbody>
</table>
Table 11. Patient and treatment characteristics. Characteristics have been categorised according to the information they provide. Median values are given for continuous characteristics.

Table 12. Systematic ($\Sigma_{laser}$) laser set-up errors for patients grouped using patient and treatment related characteristics. For $\Sigma_{laser}$, P-values from the non-parametric Levene’s test are given. Values are only shown for characteristic that gave a significant difference between patient groups ($p<0.05$, without Holm-Bonferroni adjustment). The directions in which differences occur are indicate (SI denotes superior-inferior, LR denotes left-right and AP denotes anterior-posterior).

Table 13. Systematic bony anatomy verification set-up errors ($\Sigma_{BA}$) for patients’ groups determined using patient and treatment related characteristics. P-values for non-parametric Levene’s test are given. Data only given for characteristic that gave a significant difference in systematic bony anatomy verification error between patient groups ($p<0.05$, without Holm-Bonferroni adjustment). The direction in which errors were different is indicated.

Table 14. Radiotherapy treatment planning constraints for IMPORT HIGH. Bold indicates mandatory constraints. Where two dose levels are given they are for the 48Gy or 53Gy test arm doses.

Table 15. Volumes of the breast receiving 95% of prescribed dose from plans based on 5 mm and 8 mm PTV_TB margins. Data are given as median (range) and presented in absolute volume (cc) and as a percentage of the whole breast volume. Differences between 95% volumes for the two PTV_TB were statistically significant ($p<0.01$) and are given in column 4.

Table 16. Dosimetric data given as median (range) for each of the assessment criteria. All differences were found to be statistically significant ($p<0.01$).
<table>
<thead>
<tr>
<th>Figure</th>
<th>Caption</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>IMPORT LOW trial schema.</td>
<td>52</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>IMPORT HIGH trial schema.</td>
<td>54</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>The two step dose volume histogram model.</td>
<td>58</td>
</tr>
<tr>
<td>Figure 4.</td>
<td>Lyman Kutcher Burman Model - The probability of moderate-severe breast fibrosis versus biological equivalent dose using $\alpha/\beta$ of 3 Gy ($BED_3$). The solid line is based on the best-fit parameters ($BED_3 = 132$ Gy and $m= 0.35$) and the dashed lines are upper and lower 95%CI. The summative toxicity data of the three dose fractionations ± boost at five years from the START pilot trial are plotted.</td>
<td>65</td>
</tr>
<tr>
<td>Figure 5.</td>
<td>Niemierko Model - The probability of moderate-severe breast fibrosis versus biological equivalent dose using $\alpha/\beta$ of 3 Gy ($BED_3$). The solid line is based on the best-fit parameters ($BED_3 = 136.4$ Gy and $\gamma_{50}= 0.9$) and the dashed lines are upper and lower 95%CI. The summative toxicity data of the three dose fractionations ± boost at five years from the START pilot trial are plotted.</td>
<td>66</td>
</tr>
<tr>
<td>Figure 6.</td>
<td>Set-up errors and PTV margin in radiotherapy.</td>
<td>75</td>
</tr>
<tr>
<td>Figure 7.</td>
<td>Standard verification technique compares digitally reconstructed radiograph and mega-voltage portal image.</td>
<td>76</td>
</tr>
<tr>
<td>Figure 8.</td>
<td>Measurement of central lung distance (CLD) in black and cranio-caudal distance (CCD) in blue on a DRR.</td>
<td>77</td>
</tr>
<tr>
<td>Figure 9.</td>
<td>Bony anatomy and clip based verification using kV-CBCT, MVCT and 2D-kV.</td>
<td>86</td>
</tr>
</tbody>
</table>
Figure 10. Bland-Altman plots of average of bony anatomy and clips mean set-up error versus difference between mean clip set-up error mean bony anatomy set-up error in the LR, SI and AP directions. Solid line indicates mean difference between mean clip and mean bony anatomy set-up errors (the bias) and the dotted lines represent the limits of agreement (±1.96SD).

Figure 11. Tumour bed PTV margins required for the different imaging verification protocols considered in this study. Margins are given for the LR, SI and AP directions.

Figure 12. Schematic diagram to show (a) TB location viewed on axial CT slice (1(blue) = medial, 2(pink) = chest wall, 3(green) = anterior and 4(yellow) = lateral) and (b) TB location in the SI (superior-inferior) direction viewed on sagittal CT slice (1 = superior, 2 = middle and 3 = inferior).

Figure 13. Section through a patient’s treatment plan showing CTV_TB (red), PTV_TB = 5 mm (yellow) and PTV_TB = 8 mm (pink).

Figure 14. A) shows an original PTV_TB with 5 mm margin and in close proximity to the lung. When PTV_TB is increased by a further 3 mm it expands into the lung B) and to achieve coverage requires an increase in the width of the tangential fields by 4 mm C) which in turn increasing the dose to heart.
List of abbreviations

BASO  British Association of Surgical Oncology
BCS  breast conserving surgery
BEUD  biologically equivalent uniform dose
CBCT  cone-beam computed tomograph
CRT  conformal radiotherapy
CT  computed tomograph
CTV  clinical target volume
CTV_TB  tumour bed clinical target volume
DVH  dose-volume histogram
EBCTCG  Early Breast Cancer Trialists Collaborative Group
EORTC  European Organization for Research and Treatment of Cancer
EQD2  equivalent dose in 2 Gy fractions
Gy  Gray, unit of radiation dose equal to 1 Joule per kilogram of tissue
HDR  high dose rate
IGRT  image guided radiotherapy using clip-based verification
IMPORT  intensity modulated and partial organ radiotherapy trial
IMPORT HIGH  intensity modulated and partial organ radiotherapy trial – higher risk patient group
IMPORT-IGRT  intensity modulated and partial organ radiotherapy trial – image guided radiotherapy study
IMRT  intensity modulated radiotherapy
IORT  inter-operative radiotherapy
IVP  image verification protocol
MC  Monte Carlo
MV  megavoltage
NTCP  normal tissue complication probability
NTD_{50}  tolerance dose which results in 50% chance of tissue injury
PBI  partial breast irradiation
PTV  planning target volume
PTV_TB  tumour bed planning target volume
QOL  quality of life
QUANTEC  Quantitative Analyses of Normal Tissue Effects in the Clinic
RT radiotherapy
START Standardisation of breast radiotherapy
WBI Whole breast irradiation
Scientific Summary

Background
The role of breast radiotherapy after breast conserving surgery is well established with the 2005 systematic overview of the Early Breast Cancer Trialists Collaborative Group demonstrating a 70% proportional reduction in local tumour recurrence risk following radiotherapy for patients treated with breast conserving surgery for early stage breast cancer.

A wealth of evidence confirms that most recurrences occur close to the primary tumour: in the region referred to as the tumour bed. For this reason a higher radiotherapy dose may be given to the tumour bed compared to the rest of the breast. This extra tumour bed “boost” typically reduces local relapse risk by 50%, at the expense of a 30% increase in the risk of moderate/severe breast fibrosis and is usually given after whole breast radiotherapy. New treatment developments include increasing dose to the tumour bed during whole breast radiotherapy (synchronous integrated boost) and simply irradiating the region around the tumour bed (partial breast radiotherapy), for patients at high and low risk of tumour recurrence respectively.

Currently, standard imaging uses bony anatomy to ensure accurate delivery of whole breast radiotherapy. In addition, a relatively wide safety margin of normal tissue is added to the breast to account for uncertainties in its position on each day of treatment. New imaging techniques use titanium clips implanted in the tumour bed during surgery, which are imaged with x-rays during treatment. This is called clip-based image guided radiotherapy (IGRT) and has been used in conjunction with synchronous integrated boost and partial breast radiotherapy as it is perceived to locate the tumour bed more accurately than standard imaging. This perception has led to the use of smaller safety margins around the tumour bed under the premise that the smaller volume irradiated will reduce late normal tissue toxicity (mainly fibrosis) and facilitate dose escalation, which may reduce tumour recurrence. Despite this shift in breast radiotherapy practice, two questions remain largely unanswered. Firstly, what is the accuracy of clip-based IGRT compared to standard imaging? Secondly, if clip-based IGRT irradiates a smaller volume of normal breast tissue around the tumour bed, can we predict how this would reduce side effects?
The UK IMPORT HIGH trial provided a unique opportunity to answer the above questions and is led by members of the group involved in this study. It is a randomised trial of radiotherapy dose escalation using a synchronous integrated boost, in women at higher than average risk of local cancer recurrence after breast conserving surgery. The programme of work presented in this report is a sub-study of the IMPORT HIGH trial. There was no intervention in patients’ treatment: IMPORT HIGH patients received clip-based IGRT as routine; standard imaging data was obtained from IGRT images. This novel sub-study design allows direct comparison of clip-based IGRT with standard imaging, but does not pose the ethical dilemma of randomising patients to potentially less accurate imaging for synchronous integrated boost radiotherapy.

**Objectives**

The primary objective was:

- To compare the spatial accuracy of breast radiotherapy based on imaging i) titanium surgical clips implanted in the tumour bed (clip-based IGRT) and ii) bony anatomy and lung position during curative radiotherapy for early breast cancer (standard imaging).

Secondary objectives were:

**Comparison of standard imaging with clip-based IGRT for:**

(i) Adequate RT safety margins around the tumour bed to avoid geographical miss

(ii) Volume of breast tissue irradiated around tumour bed

(iii) Estimated breast toxicity following development of a normal tissue control probability model

(iv) Time taken for each imaging method

**Methods**

This project was a staged programme of work with five main studies. These may be split into two sets. The first set involved study of the evidence for a dose volume effect in breast radiotherapy. The second set involved an analysis of the effects of clip-based IGRT on treatment margins.
The set of studies to evaluate evidence for a dose volume effect in breast radiotherapy had two component studies. The first was a review of the published literature and the second was a quantitative analysis of dose volume effect for breast tissue.

The literature review evaluated evidence from a range of radiotherapy studies. These included randomised trials evaluating a boost to the tumour bed versus no boost, with the boost delivered via a range of modalities and approaches, including brachytherapy, cobalt-60, inter-operative irradiation, electrons and photons. A second area of analysis of the literature was the evidence from studies of partial breast irradiation, which is a mode of treatment with current clinical and research activity. A third area of analysis was evidence from breast fractionation studies.

In the second study, data from two large randomised trials were analysed: the Cambridge IMRT (intensity modulated radiotherapy) trial and the EORTC 22881-10882 “boost versus no boost” trial. The Cambridge trial was a single centre study, which recruited 1145 patients with stage T1-T3N0-1M0 invasive breast cancer or ductal carcinoma in situ. Patients received whole breast radiotherapy, followed by an electron boost to the tumour bed in selected cases (n=728). Breast fibrosis was assessed at 2 and 5 years after completion of radiotherapy. The EORTC study was a multi-centre trial that recruited 5569 patients with stage T1-T2N0-1M0 invasive breast cancer. Patients received whole breast radiotherapy and were randomised to three boost levels: no boost (n=2657), ii) 10 Gy boost (n=126), iii) 16 Gy boost (n=2661), and iv) 26 Gy boost (n=125). Breast fibrosis was assessed clinically at follow up. The relationship between partial breast volume irradiated to high dose and probability of moderate or severe fibrosis was fitted using two standard NTCP (normal tissue complication probability) models: the Lyman Kutcher Burman (LKB) and Niemierko models. These models use three parameters to describe the dose response: the uniform dose to the whole breast to produce 50% complication probability, the steepness of the dose response and the volume effect.

The second set of studies examined the effects of clip-based IGRT. It was carried out as a sub-study of the IMPORT HIGH national trial. The clip-based IGRT approach used in IMPORT trial was the use of titanium surgical clips implanted at the time of breast conserving surgery and imaged using x-rays. The first study compared the clip-based IGRT method with two other approaches: the use of x-ray imaging of bony anatomy (standard
imaging) and the use of laser-based set-up using skin markers (no imaging). In the first analysis the set-up accuracy of these methods was analysed and the resulting safety margins for set-up error needed were determined. The time required to perform image matching of clips and bony anatomy was also measured and recorded. A second study evaluated the patient and treatment characteristics that influenced the resulting set-up errors. The third study evaluated the effects of the margins required for the three set-up methods on the radiotherapy planning of the patient’s treatment.

218 patients recruited by five centres to the IMPORT HIGH trial contributed to this study. The centres used a range of imaging methods to visualise the titanium clips and bony anatomy. Centre A used kV cone-beam CT (n=79), B used megavoltage-energy CT (n=40) and C, D and E used 2D kV planar imaging (n=39, 30 and 30, respectively).

Patient random and systematic set-up errors were measured for bony anatomy and clip-based IGRT. The differences between the two measurement sets were used to generate delta errors which described the extra uncertainty produced by the use of bony anatomy matching in the absence of clip-based IGRT. Differences in set-up errors, delta errors and times between centres, imaging modalities and imaging protocols were investigated. Population random and systematic set-up errors was determined and used to generate the necessary margins for error to achieve target coverage, using standard margin formulae and for a variety of image verification protocols.

Patient and treatment characteristics that influence set-up accuracy were studied using patient characteristics of position of the tumour bed and breast volume. Surgery characteristics included seroma visibility, surgery closing technique, number of clips and clip position. Radiotherapy characteristics included IMPORT HIGH trial arm, time between surgery and chemotherapy and time between chemotherapy and radiotherapy.

The effects of the different safety margins using clip-based IGRT and standard imaging were studied by re-planning 60 patients from the IMPORT HIGH trial. Treatment plans were generated for two planning target volume (PTV) margins: 5 mm (achievable with clip-based IGRT) and 8 mm (required for bony anatomy based verification). Two types of plan were generated: 30 patients were planned using a sequential, conformal photon boost to the tumour
bed and 30 using the simultaneous integrated boost technique. The plans were generated to fit the dose constraints required by the IMPORT HIGH trial.

**Results**

In the literature review, one of the strongest pieces of evidence for a dose-volume effect was from a study by Borger et al using low dose iridium implants. This study found evidence that for every 100 cm$^3$ increase in the volume of the boost region, the risk of fibrosis increased by a factor of four and that a two-fold increase in boost volume results in an 11% reduction in the normal tissue tolerance dose. Other studies supporting volume effect for breast tissue included trials comparing brachytherapy based partial breast irradiation (PBI) and intra-operative radiotherapy (IORT) with whole breast irradiation. The brachytherapy and intra-operative dose distribution can differ from the external beam radiotherapy and therefore, it is unclear whether these results can be extrapolated to external beam techniques. There is some evidence to support volume effect using external beam techniques. The Royal Marsden Gloucester trial used an electron boost and showed that for every Gy increase in boost dose, the risk of moderate to severe breast induration increases by 1%. In comparison, a 1Gy increase to the whole breast can increase the risk of moderate-severe breast induration by 3%, indicating a dose-volume effect. Two large studies, IMPORT Low and Danish Breast Cancer Cooperative Group trial used external beam radiotherapy for partial breast irradiation and will provide more robust data on dose-volume effect in the near future.

Individual patient data of 5856 patients from the Cambridge trial and EORTC trial was used to develop the normal tissue complication probability (NTCP) model of breast fibrosis. The best fit for the Niemierko model gave a value for the biologically equivalent uniform dose (BEUD) dose to the whole breast which produces a 50% complication rate of 136.4 Gy. The parameter describing the steepness of the dose response was $\gamma_{50} = 0.9$ and the parameter for the volume response was $n = 0.011$. The best fit for the LKB model was (BEUD$_{50}$=132 Gy, $m = 0.35$ and $n = 0.012$). The $n$ parameter describing the volume effect ranges between 0 (for no volume effect) and 1 (for a strong volume effect). Hence, these results, from both models strongly imply that the risk of moderate or severe breast fibrosis is mainly associated with radiotherapy dose and that the change in volume of tissue irradiated does not change the risk of breast fibrosis. These results were validated on an independent dataset from the START trial. One of the secondary objectives of the programme was to estimate the reduced risk of late adverse effects resulting from the smaller tissue volume irradiated. However, based on
the results of “no volume effect”, it was not possible to predict a reduction in risk of breast fibrosis if a smaller tissue volume is irradiated. Clearly, any model has limitations and the mature results from the clinical trials addressing this question are awaited.

The primary research objective of this study was to compare the accuracy of clip-based IGRT and standard imaging using bony anatomy. The random and systematic set-up errors for bony anatomy and clip-based IGRT were found to be 3 mm averaged over the five centres, with no strong evidence for differences between the centres. The delta errors (difference between clips and bony anatomy) were found to be between 2 and 3 mm. The margin formulae showed that the use of no imaging (i.e. laser-based set-up) requires a PTV margin of 8-10 mm, the use of standard imaging allows this to be reduced to 7-9 mm and the use of clip-based IGRT with a suitable verification protocol allows the margin to be reduced to 4-5 mm. The time taken to perform clip match was quicker than bony anatomy match using 2D-KV technique, but not when using cone beam CT imaging (secondary objective).

For the study of patient, surgery and radiotherapy characteristics that influence set-up errors, laser-based set-up (no imaging) was found to be significantly influenced by breast volume, seroma visibility and surgical closing technique. Bony anatomy (standard imaging) based set-up was found to influenced by both breast volume and tumour bed axial position.

The results of the re-planning study showed that the reduced margins that were achievable with clip-based IGRT compared to standard imaging (5mm versus 8mm) led to a reduction of 29 cm³ (range 11 – 193 cm³) in the volume of breast tissue receiving a high dose. Using clip-based IGRT margin (5mm), 56 of the 60 cases met all the IMPORT High treatment planning criteria. Using standard imaging margin (8mm), four sequential boost plans and 10 concomitant boost plans breached mandatory planning constraints. The use of smaller PTV margins with clip-based IGRT also allowed a small reduction in the radiotherapy dose to the contralateral breast, heart and lung.

Conclusions and implications for clinical practice

This research demonstrates the benefits of clip-based IGRT over standard imaging, with a reduction in PTV margins. Margins less than 8 mm cannot be safely used without clip-based IGRT for patients receiving concomitant tumour bed boost as there is a risk of geographical miss of the tumour bed being treated within the high dose region.
The existing literature suggests a volume effect for breast tissue, but our normal tissue complication probability (NTCP) model could not demonstrate a volume effect for breast fibrosis. We anticipate mature results from the ongoing clinical trials to provide a definitive answer. In principle, these smaller, but accurately placed margins may also influence local control rates, but again this needs to be evaluated from mature clinical trial data in the future.

[2344 words]
**Plain English Summary**

Whole breast radiotherapy is the standard treatment for breast cancer following breast-conserving surgery (lumpectomy). Cancer recurrences are most likely to occur near the original cancer: the tumour bed. A new technique aims to reduce recurrence by delivering a higher dose to the tumour bed (“boost”) during whole breast radiotherapy. Currently, X-rays of the ribcage (standard imaging) are used to ensure accurate delivery of breast radiotherapy. Newer imaging using surgical clips within the tumour bed (clip-based imaging) may be preferable for boost radiotherapy.

The main objective was to compare accuracy of radiotherapy boost with standard and clip-based imaging. The bigger “safety margin” required around the tumour bed was calculated and a mathematical model was constructed to estimate whether the extra volume irradiated caused more side effects. 218 patients receiving breast radiotherapy within a national breast boost trial were studied: all had clip-based imaging, but standard images of the ribcage were available for comparison.

Results show that clip-based imaging is more accurate than standard imaging for boost radiotherapy and safety margins are 5mm and 8 mm, respectively. The volume of breast tissue irradiated decreased by 29 cm$^3$ (range 11-193 cm$^3$) using clip-based imaging, but estimation of side effects was not possible using the model.

In conclusion, margins less than 8 mm cannot be used safely without clip-based imaging for patients receiving boost radiotherapy as the higher dose boost treatment may “miss” the tumour bed. Smaller margins may reduce both cancer recurrence and side effects, but long-term results from on-going trials are needed.

[250 words]
Chapter 1 - Introduction

Structure of this report

The work in this report is based on a sub-study of a national randomised controlled trial (RCT) called IMPORT HIGH (intensity modulated and partial organ radiotherapy trial – higher risk patient group) (CRUK/06/003). This trial was funded by Cancer Research UK under grant number (C1491/A16831) The sub-study is known as IMPORT-IGRT (intensity modulated and partial organ radiotherapy trial – image guided radiotherapy study).

The work described is a staged EME (Efficacy and Mechanism Evaluation programme) study and hence the structure of the report reflects the stages of the programme.

The report starts with a review of the background to the problem of target localisation in breast radiotherapy and the expected advantages of image guided radiotherapy to solve various aspects of the problem. The research objectives are listed at the end of this first chapter.

Subsequent chapters then address each of the components of the study. These generally fall into two groups: i) radiobiological side-effects of radiotherapy on normal breast tissue and ii) a study of the role of image guidance to improve treatment accuracy, reduce margins for error, and consequently reduce effects on normal tissue. Two chapters discuss radiobiology. Chapter 2 is a critical review of the literature on the relationship between irradiated volume of breast tissue and late breast tissue complications and chapter 3 is the analysis of pooled results from two randomised trials. Three chapters discuss aspects of the role of clip-based IGRT. The first of these, chapter 4, presents a comparison of standard imaging and clip-based IGRT for 5 centres participating in the IMPORT HIGH trial. Chapter 5 presents an analysis of factors that influence the relationship between set-up errors and margins generated using visual set-up, standard imaging and IGRT. In this analysis, clip-based IGRT is taken as the gold standard. Chapter 6 presents an analysis of the volume of tissue spared high dose irradiation using IGRT compared with standard imaging approaches.

The final chapter summarises the findings of this project and presents the key conclusions from this research programme.
Background

Radical RT following breast conserving surgery (BCS) is a proven alternative to mastectomy for a majority of women developing breast cancer. The success of the treatment was highlighted in the 2005 systematic overview of the Early Breast Cancer Trialists Collaborative Group (EBCTCG)\(^1\) which showed a 70% proportional reduction in local tumour recurrence risk following radiotherapy for patients treated with BCS for early stage breast cancer.

For every 100 women treated, radiotherapy prevents 20 local cancer relapses in the breast and 5 cancer-related deaths 10 years post-treatment\(^1\). Approximately 30,000 women are given radiotherapy for early breast cancer in the UK every year, resulting in the prevention of 6,000 local tumour relapses and 1,500 deaths from cancer. An estimated 1,500 local tumour relapses occur despite radiotherapy, with a disproportionate number affecting women below fifty years old, who have three times the risk of local relapse risk than women in older age groups\(^2\). A wealth of evidence confirms that most relapses occur close to the primary tumour: in the region referred to as the tumour bed\(^3\). This is the reason for giving a higher dose of radiotherapy to the tumour bed than to the rest of the breast. This extra dose to the tumour bed is called the boost and typically reduces local relapse risk by 50% at the expense of a 30% increase in moderate or severe hardening of breast tissue due to fibrosis\(^2,4\). Currently used protocols require a wide margin of healthy tissue to be added around the tumour bed to compensate for substantial (5-10 mm) variations in the patient position from one day of treatment to the next, which consequently limits the radiation dose that can be safely delivered. The challenge is to safely reduce the volume of healthy tissue included in the boost treatment in order to reduce the risk of late complications with the consequence of allowing dose escalation to be done safely and potentially higher cure rates. The hypothesis tested in this programme was that modern radiotherapy technology (including treatment machines equipped with on-line x-ray imaging facilities that can monitor the position of internal organs relative to the treatment beam, accurately) allows a substantial reduction in the safety margin around the tumour bed, thus reducing both exposure of healthy tissue and chronic complication with the inherent possibility of safe dose escalation.

The UK IMPORT HIGH trial provided a unique opportunity to test this novel approach. IMPORT HIGH is led by members of the group involved in this study. It is a randomised trial
of radiotherapy dose escalation in women at higher than average risk of local cancer recurrence after surgery. The two main challenges for this current project are i) direct measurement of the magnitude of tumour bed margin reduction and therefore tumour bed boost volume reduction achieved by tumour bed imaging and ii) estimation of the reduction in rates of moderate and severe fibrosis (breast hardening). The success of this project was judged based on these two challenges.

Radiotherapy delivery for breast cancer is a two-stage process. Firstly the patient has a CT (computed tomography) scan of her breast and chest region. During this scan the patient is lying in a comfortable position that is reproducible during subsequent visits for treatment. The CT dataset obtained is a basis for planning the treatment. This involves finding the optimum method of targeting the therapeutic radiation beam to deliver the desired dose distribution. Once the treatment plan is available, the patient commences a treatment programme consisting of a sequence of daily treatment visits, called fractions, typically over a period of 3 weeks. The positioning of the patient during these fractions is the same as that for the planning CT scan. This is achieved using the positioning of tattoos on the patient’s skin surface relative to the treatment beam (with the help of a set of laser lights that are fixed relative to the position of the treatment beam).

The treatment beam is a high-energy x-ray beam. This is chosen as it penetrates well through tissue and interacts fairly uniformly with various tissue types (e.g. fat and muscle). A consequence of this is that the radiation dose distribution is well behaved in the different tissue types but the high energy x-ray beam is not optimal for imaging, which ideally requires substantial differences in how the beam interacts with different tissues. During most standard approaches to treatment, images are acquired using the high energy x-ray beam, but these show only tissues whose density is greatly different to soft tissues, e.g. the ribs and lungs, not the soft tissues of the breast in the region where the tumour bed is. The images are compared with the planning CT images, collected before treatment, to identify discrepancies in the positioning of ribs and lung within the beam, referred to as set-up errors. If set-up errors are found to exceed predefined limits, then the patient is re-positioned, before the next treatment fraction is delivered. This repositioning is of the order of several mm. It is small on the scale of the human body but significant for radiotherapy accuracy. A major problem with this approach is that the positions of the lungs and ribs do not necessarily predict with
sufficient accuracy the location of the tumour bed. As a consequence wide margins of healthy tissue need to be added to the tumour bed boost volume.

Standard practice uses the high energy x-ray beam produced by the treatment machine to identify the position of the ribs and lungs within the beam. The high energy x-rays are too penetrating to show the soft tissues of the breast, also the tumour bed can move more than a centimetre relative to the bony anatomy and lungs. The inability to directly visualise the tumour bed means that its position cannot be measured directly, requiring a safety margin to avoid geometric miss and hence the exposure of a volume of normal breast tissue to higher doses than needed. The volume of tissue including this safety margin is known as the tumour bed boost volume. The need for this safety margin can substantially increase the volume of tissue irradiated.

This larger volume of healthy tissue irradiated around the tumour bed also places limits on the total dose that can be safely delivered. The hypothesis under test in this study is that if the tumour bed is imaged directly during treatment, the volume of healthy tissue in the tumour bed boost volume can be reduced, leading to a smaller volume of healthy tissue irradiated and hence to fewer treatment complications. An alternatively strategy, in women at highest risk of local recurrence, is to allow safe dose escalation to the tumour bed, with expected better local cancer cure rates. Before justifying these expectations, a simple surgical technique will be described that has enabled the tumour bed to be imaged directly on each day of a patient’s radiotherapy.

As part of the enabling infrastructure for the IMPORT HIGH study we have asked UK breast surgeons to attach standard titanium surgical clips to the walls of the tumour excision cavity. This accurately demarks the position of the tumour bed. Pilot work was undertaken in preparation for the IMPORT HIGH trial and as a consequence clips are now recommended for all patients undergoing breast conservation surgery by the British Association of Surgical Oncology. The key advantage of this approach is that surgical clips can be visualised directly using the low energy x-ray facility on the treatment machine. Furthermore it has been confirmed that boost treatments verified using the position of clips are more accurate than those only relying on imaging of the positioning of the ribs and lungs.
Discrepancies between the expected, planned positions of the surgical clips (based on the planning x-ray CT scan performed before the treatment) and the actual positions of the clips (based on imaging on the treatment machine) are corrected using small movements of the patient, of typically a few millimetres. The use of x-ray imaging of surgical clips to verify radiotherapy accuracy in real-time, i.e. immediately before each treatment fraction is given, is referred to as clip-based image guided radiotherapy (clip-based IGRT). The pilot study for the IMPORT HIGH trial suggests that clip-based IGRT is likely to allow smaller safety margins of healthy breast tissue around the tumour bed.

If the standard imaging method is used to verify the accuracy of the patient’s treatment no information about the tumour bed is available. In a comparison of clip-based IGRT and bony anatomy set-up (i.e. the standard imaging approach), an additional safety margin of 4.5 - 5.5 mm was needed for standard imaging. Another study found the set-up error to be on average 4mm (with 3 mm variation at 1 standard deviation) greater when using standard set-up compared to clip-based IGRT. In a recent review of the growing literature reporting significant changes in the size and position of the tumour bed, Kim et al. highlighted the fact that clinical factors such as the time between surgery and chemotherapy and the planning CT scan influence the size of the change in the tumour bed that occurs during treatment. If large changes occur then standard imaging is unable to detect this and consequently errors in patient positioning will increase in the presence of such changes. For any subset of patients, for whom there are clinical factors that result in large changes in the size and position of the tumour bed we would expect that the inaccuracy of standard imaging compared to IGRT would be greater and subsequently greater margins would need to be applied to ensure good target coverage. An example of this is the sub-set of patients who receive chemotherapy and consequently have a longer time interval between surgery and the start of radiotherapy. A 5 mm safety margin is added around the tumour bed in the IMPORT HIGH trial, compared to a 10 mm margin when standard set-up is used. In summary, reducing margins around the tumour bed translate into smaller volumes exposed to high doses and fewer expected late side effects, such as hardness and tenderness of the breast. Clinical evidence justifying this expectation is now discussed.
Risks and benefits in breast radiotherapy

The dose responses for tumour control and normal tissue damage in radiotherapy are long established and well understood. In the UK Standardisation of Breast Radiotherapy (START) trials (conducted by members of the IMPORT HIGH trial group and the group undertaking this study) physical morbidity was defined in terms of breast shrinkage, distortion and fibrosis. This was scored by i) independent expert observers using serial clinical photographs, ii) examination by physicians in the clinic and iii) by patient self-assessment at regular time-points over 5 years of follow up\(^4,11\). We reported one third of women with minor or marked change in photographic breast appearance (in terms of shrinkage and distortion). These chronic effects increase in incidence and severity even after 5 years post radiotherapy. The photographic changes were in accordance with prospective patient self-assessments of adverse effects, including moderate or marked change in skin appearance (reported by around 30% of patients), breast hardness (over 40%), and breast shrinkage (over 20%). A study of 254 patients undergoing breast conserving surgery and radiotherapy reported that physical changes in breast tissue had a marked bearing on subsequent psychological outcome\(^12\). Patients completed questionnaires assessing satisfaction with treatment outcome and scored psychosocial morbidity using the Hospital Anxiety and Depression (HAD) scale, the Body Image questionnaire and the Rosenberg Self-esteem scale. There was a strong association between the breast appearance and levels of anxiety (r = -0.81, P <0.001), depression (r = -0.7, P <0.001), body image (r = -0.4, P <0.001), sexuality (chi\(^2\) = 22, P = 0.001) and self-esteem (r = -0.64, P <0.001). Similar findings were recorded in the START trials, which detected a significant association between body image and anxiety and depression\(^13\).

The START trials confirmed a steep dose response for radiotherapy complications: a 10% increase in whole breast radiotherapy dose doubled the rate of late adverse effects. A clear volume response was also observed, with the adverse effect of radiotherapy on normal tissue varying according to the volume of tissue irradiated. The magnitude of the volume response can be generated in different ways, most directly by comparing the outcome of the same dose schedule delivered to different partial volumes of breast tissue. In a retrospective study of a radiotherapy boost dose delivered using radioactive implants in 404 patients, a 4-fold increase in risk of breast fibrosis (hardness) was reported for each 100cm\(^3\) increment in boost volume, suggesting a very steep volume response\(^14\) and confirming the desirability to review the evidence for a volume effect.
These findings are consistent with univariate analysis of 364 patients randomised to a radiotherapy tumour bed boost dose after whole breast radiotherapy, which reported a hazard ratio for poor cosmesis of 0.45 (95% CI, 0.29 – 0.76) for boost volumes dichotomised to below 200 cm³ compared to boost volumes above 200 cm³. Another approach to quantify the increased risk of late side-effects is to compare the increased risk of late side-effects after a boost dose to the tumour bed compared with the same dose delivered to the whole breast. Such an analysis has been performed in 723 patients in the START pilot trial randomised to tumour bed boost dose versus no tumour bed boost. Patients randomised to a tumour bed boost of 15.5 Gy in 7 fractions had a 17% higher risk of moderate or marked breast hardness at 10 years. The same randomised trial compared two dose levels of whole breast radiotherapy. The dose of whole breast radiotherapy causing a 17% increased risk of breast hardening was estimated to be 4.5 Gy. This value is much lower than the 15.5 Gy that causes the same level of breast hardening when given to a boost volume of about 200 cm³, representing 20-30% of whole breast volume.

From previous work including our own pilot study, we estimated that the margins of a conventional tumour bed boost volume can be safely reduced by approximately 5mm in all spatial dimensions using IGRT. From our pilot study the average tumour bed boost volume required is approximately 70 cm³, reduced from 110 cm³, a reduction of the total volume of the average breast boost dose by approximately 40 cm³, and was expected (after Borger et al.) to reduce the risk of moderate to severe fibrosis by up to a factor of 1.7.

We considered a randomised trial to be a cumbersome and expensive technology to apply to address the problems discussed above. On the other hand, the lack of empirical research data justifying the widespread use of clip-based IGRT meant that the necessary resources to implement clip-based IGRT in routine clinical practice are not available, and the position was that expensive and potentially valuable equipment was often left idle.

Whatever research methodology is used, the preferred primary endpoint should measure treatment accuracy. Gains in accuracy can be derived directly (without assumptions) from data collected in an ongoing clinical trial that uses clip-based IGRT to verify treatment accuracy as part of its technical quality assurance protocol. The IMPORT HIGH trial provided a very reliable context in which to test the hypothesis that more accurate treatment
verification allows a substantial reduction in the volume of breast tissue exposed to high boost doses of radiotherapy. By generating direct estimates of the mean volume of breast spared by daily clip-based IGRT, it should be possible to estimate the expected reductions in late adverse effects. These estimates were largely based on published results of randomised trials conducted by members of our collaboration\textsuperscript{4,11,16,17}. A consequence of this is the possibility to estimate the degree to which dose could be safely escalated in the group of patients at highest risk of local recurrence, and the predicted consequent benefits in terms of improved local tumour control.

The study aimed to quantify the benefits of clip-based IGRT using titanium clips in breast cancer patients using a study design that did not jeopardise patient care. Accurate tumour bed localisation has been shown to be important to ensure the accuracy of whole breast radiotherapy\textsuperscript{7,18}. Also the results of this study should be applicable for all patients with breast cancer, including those prescribed partial breast radiotherapy. Proven, quantified benefit from clip-based IGRT for breast cancer patients with higher risk of recurrence would justify the routine adoption of clip-based IGRT in all UK centres and hence ensure equity of access and optimal treatment for all breast cancer patients.

**Research Hypotheses**

1. Clip-based IGRT provides more accurate method of locating the tumour bed compared to standard imaging.
2. Clip-based IGRT allows smaller margins compared to standard imaging, which translate into less toxicity and improved quality of life for patients.

**Research objectives**

The primary objective was:

To compare the spatial accuracy of breast radiotherapy based on imaging i) titanium surgical clips implanted in the tumour bed (clip-based IGRT) and ii) bony anatomy and lung position during curative radiotherapy for early breast cancer (standard imaging).

Secondary objectives were:
Comparison of standard imaging with clip-based IGRT for:

I. Adequate RT safety margins around the tumour bed to avoid geographical miss

II. Volume of breast tissue irradiated around tumour bed

III. Estimated breast toxicity following development of a normal tissue control probability model

IV. Time taken for each imaging method
Chapter 2 – A critical review of the relationship between irradiated breast volume and late breast tissue complications

This chapter describes a critical review of the literature which aimed to understand the quality and scope of evidence for the relationship between irradiated breast volume and late breast tissue complications. This work supported a secondary objective of this study, the estimation of the reduction in risk of late breast toxicity which results from using clip-based IGRT, and helped to inform clinical recommendations based on all study outcomes.

This chapter is based on the peer-refereed scientific journal paper:

Introduction

In radiotherapy the aim is to deliver a tumoricidal dose for optimal loco-regional control whilst maintaining relative sparing of the surrounding normal tissues. Accurate knowledge of the tumoricidal and tolerance doses to the various tissues along with the effects of irradiating partial volumes of organs (i.e. the dose volume effect) is essential for all types of modern radiotherapy, including conformal, IMRT and IGRT.

Emami et al.\textsuperscript{19} pioneered this field with a comprehensive review of the current knowledge of radiation tolerance doses for normal tissues. This included quantification of late normal tissue complication probability (NTCP) as a function of the volume of the organ irradiated. Although this review was informative it was limited by the available data, with most of the dose volume data based on interpolated or extrapolated from whole organ data, or based on the expert experience of the involved clinicians. Since that work was published, an update on the dose volume effect of radiation on the normal tissues has been published in the “Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC)” report\textsuperscript{20}. This can be utilised in treatment planning to estimate the effect of irradiated volume on tissue tolerance\textsuperscript{21,22}. 
For many years, the radiation dose-volume effect for the breast has been exploited in the boost treatment for breast cancer patients who are at high risk of recurrence. This involves treating a small volume of breast tissue to a higher dose with improved local control rates\(^2\). More recently, the dose-volume effect has been exploited in trials of Partial Breast Irradiation (PBI) for patients at low risk of recurrence. In PBI the irradiated volume is confined to the region around the tumour bed and the aim is to reduce toxicity whilst maintaining the local control rate. However, there is paucity of published data on the dose-volume effect of irradiation on breast tissue including the QUANTEC report.

The critical review presented in this chapter of the report evaluates the evidence for a relationship between the volume of breast tissue irradiated and late complications rates. These complications include cosmesis, breast fibrosis, breast induration and telangiectasia. It also explores the evidence for a modest reduction of dose to part of the breast allowing dose escalation to the tumour bed, with lower NTCP expected.

**Methods**

Medline and Embase were used to perform a literature search with the following search strategy “Breast neoplasm” AND “radiotherapy OR Irradiation”. This was combined with “AND fibrosis”, “AND cosme*”, “AND side effect*”, “AND toxicity”, “AND shrinkage” and “AND normal tissue”. The search was then expanded to include related articles and a reference list of articles and was conducted from 1st January 1975 up to 1st May 2012.

**Impact of boost volume on breast tissue complications**

**EORTC 22881-10882 “boost versus no boost” trial (level I evidence)**

In the EORTC “boost versus no boost” trial 5318 patients with early breast cancer were randomised between a tumour bed boost of 16 Gy and no boost after whole breast irradiation (WBI)\(^2\). The boost dose was delivered using electrons or tangential photon fields with a daily fractionation of 2Gy, or using Ir192 implant at a dose rate of 0.5 Gy per hour. The results showed that, at 10 years, the use of tumour bed boost of 16 Gy increased the rates of moderate to severe breast fibrosis by 15% (28.1% versus 13.2%; p <0.0001). 251 patients with microscopically incomplete tumour excision were also randomised to either a low dose
boost of 10 Gy (126 patients) or a high dose boost of 26 Gy (125 patients). The cumulative incidence of moderate/severe fibrosis for low dose and high dose boost at ten years were found to be respectively 24% and 54%. Thus dose escalation of 16 Gy to the boost volume in the incomplete tumour excision group increased the rates of moderate or severe fibrosis by 30%, compared with a 15% increase in the complete excision group for the same 16 Gy increase in dose.

The boost volume for the complete excision group was tumour bed plus 1.5 cm margin as compared to tumour bed plus 3 cm margin in the incomplete tumour excision group. This suggests a dose volume relationship for breast tissue, as an increase in irradiated breast volume in the incomplete excision group doubled the risk of moderate/severe fibrosis for the same dose escalation of 16 Gy. However, it may be the case that the increase risk of breast fibrosis is secondary to a combination of larger boost volume and a steeper dose response curve as the total dose was increased up to 76 Gy in the incomplete excision group. The trial group also reported that on univariate analysis, patients with large boost volume are more likely to develop sub-optimal cosmesis at 3 years and breast fibrosis at 10 years. However, boost volume was not a significant variable affecting fibrosis and cosmesis in multivariate analysis.

Brachytherapy boost (level IV evidence)

Borger et. al. reported on the dose-volume effect of brachytherapy boost for breast fibrosis. The patient group was 404 patients who were treated with external beam radiotherapy (50 Gy in 2 Gy daily fractions to the whole breast). This was followed by an iridium implant boost (dose rate 0.57± 0.11 Gy/hour). Brachytherapy doses fell into three groups: 15 Gy for 101 patients, 25 Gy for 301 patients and 20 Gy for 2 patients. With a median follow-up period of 70 months, a four times higher risk of fibrosis was observed for each 100 cm³ increase in irradiated boost volume, and a tenfold higher risk of fibrosis was observed when the total dose exceeded 79 Gy compared to doses below 70 Gy.

In another study from Georgetown University Medical Centre, McRae et al. also reported on the relationship between brachytherapy boost volume and soft tissue complication. Retrospective brachytherapy plans for 5 patients with radiation induced soft tissue damage were compared with 51 patients who did not experience severe complication after WBI followed by Iridium-192 boost. The mean boost volume for patients who developed soft
tissue damage was significantly higher for all dose levels between 10 Gy and 50 Gy when compared to patients with no reported complications (p<0.05), suggesting a relationship between volume and NTCP at any dose studied. Similarly, Olivotto et. al.\textsuperscript{28} reported an association between the brachytherapy boost volume and late cosmetic outcome for 497 patients who received WBI (46 to 50 Gy over 4.5 to 5 weeks) followed by a low dose rate boost with Iridium-192 to bring the tumour bed dose to 60Gy. At a median follow up of 76 months, the boost volume boost, measured by the number of Iridium seeds used, was found to be a significant factor for fair/poor cosmesis. Patients with <70 seeds had a 15% risk of fair/poor cosmesis compared to 38% for patients with ≥100 seeds (p<0.01). The use of a greater number of seeds implies a larger volume of irradiated breast tissue, and hence a radiation volume effect for cosmesis. Several other single and multi-centre studies have reported on the relationship between volume of brachytherapy boost and NTCP risk and are summarised in Table 1.

\textit{Intra-operative RT (IORT) boost using low energy X-ray (level IV evidence)}

IORT uses low energy x-ray of 50 kV and can be used to deliver a single-fraction, high-dose radiation boost to the tumour bed after lumpectomy. Advocates for IORT cite several potential advantages of using this approach: including delivery of radiation immediately after surgery to prevent tumour cell proliferation; change in cytokines pattern into a less stimulating microenvironment, which is expected to reduce the local recurrence rates; and reduced risk of geographical miss\textsuperscript{29,30}.

The University of Heidelberg reported on late toxicity data (at 3 years) for 79 cases treated with the IORT method\textsuperscript{31}. All patients received 20 Gy intra-operative boost using a 50 kV x-ray set followed by 46-50 Gy in 2Gy daily fraction of WBI with or without supra/infra-clavicular fossa irradiation. 35% of patients developed grade 2-3 breast fibrosis. They observed that size of the applicator used for IORT significantly correlated with late breast fibrosis (spearman rank correlation coefficient 0.496, p<0.001). A larger applicator size would imply a larger volume of irradiated breast tissue, which in turn would suggest a radiation volume effect on late breast tissue toxicity.

\textit{Cobalt unit based boost (level IV evidence)}

Dewar et. al. from the Institute Gustave-Roussy reported on cosmetic outcome after breast-conserving surgery and radiotherapy\textsuperscript{32}. 592 patients received WBI (45 Gy in 2.5 Gy per
fraction, four times weekly). They were treated using two tangential fields, with each field treated on alternate days. This was followed by tumour bed boost of 15Gy in 6 fractions using on a cobalt unit. Multivariate analysis showed that in addition to applied dose per fraction, the area of field to the tumour bed (>30 cm$^3$) was associated with an increased risk of fibrosis (p<0.02) and telangiectasia (p<0.01).

*Other boost studies (level IV evidence)*

The Fox Chase Cancer Centre, Philadelphia recently presented a reported on tumour bed boost parameters associated with overall cosmesis and fibrosis for a group of 3186 patients who were treated at their centre from 1970-2008$^{33}$. All patients received whole breast radiotherapy (46-50 Gy) followed by a tumour bed boost of 10-18 Gy using either electrons or photons. Median follow-up was 78 months. Smaller boost cut-out size was found to be a borderline predictor of excellent cosmesis (p=0.05) and lower risk of breast fibrosis (p<0.0001) based on univariate analysis. Neither fibrosis nor cosmesis were found to remain significantly associated with higher field size on multivariate analysis. However, no information was available on the size of the treated boost volume and no distinction was made between physician and patient cosmetic score in their report.
<table>
<thead>
<tr>
<th>First author, Institute and radiation technique</th>
<th>Number of patients (median follow up)</th>
<th>TNM/stage</th>
<th>Comments on NTCP assessment</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borger et. al.\textsuperscript{14} Netherlands Cancer Institute</td>
<td>404 patients median follow up 70 months (range 30-133 months)</td>
<td>Stage 1-2</td>
<td>Four trained physicians scored fibrosis by palpating induration in the tumour bed. Four-scale scoring system: no fibrosis = no difference in consistency between the two breasts, grade 1 = a small difference, grade 2 = a moderate difference, grade 3 = a large difference. The scores of the four investigators were averaged to obtain the final result per patient</td>
<td>Implant volume (100% dose) associated with risk of fibrosis Odds ratio 4.2 (95% CI 2.3-8.0) per 100cm\textsuperscript{3} increase in boost volume</td>
</tr>
<tr>
<td>McRae et al.\textsuperscript{27} Georgetown University Medical centre, Washington</td>
<td>56 patients with a minimum follow up of 2.5 years</td>
<td>Stage 1-3</td>
<td>Radiation injury to connective tissue or fat necrosis requiring prolonged medical or surgical management</td>
<td>Mean boost volume significantly higher for all dose level between 10Gy and 50Gy for patients who developed soft tissue damage as compare to patients with no reported complications (p&lt;0.05)</td>
</tr>
<tr>
<td>Study</td>
<td>Institution/Methodology</td>
<td>Patient Details</td>
<td>Site/Stage Details</td>
<td>Follow-Up Details</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>Dewar et al.</td>
<td>Institut Gustave-Roussy, France WBI 45 Gy in 2.5 Gy per fraction using two tangential fields followed by tumour bed boost of 15 Gy in 6 fractions using one to two fields on the cobalt unit</td>
<td>592 patients mean follow up 78 months (standard deviation 35 months)</td>
<td>T1-2 N0-1 Fibrosis and/or telangiectasia of the whole breast/the tumour bed graded as absent, slight, moderate or severe by the radiation oncologist. Cosmetic outcome graded as excellent, good, fair and poor</td>
<td>Area of field to the tumour bed (&gt;30 cm³) associated with increased risk of fibrosis (p&lt;0.02) and telangiectasia (p&lt;0.01) on multivariate analysis. No relationship between cosmesis and area of field to the tumour bed</td>
</tr>
<tr>
<td>Olivotto et al.</td>
<td>Joint Center for Radiation Therapy, Boston WBI 46-50 Gy in 4.5-5 weeks, followed by low dose rate Iridium-192 boost (10-27 Gy)</td>
<td>497/593 with Iridium-192 boost Median follow up 76 months (range 37-186 months)</td>
<td>T1-2 N0-1 Overall cosmesis scored as excellent, good, fair or poor by the physician. Excellent if treated breast looked the same as the opposite breast, good if minimal but identifiable effects of radiation, fair when significant effects of radiation and a poor if severe normal tissue sequelae</td>
<td>Boost volume measured by number of Ir-192 seeds associated with increased risk of fair/poor cosmesis (p&lt;0.0001 for trend)</td>
</tr>
<tr>
<td>Clarke et. al.</td>
<td>Paul A. Bissinger Memorial Center for</td>
<td>64/78 patients with Iridium 192 boost Median follow up</td>
<td>Stage1-2 Cosmetic result scored as excellent (treated breast looked the same as the opposite breast), satisfactory (mild to</td>
<td>6% patients developed moderate/severe fibrosis with no correlation between fibrosis</td>
</tr>
<tr>
<td>Study</td>
<td>Institution</td>
<td>Treatment Details</td>
<td>Follow-Up</td>
<td>Cosmetic Score</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>-------------------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Radiation Therapy, Stanford WBI</td>
<td>45-55 Gy in 1.8-2.5 Gy per fraction followed by low dose rate Iridium-192 boost (18-25 Gy)</td>
<td>42 months (range 30-120 months)</td>
<td>moderate breast asymmetry with &lt; 1/3 volume loss secondary to surgery or retraction from fibrosis) or unsatisfactory (marked breast asymmetry or severe fibrosis with &gt;1/3 volume loss). Breast fibrosis scored as mild, moderate or severe.</td>
<td>and implanted boost volume. Surgical factors like poorly planned excision scar and large volume excision main factors for unsatisfactory cosmesis.</td>
</tr>
<tr>
<td>Wazer et. al.</td>
<td>Tufts University School of Medicine, Boston WBI</td>
<td>50 to 50.4 Gy at 1.8-2 Gy per fraction followed by low dose rate Iridium-192 boost of 20Gy</td>
<td>127 patients Median follow up 80 months (standard deviation 34 months)</td>
<td>Cosmetic score scored by two separate examiners as excellent = perfect symmetry and no visible distortion, good = slight distortion, visible telangiectasia or absent nipple-areolar complex, fair = moderate distortion, hyper pigmentation, prominent skin retraction, oedema or telangiectasia and poor = marked distortion, oedema, fibrosis, severe hyper-pigmentation. The lowest score was used</td>
</tr>
<tr>
<td>Wronczewska et. al.</td>
<td>Nicolaus Copernicus</td>
<td>54 patients Mean follow up 65</td>
<td>Cosmesis and breast fibrosis assessed by two doctors independently and</td>
<td>Boost volume receiving 100% (V100) was significantly</td>
</tr>
<tr>
<td>University, Poland</td>
<td>WBI 50-50.4Gy in 1-8-2Gy fraction followed by high dose rate Iridium-192 boost of 5-20Gy</td>
<td>months (range 41-89 months)</td>
<td>compared to the contralateral breast</td>
<td>associated with risk of breast fibrosis (p=0.0236)</td>
</tr>
</tbody>
</table>

*Table 1: Effect of brachytherapy boost volume on NTCP.*
Partial breast irradiation studies

WBI is considered the current standard of care following breast-conserving surgery and in the last decade, PBI has been explored as an alternative treatment to WBI in low risk patients. PBI involves irradiation of the volume of breast tissue around the tumour bed and is currently under investigation in several randomised phase II and III trials (shown in Table 2). This treatment approach is based on the rationale that the majority of local recurrences are located close to the area of surgical resection/the index quadrant, also the foci of breast disease outside the index quadrant are often new primary tumours and that irradiating a limited volume of breast would reduce treatment related morbidity.

Randomised controlled trials of Partial Breast Irradiation versus Whole Breast Irradiation (level I evidence)

Four randomised controlled trials (RCT) comparing WBI versus PBI have reported on their outcome.

A group from The Christie hospital reported in 1993. Their study randomised 708 patients with breast tumours ≤ 4 cm in diameter to PBI or WBI plus regional lymph nodes irradiation. PBI involved irradiating the tumour bed (with average field size 8 cm x 6 cm) to 40-42.5 Gy in 8 fractions over 10 days using electrons. WBI involved treating the whole breast to 40Gy in 15 fractions over 21 days using tangential fields with a matched field for regional lymph nodes. After a median follow up of 65 months, recurrence rates were higher in the PBI arm as compare to WBI arm (19.6% versus 11%; p=0.0008). The possible reasons for higher recurrence rates in the PBI arm were difficulty in defining the target volume, leading to geographical miss and the inclusion of patients with infiltrating lobular carcinoma and ductal carcinoma with an extensive intra-ductal component. Patients with PBI were also found to have significantly higher rates of marked breast fibrosis (14% versus 5%) and telangiectasia (33% vs. 12%) compared to WBI.

In another study, the Yorkshire Breast Cancer Group randomised 174 patients between WBI (with 40 Gy in 15 fractions over 21 days followed by a tumour bed boost of 15 Gy in 5 fractions) and PBI, which used a variety of techniques, including direct cobalt or caesium beams, electrons or a small mega-voltage tangential pair to a dose of 55Gy in 20 fractions
over 28 days\textsuperscript{40}. The trial closed prematurely due to poor accrual and higher loco-regional recurrence rates in the PBI group compared to the WBI group (24\% versus 9\%). It may be that the higher recurrence rate in the PBI arm was secondary to difficulty in accurate definition of the tumour bed. Treatment related morbidity with PBI and WBI has not been reported. These two studies pioneered the concept of PBI at a time when patient selection and tumour bed localisation were still very much under development. Subsequent randomised trials have more developed technology and used more stringent protocols for both of these factors.

The Hungarian National Institute of Oncology PBI trial\textsuperscript{41} and TARGIT trial\textsuperscript{42} have reported their outcomes more recently. The Hungarian PBI trial randomised 258 patients with stage T1 N0-1 Grade ≤2 breast cancer to WBI or PBI following breast-conserving surgery\textsuperscript{41}. WBI was carried out using Cobalt or photon beams to deliver a dose of 50 Gy in 2 Gy daily fractions, whilst and PBI was delivered using high dose rate (HDR) Iridium-192 brachytherapy (for 85 patients) and a dose of 36.4 Gy in 5.2 Gy per fraction over 4 days or electrons (for 40 patients) to a dose of 50 Gy in 2 Gy daily fractions prescribed to the 80\% isodose level. At a median follow-up of 122 months, the local recurrence rates were not significantly different in the two arms of the trial. The cosmetic results using Harvard criteria\textsuperscript{43} were found to be favourable in the PBI arm. The rate for a cosmesis score of excellent to good was 81\% for the PBI group and 63\% for the WBI group (p=0.0015).

The TARGIT-A trial randomised 2232 patients with early breast cancer to either i) WBI (40–56 Gy) with or without a boost of 10–16 Gy or ii) intra-operative PBI using low energy x-rays (50 kV and a dose of 20Gy to the tumour bed attenuating to 5–7Gy at 1 cm depth\textsuperscript{42}). Patients with adverse histological features, which included invasive lobular carcinoma or an extensive intra-ductal component, also received WBI without a boost in the PBI arm. At two years, the local recurrence rate was similar with no significant difference in the rate of toxicity, but the type of toxicity was significantly different between the two trial arms. The WBI arm had higher RTOG grade 3-4 toxicity for dermatitis, telangiectasia or breast pain (2.1\% versus 0.5\%; p=0.002), whereas patients receiving intra-operative PBI experienced a different range of side effects. Breast seroma needing more than three aspirations was more common in the IORT PBI group (2.1\% versus 0.8\%; p=0.012) and more patients reported skin breakdown or delayed healing, required surgical evacuation of haematoma and
intravenous antibiotics or surgical intervention for infection. Cosmetic results have yet to be reported.

**Case-matched pair studies (level III evidence)**

There have been four case match pair studies that have compared breast tissue complications for partial and whole breast irradiation.

Polgar et al. selected 45 patients prospectively with stage T1N0-1 breast cancer who were treated with PBI using HDR Iridium-192 implants to a dose of 30.3-36.4 Gy delivered in 7 fractions over 4 days and matched them to 80 patients (eligible for PBI) treated with WBI 50 Gy in 2 Gy daily fractions with or without a tumour bed boost of 10 to 16 Gy. Analysis at a median follow up of 7 years, showed no significant difference in the ipsilateral breast recurrence rates in the two groups. Excellent or good cosmesis measured using Harvard criteria was seen in 84.4% patients in the PBI arm and 68.3% patients in the WBI arm (p=0.04). However, a trend of increased incidence of RTOG grade 2-3 fibrosis was seen in the PBI group compared to the WBI group without a boost dose (20% versus 5.8%; p=0.06).

The William Beaumont group matched 174 patients treated with PBI (with low dose rate I125 implants delivering 50 Gy over 96 hours at a dose rate of 0.52 Gy/hour or HDR implants delivering 32 Gy in 8 fractions, each separated by 6 hours), with 174 patients treated with WBI with a median total tumour bed dose of 60 Gy. At a follow up of 36 months, cosmetic outcome was more favourable in the PBI group than the WBI group (excellent or good cosmesis was seen in 90% versus 83% of patients) This was not found to be statistically significant (p=0.17).

King et. al. matched 51 patients treated with PBI delivered with low dose rate Ir192 implants to achieve 45 Gy over 4 days (or HDR implants of 32 Gy in 8 fractions over 4 days) to 94 patients treated with WBI to a mean dose of 59 Gy following breast-conserving surgery. A blinded panel of experts scored photographic assessment of cosmesis on a four-part scale (excellent, good, fair, poor). At 20 months follow up, 75% patients in the PBI group and 84% patients with WBI had excellent or good cosmesis (not statistically significant). Grade I and II treatment complications including skin erythema, desquamation, discoloration, hyperpigmentation, dimpling; breast pain, tenderness, shrinkage or fibrosis were significantly more common in the WBI arm than the PBI study arm (80% versus 22%, p=0.001). Grade III
treatment complications requiring surgical intervention were not found to be different in the two groups (8% versus 5%, p=not significant).

Tata Memorial Hospital, India matched 27 patients treated with PBI using HDR brachytherapy (34 Gy in 10 fractions over 6-8 days) with 67 patients treated with WBI (45 Gy in 25 fractions over 5 weeks followed by a tumour bed boost using electrons (15 Gy in 6 fractions or interstitial HDR brachytherapy with a single 10 Gy fraction)\(^4\)). They reported that at a median follow-up of 43 months, cosmetic outcome was superior in the PBI group compared to the WBI group (excellent or good cosmesis was 88.9% versus 56%; p=0.003). No significant difference was seen in the rates of moderate or severe breast fibrosis.

**Effect of treatment volume on NTCP in PBI series**

There are several publications reporting on the efficacy and low toxicity achievable with PBI, but only a few evaluate the impact of treatment volume on NTCP. The current literature on the volume effect of PBI for 3D-CRT/IMRT, electrons and brachytherapy is summarised below.

**3D-CRT/IMRT based PBI (level IV evidence)**

Jagsi et. al.\(^4\)\(^8\) presented the cosmetic outcome of 32 patients treated with PBI using IMRT at deep inspiration breath hold. The patients received 38.5 Gy twice daily over five consecutive days. At a median follow up of 2.5 years, 22% patients were scored as having unacceptable cosmesis. Retrospective comparison between patients with acceptable and unacceptable cosmesis showed that the mean percentage volume of the breast receiving a minimum of 100% of the prescribed dose i.e. 38.5 Gy (V100) was lower in patients with acceptable cosmesis as compared to patients with unacceptable cosmesis (15.5% versus 23.0%; p=0.02). The mean percentage volume of breast receiving a minimum of 50% of the prescribed dose i.e. 19.25 Gy was also smaller in the group with acceptable cosmesis compared to unacceptable cosmesis group (p=0.02).

Hepel et. al.\(^4\)\(^9\) also reported on a positive correlation between the volume of breast tissue treated with PBI and cosmesis outcome. 60 patients were treated with PBI to a dose of 38.5 Gy in twice daily fractionations over one week using 3D-CRT. At a median follow up of 15 months, 18% patients developed fair or poor cosmesis and 25% developed Grade 2-4 subcutaneous fibrosis. In univariate analysis, the ratio of the size of 3D-CRT target volume to
the whole breast volume was found to correlate with fair or poor cosmesis (p=0.02) and with grade 2-4 subcutaneous fibrosis (p=0.10). Reference 48 and 49 suggested an association between breast volume irradiated in PBI and normal tissue complication rates.

In contrast, Chen and colleagues from the William Beaumont group reported no association between overall cosmesis and the ratio of the size of 3D-CRT target volume to the whole breast volume. In their study 94 patients received PBI with a dose of 38.5 Gy in twice daily fractions over five consecutive days using 3D-CRT. Of the 56 patients with cosmesis assessment of greater than 48 months, 11% had fair to poor cosmesis and 3% had Grade 3 fibrosis with no association between cosmesis or subcutaneous toxicity and this ratio.

**Single source brachytherapy and multi-source brachytherapy (level IV evidence)**

Multi-source brachytherapy has been used for PBI for many years. Most publications on the results of this technique focus on local control rates and there is limited reporting of normal tissue toxicity. Some have reported on factors associated with normal tissue toxicity and have commented on a positive correlation between NTCP and the implant volume. Yeo et. al. reported on the efficacy and safety of PBI using multi-source brachytherapy for 48 patients with a median follow up of 53 months. A dose of 34Gy in 10 fractions over five days was delivered to the tumour bed plus a margin of 1 to 2 cm. 14% of patients developed Grade 2 subcutaneous toxicity with V100 and V150 significantly higher in these patients (p=0.018 and 0.034 respectively). No patients were found to have poor cosmesis.

Wazer et al. studied late toxicity and long term cosmetic outcome after multi-source brachytherapy PBI using pooled data from Tufts University, Brown University and Virginia Commonwealth University. The number of dwell positions, a determinant of total volume of implanted breast tissue, was found to correlate with late cosmetic outcome (p=0.04).

Lawenda and colleagues found no association between implant volume and overall cosmetic outcome for 48 patients treated with LDR brachytherapy at their centre from 1997-2001. The purpose of the study was to evaluate the effects of dose escalation in PBI. The dose was escalated in three groups of 50 Gy, 55 Gy and 60 Gy and implant volume was divided into four groups. A trend between dose escalation and fibrosis was seen (not found to be significant). They also observed a decline in the incidence of breast fibrosis with increase in implant volume, a finding contrary to other published literature.
<table>
<thead>
<tr>
<th>Trial/Institute</th>
<th>Control arm (WBI)</th>
<th>Test arms (PBI): treatment modality</th>
<th>Median follow up (months)</th>
<th>Target accrual</th>
<th>Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christie group trial&lt;sup&gt;39&lt;/sup&gt;</td>
<td>WBI 40Gy in 15 fractions with matched field for regional nodes</td>
<td>PBI: 40-42.5Gy in 8 fractions using electrons</td>
<td>65 months</td>
<td>708</td>
<td>Yes</td>
</tr>
<tr>
<td>Yorkshire Breast Cancer Group trial&lt;sup&gt;40&lt;/sup&gt;</td>
<td>WBI 40Gy in 15 fractions with 15Gy boost</td>
<td>PBI using direct cobalt, caesium or electrons beam or a small mega-voltage tangential pair to a dose of 55Gy in 20 fractions</td>
<td>96 months</td>
<td>174 (pre-mature closure)</td>
<td>Yes</td>
</tr>
<tr>
<td>Hungarian National Institute of Oncology&lt;sup&gt;41&lt;/sup&gt;</td>
<td>WBI using Cobalt or photons beam to a dose of 50Gy in 25 fractions over 5 weeks</td>
<td>HDR Ir-192 (85 pts) to a dose of 36.4Gy in 7 fractions over 4 days or Electrons (40 pts) to a dose of 50Gy in 25 fractions prescribed to the 80% isodose</td>
<td>66</td>
<td>258</td>
<td>Yes</td>
</tr>
<tr>
<td>TARGIT&lt;sup&gt;42&lt;/sup&gt;</td>
<td>WBI 40–56Gy with optional boost of 10–16Gy</td>
<td>PBI: 20Gy single fraction using Intra-operative 50 KV photons</td>
<td>24 months</td>
<td>2232</td>
<td>Yes</td>
</tr>
<tr>
<td>ELIOT&lt;sup&gt;55&lt;/sup&gt;</td>
<td>WBI 50Gy in 25 fractions with 10Gy boost</td>
<td>PBI: Intra-operative electrons 21Gy in single fraction</td>
<td>NA</td>
<td>1300 (closed 2007)</td>
<td>No</td>
</tr>
<tr>
<td>IMPORT LOW&lt;sup&gt;56,57&lt;/sup&gt;</td>
<td>WBI 40Gy in 15 fractions, no boost</td>
<td>Arm 1: 36Gy in 15 fractions to the low risk volume of the breast and 40Gy in 15 fractions to the</td>
<td>NA</td>
<td>2000 (closed 2010)</td>
<td>No</td>
</tr>
<tr>
<td>Study</td>
<td>WBI Dose and Schedule</td>
<td>PBI Dose and Schedule</td>
<td>N/A</td>
<td>Patient Number</td>
<td>Activated Year</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>-----</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>GEC-ESTRO[^58]</td>
<td>50-50.4Gy in 25-28 fractions with 10Gy optional boost</td>
<td>PBI: 32 Gy in 8 fractions or 30.3Gy in 7 fractions HDR or 50Gy PDR</td>
<td>NA</td>
<td>1170</td>
<td>2004</td>
</tr>
<tr>
<td>NSABP-39[^59]</td>
<td>50-50.4Gy in 25-28 fractions with 10-16Gy optional boost</td>
<td>PBI: 34Gy in 10 fractions over five days using single/multi-source brachytherapy or 38.5Gy in 10 fractions over 5 days using 3D-CRT</td>
<td>NA</td>
<td>4300</td>
<td>2005</td>
</tr>
<tr>
<td>RAPID[^60]</td>
<td>42.5Gy in 16 fractions with optional 10Gy boost</td>
<td>PBI: 38.5Gy in 10 fractions BD over 5-8 days using 3D-CRT</td>
<td>NA</td>
<td>2128</td>
<td>2006</td>
</tr>
<tr>
<td>IRMA[^61]</td>
<td>45Gy in 18 fractions or 50Gy in 25 fractions or 50.4Gy in 28 fractions with optional 10 – 16Gy boost</td>
<td>PBI: 38.5Gy in 10 fractions BD over 5 days using 3D-CRT</td>
<td>NA</td>
<td>3302</td>
<td>2007</td>
</tr>
<tr>
<td>SHARE&lt;sup&gt;53&lt;/sup&gt;</td>
<td>WBI 50Gy in 25 fractions + 16 Gy boost or WBI 40-42.5Gy in 15-16 fractions without boost</td>
<td>PBI: 40Gy in 10 fractions BD over 5 to 7 days using 3D-CRT</td>
<td>NA</td>
<td>2796 (activated 2010)</td>
<td>No</td>
</tr>
</tbody>
</table>

*Table 2: Phase II-III randomised controlled trials comparing WBI versus PBI*
Breast fractionation studies

The Royal Marsden Hospital and Gloucestershire Oncology Centre (RMH/GOC) trial\(^\text{16}\) randomised 1410 early breast cancer patients between three fractionation schedules for WBI. The control arm treatment was 50Gy in 25 fractions delivered in 5 weeks. Two test arms were used: a) 39Gy in 13 fractions over 5 weeks, and b) 42.9Gy in 13 fractions over 5 weeks. The equivalent doses in 2 Gy fractions (EQD2) (using an \(\alpha/\beta\) ratio of 3.1 Gy for palpable breast induration) were calculated to be: 46.7 Gy and 53.8 Gy respectively. The risk of moderate to severe induration at 10 years for the two test arms was 27\% and 51\%, respectively, suggesting a 24\% increased risk of induration with a dose escalation of 7 Gy to the whole breast (or 3.3 \% increased risk per Gy). Compared to this fractionation effect, an escalated dose to the tumour bed alone i.e. a boost of 15.5 Gy delivered in 7 fractions (EQD2 of 16 Gy) increased the risk of induration by 17\% (equivalent to a 1.05\% increase per Gy). The increased risk of induration, per Gy of dose seen with increased breast volume irradiated indicates a radiation volume-effect for breast tissue.

Dose modulating effect on the breast

The dose volume effect in normal-tissue can be exploited therapeutically by radiating a small volume of tissue to a higher dose and reducing the overall dose to the rest of the organ. This has been successfully demonstrated in prostate cancer radiotherapy with IMRT\(^\text{64}\). A trial at Saint George and Wollongong in Sydney suggests that this modulation effect is also present in breast tissue\(^\text{65}\). This trial randomised 688 patients with stages T1-2N0-1 breast cancer between a standard arm of WBI (50 Gy in 2 Gy daily fractions and no boost) and test arm of WBI (45 Gy in 1.8 Gy daily fractions plus a 16 Gy tumour bed boost). The overall cosmesis was scored by a five-person panel using digital photographs with a scale of excellent, good, fair and poor. 79\%
patients in the test arm with boost and 68% patients in the standard arm had excellent or good cosmesis (p=0.016). The rate of moderate to severe breast fibrosis at five years was similar in the treatment arms. These results are contrary to the current literature of worse cosmetic outcome and higher rates of breast fibrosis with additional boost radiation. A possible explanation for these results is that a modest reduction in dose to the whole breast allowed dose escalation to the tumour bed without the expected increase in normal tissue toxicity and provides evidence of a volume effect.

**Limitations of this review**

Late breast tissue toxicity post radiotherapy is influenced by several patient and treatment related factors. Many of the studies reviewed have not accounted for other confounding factors including extent of surgical excision, total delivered dose, dose fractionation, post-operative complications and brachytherapy dose inhomogeneity, e.g. surgical excision volume and baseline surgical cosmesis are significant factors affecting cosmesis. Larger surgical excision would also imply larger brachytherapy boost and/or target volume and a larger applicator size for IORT. Based on the current reports, it is difficult to draw strong support on the independent volume effect on late breast tissue complications.

A variety of treatment approaches have been used including photons, electrons, intra-operative techniques and brachytherapy. In addition, the reported studies have used different endpoints (fibrosis, cosmesis and telangiectasia) with several different scoring methods and a range of periods at which follow-up were obtained. These factors all make it difficult to draw firm conclusions on the dose-volume relationship for breast tissue.
Conclusions

Quantitative effect of treatment volume
The study by Borger et al. which used low dose rate iridium implants provided the most robust quantitative data on the dose-volume relationship. For every 100 cm³ increase in the boost volume, the risk of fibrosis increased by a factor of four and a two-fold increase in boost volume results in an 11% reduction in tolerance dose (NTD₅₀). It is however difficult to be certain as to how the low dose rate brachytherapy data can be extrapolated to other techniques: HDR brachytherapy, electron and photon boost techniques. The RMH/GOC trial which used electron boost provides indirect quantitative information on the dose volume relationship for NTCP. For every Gy increase in boost dose, the risk of moderate to severe breast induration was found to increase by 1% as compared to 3% when the whole breast dose is increased by 1 Gy.

Qualitative effect of treatment volume
The results from the Hungarian PBI trial and TARGIT trial provides strong qualitative evidence of a dependence of NTCP on volume irradiated. These studies report both improved cosmetic outcome and reduced NTCP in the PBI arm compared to WBI. However, these are significant differences in the radiotherapy techniques and fractionation schedules used by these two groups, making it difficult to draw quantitative conclusions on the radiation volume effect on breast tissue. The other reported randomised trial from Christie reported higher rates of breast fibrosis and telangiectasia in the PBI arm. A dose-response relationship for late radiation effects including telangiectasia and breast fibrosis is well known and these dissimilar results may be explainable if one calculates the 2Gy equivalent dose (EQD2) for the PBI and WBI groups using an α/β ratio of 3.1 for fibrosis. The WBI group received
a lower dose of 45 Gy (EQD2), compared to 63-70 Gy for the PBI group in the Christie study.

The four matched case series\textsuperscript{44-47} which compared PBI and WBI also showed favourable cosmesis and lower NTCP risk with PBI with the exception of higher grade 2-3 fibrosis observed in the Hungarian series\textsuperscript{44}. It is possible that significant dose heterogeneity with the use of Ir192 implants could explain the increased grade 2-3 fibrosis in the PBI arm in the Hungarian series. As in the randomised trials, this study evaluated PBI and WBI using different radiotherapy techniques and fractionation.

**Future work**

The current literature suggests that treatment volume is an important parameter affecting late breast tissue complications but that more robust data is needed. This is expected to come from the below mentioned randomised trials which will quantify the impact of volume of breast irradiated on NTCP.

**Randomised controlled trials of Partial Breast versus Whole Breast Irradiation**

IMPORT LOW and The Danish Breast Cancer Cooperative Group trial (which has not yet reported) are two randomised trials comparing PBI and WBI, with volume of breast irradiated as the solitary randomisation variable. IMPORT LOW is a randomised Phase 3 trial comparing WBI with two dose levels delivered as PBI using IMRT in women with low risk of recurrence from their breast cancer. It has completed target accrual of 2000 patients in 2010\textsuperscript{56-57}. The control arm is WBI delivering 40 Gy in 15 fractions over 3 weeks to the whole breast. Test arm 1 delivers synchronously 40 Gy in 15 fractions to the partial breast PTV and 36Gy in 15 fractions to the remainder of the whole breast. Arm 2 uses PBI to deliver 40 Gy in 15
fractions to the partial breast PTV alone (see figure 1). The primary endpoint is local tumour control in the treated breast. Secondary endpoints include location of tumour relapse, occurrence contralateral primary tumours, regional and distant metastases, late adverse effects in normal tissues, quality of life (QOL) and economic evaluation.

The trial implemented by the Danish Breast Cancer Cooperative Group is a Phase 2 study comparing PBI to WBI in low risk breast cancer patients. Both treatment arms receive 40Gy in 15 fractions over 3 weeks. The primary endpoint for this study is grade 2-3 breast fibrosis after radiotherapy. Secondary endpoints are other late morbidity, local recurrence and genetic risk profiling for development of late radiation morbidity. The results on these two trials in terms of late normal tissue effects will not become available for several years, but are expected to provide definitive data regarding the effects of irradiated breast volume on normal tissue effects.
The 3D-CRT/IMRT based PBI series\textsuperscript{48-50} discussed above have produced conflicting reports on the relationship between the treated volume and NTCP. The mature data from the on-going Phase 3 NSABP B-39/RTOG 0413 trial (n=4300) is expected to provide more definitive data on whether an association between breast volume irradiated in APBI and normal tissue complications exists.

\textbf{CONTROL}:
Whole breast irradiation, 40Gy in 15 fractions over 3 weeks

\textbf{Arm 1}:
36Gy in 15 fractions to the low risk volume of the breast and 40Gy in 15 fractions to the index quadrant over 3 weeks

\textbf{Arm 2}:
Partial breast irradiation, 40Gy in 15 fractions over 3 weeks to the index quadrant only

\textit{Figure 1: IMPORT LOW trial schema}
Dose modulation effect on the breast

This dose modulating effect on the breast tissue is further investigated in the IMPORT HIGH trial (of which this study is sub-study)\textsuperscript{56,57}. The trial randomises patient at higher risk of recurrence between three groups: a) a standard arm of 40 Gy in 15 fractions to the whole breast over 3 weeks with a sequential tumour bed boost of 16 Gy in 2 Gy daily fractions; b) test arm 1 of 36 Gy in 15 fraction to the low risk volume of the breast, 40 Gy in 15 fractions to the index quadrant plus a concomitant tumour bed boost of 48 Gy in 15 fractions; and c) Test arm 2 of 36Gy in 15 fractions to the low risk volume of the breast, 40 Gy in 15 fractions to the index quadrant plus a concomitant tumour bed boost of 53 Gy in 15 fractions. The trial planning schema is shown in figure 2. This trial tests the hypothesis that decreasing the radiation the dose to the whole breast volume by a very small amount (40 Gy to 36 Gy) and treating an iso-effective dose to the index quadrant and tumour bed (Test arm 1), may result in less normal tissue side effects compared to the control group. It will also test if decreasing the radiation dose to the whole breast tissue by a very small amount allows dose escalation to the tumour bed (which area of highest risk of local recurrence) without an increase in normal tissue side effects (Test arm 2).
**Figure 2: IMPORT HIGH trial schema**

**Control**: 40Gy in 15 fractions to the whole breast followed by a sequential photon boost of 16Gy in 8 fractions to the tumour bed (Total- 56Gy in 23 fractions sequential dose)

**Test Arm 1 and 2**: 36Gy in 15 fractions to the low risk volume of the breast, 40Gy in 15 fractions to the index quadrant and dose escalation to the tumour bed with two dose levels of 48Gy and 53Gy in 15 fractions as concomitant boost
Chapter 3 – Normal tissue complication modelling for breast tissue

This chapter describes the development of a normal tissue complication probability (NTCP) model which may be used to relate the volume of normal tissue irradiated to the risk of toxicity. This model is required to address a secondary objective of this study: to estimate the reduced risk of late adverse effects resulting from the smaller tissue volume irradiated.


Introduction

The goal of this stage of the work was to develop an NTCP model for breast tissue and predict the probability of complication to quantify the dose-volume effect for a non-uniform irradiation of the patient’s breast. Fibrosis is a common sequela of breast RT and adversely affects overall cosmesis, it can be assessed using a scoring system and has been shown to impact on patient physical and psychological wellbeing.\textsuperscript{12} Hence, this work relates to the development of NTCP model for moderate-severe breast fibrosis.

The hypothesis in this work was that breast tissue displays a significant dose-volume effect to radiation which manifests itself as moderate to severe fibrosis and that an NTCP model can effectively predict the probability of breast fibrosis based on the interaction of radiation dose and the treatment volume.
Materials and methods

This analysis required diverse datasets with a range of dose and volume data plus quantitative toxicity endpoint data. Randomised clinical trials (RCTs) were felt to provide the most robust data to this end. In addition pooling data from a set of RCTs was felt to increase the diversity of the dataset and enable generalisation of the findings to a broader population. Moderate to severe breast fibrosis was chosen as the toxicity endpoint for this study.

The principal investigators of three trials kindly agreed to collaborate by sharing their patient data.

i) Dr. Charlotte Coles from the Cambridge Breast IMRT trial

ii) Professor Harry Bartelink from the EORTC 22881-10882 “boost versus no boost” trial

iii) Dr. Peter Graham from the St. George and Wollongong trial

Preliminary assessment of the available data, showed that the moderate-severe breast fibrosis rate of 3% with 50 Gy WBI in the St George and Wollongong trial was smaller than that in the published literature (including the Cambridge and EORTC trials). Hence, for this study, only individual patient data from the EORTC 22881-10882 “boost versus no boost” trial and the Cambridge Breast IMRT trials were pooled.

Patient cohort details and toxicity scoring – Cambridge

This was a single centre trial which recruited 1145 patients with invasive breast cancer (stage T1-T3N0-1M0) or ductal carcinoma in situ who received breast conserving surgery and radiotherapy. All patients received 40 Gy in 15 fractions over 3 weeks to the whole breast, which was followed by an electron tumour bed boost of 9
Gy in 3 fractions over 3 days in selected cases \((n = 728)\). The level of breast fibrosis was assessed clinically at 2 and 5 years after completion of RT and scored on a four point scale \((0 = \text{none}, 1 = \text{a little}, 2 = \text{moderate} \text{ and } 3 = \text{severe})\).

**Patient cohort details and toxicity scoring – EORTC**

This was a multi-centre trial that recruited 5569 patients with invasive breast cancer \((\text{stage T1-T2N0-1M0})\) who received surgery and radiotherapy. All patients received 50 Gy in 25 fractions over 5 weeks to the whole breast and were randomised by three boost levels: i) no boost \((n = 2657)\), ii) 10 Gy in 5 fractions boost \((n = 126)\), iii) 16 Gy in 8 fractions boost \((n = 2661)\), and iv) 26 Gy in 13 fractions boost \((n = 125)\). The boost was delivered using: electrons \((63\%)\), photons \((29\%)\) and low dose rate brachytherapy \((9\%)\). Breast fibrosis was assessed clinically at each follow up visit and scored on a four point scale \((1 = \text{none}, 2 = \text{minor}, 3 = \text{moderate}, \text{and } 4 = \text{severe})\).

**Exclusion criteria**

As brachytherapy may lead to significant dose heterogeneity and the boost volumes used are usually much smaller than external beam techniques\(^73\) patients with brachytherapy boost were excluded from the analysis as were patients with missing data or toxicity scores \((\text{Cambridge trial: 571 and EORTC trial: 275})\).

**Dose-Volume data**

The accuracy with which NTCP model parameters can be estimated depends on the quality of dosimetry and follow-up data. The late toxicity scores and boost volumes were recorded in both trials but limited data on the dose distributions were available. Consequently, a more simplistic two-compartment dose-volume histogram (DVH) model was used. The first step of the DVH was the tumour bed volume receiving the whole breast dose plus the boost dose and the second step of the DVH was the
remaining breast volume (whole breast volume minus tumour bed volume) receiving whole breast dose only (figure 3).

![Diagram showing two step dose volume histogram model]

**Figure 3: The two step dose volume histogram model**

DVH step 1: Tumour bed receiving the whole breast dose plus the tumour bed boost
DVH step 2: Whole breast volume minus the tumour bed receiving whole breast dose alone

The whole breast volume was only recorded in the Cambridge trial. Based on the Cambridge data, estimates of the whole breast volume for the EORTC trial patients could be estimated. Hence, a Monte Carlo (MC) simulation method was written which generated breast volume data for these patients. The MC simulation used the breast volume distribution from the Cambridge trial and an acceptance-rejection test for the randomly generated volumes to ensure the ratio of boost/breast volume was between 5-40% (the range of boost volume to breast volume ratio observed in the Cambridge data). In doing this the assumption was made that the distribution of breast volume
and the ratio boost/breast volume in the EORTC trial was the same as that in the Cambridge trial.

NTCP modelling
In this study two established radiobiological models were used: Lyman Kutcher Burman (LKB) model\textsuperscript{74} and the Niemierko model\textsuperscript{75}. Both assume that the dose-response follows a sigmoid curve. Both describe the response with three parameters:

TD50: the homogeneous dose to the organ that leads to 50\% patients experiencing the defined toxicity at 5 years

γ50/m: the steepness parameter of the dose-response curve

n: volume parameter of the organ being assessed

To estimate these parameters, each patient’s two-compartment DVH was converted into a generalised equivalent uniform dose (EUD) using the Kutcher-Burman histogram reduction method. The EUD is the dose, when delivered uniformly to the organ, will lead to the same complication probability as the actual dose distribution.

\[
EUD = \left( \sum_i v_i \left( D_i \right)^{\gamma} \right)^n
\]  \hspace{1cm} (1)

\( v_i \) is the i-th sub-volume of the organ irradiated with dose \( D_i \) in the differential dose-volume histogram.

For the volume parameter, \( n \), If \( n=1 \), the organ has a parallel structure with a strong volume dependence on late complication rate and EUD is the mean dose, If \( n=0 \), the organ has a serial structure with no volume dependence on late complication rate and EUD tends to be the maximum dose.
As complications due to radiotherapy depend on fraction size, a biologically equivalent uniform dose (BEUD₃) was generated using the EUD and α/β ratio of 3Gy in the linear quadratic model.

\[
BEUD₃ = EUD \left( 1 + \frac{EUD}{N \times \alpha / \beta} \right)
\]  

(2)

In the Lyman Kutcher Burman (LKB) model:

\[
NTCP = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\left(\frac{x^2}{2}\right)} \, dx
\]

(3)

where

\[
x = \frac{BEUD₃ - BEUD \, 50₃}{mBEUD \, 50₃}
\]

(4)

In the Niemierko model:

\[
NTCP = \frac{1}{1 + \left(\frac{BEUD \, 50₃}{BEUD₃}\right)^{4 \gamma \, 50}}
\]

(5)
The two NTCP models were written in Object Pascal (Delphi, Embarcadero technologies, San Francisco, CA, USA).

A Maximum Likelihood Estimation (MLE) method was used to find the best-fit values of the model parameters BEUD50, γ50/m and n. This method estimates the probability that the observed pattern of complications can be best described by the parameters of the model.

\[
\ln L = \sum_{y(i)=1} \ln [NTCP (TD 50 (1), m, n)] + \sum_{y(i)=0} \ln [1 - NTCP (TD 50 (1), m, n)]
\]

where \(y(i)=1\) if moderate or severe fibrosis is observed and \(y(i)=0\) if moderate or severe fibrosis is absent.

As discussed above, value of n value close to 1 suggest that the organ has a parallel structure with a strong volume dependence whilst a value of n close to zero suggests that the organ has a serial structure with little or no volume dependence on late complication rate. A full sequential parameter search was carried out using the following parameter constrains: BEUD3 (0-150), n (0.01-1.0), γ50 (0.5-3.0) and m (0.1-0.8). The 95% confidence intervals (CI) for the optimally fit parameters were obtained using the Profile Likelihood Estimation method. The parameter of interest was varied around its optimal values, while the other parameters were fixed in the MLE to generate the upper and lower 95% CI. This method takes non-linearity and an asymmetrical CI into consideration but does not account for correlations between parameters.

Results from the START-pilot trial were used to assess the goodness of fit of the predicted NTCP models. The START-pilot trial randomised 1410 patients into three
whole breast RT dose fractionations: i) 50 Gy in 25 fractions, ii) 39 Gy in 13 fractions, and iii) 42.9 Gy in 13 fractions. Patients were also sub-randomised for tumour bed boost to a dose of 14 Gy in 7 fractions using electrons. Cumulative data on moderate or severe breast induration at 5 years were used for all three whole breast dose fractionation regimes with and without the boost to test the goodness of fit. The goodness-of-fit statistic was obtained by calculating the Pearson chi-square statistic for the observed and predicted rates of breast fibrosis. The statistic is denoted as $\chi^2$ and a large values of $\chi^2$ (and small $p$-values) indicate a lack of fit of the model.

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

(7)

‘O’ is the observed rates of fibrosis and ‘E’ is the rate of fibrosis predicted by the model.

Results

Dose-volume and toxicity data for 574 patients from the Cambridge trial and 5282 patients from the EORTC trial were used for the NTCP modelling. 26.8% (154/574) patients in the Cambridge trial and 20.7% (1096/5282) patients in the EORTC trial developed moderate or severe breast fibrosis. The patient’s radiotherapy dose volume characteristics are summarised in table 3 below.
### Table 3: Dose-volume characteristics from the Cambridge and the EORTC dataset used for the NTCP model

The best estimated NTCP parameters for the Niemierko model from the MLE method were \( \text{BEUD}_{3}(50) = 136.4 \text{ Gy} \), \( \gamma_{50}=0.9 \) and \( n=0.011 \). The 95% CI for parameters were \( \text{BEUD}_{3}(50) = 132.8-140 \text{ Gy} \), \( \gamma_{50}= 0.84-0.97 \) and \( n= 0.01-0.03 \). The best estimated parameters for the LKB model were \( \text{BEUD}_{3}(50) = 132 \text{ Gy} \), \( m= 0.35 \) and \( n= 0.012 \) with 95% CI of \( \text{BEUD}_{3}(50) = 128.8-135.6 \text{ Gy} \), \( m= 0.326-0.374 \) and \( n= 0.01-0.03 \). The results of both models strongly imply that the risk of moderate-severe breast fibrosis
is strongly associated with radiotherapy dose and the effect of volume (i.e. the volume parameter) is small. The BEUD$_3$(50) values of 136.4 Gy and 132 Gy correspond to EQD2 values (equivalent doses in 2 Gy fractions) of 79.2 Gy and 81.8 Gy respectively.

The observed rates of moderate-severe fibrosis in the RMH/GOC trial were in good agreement to the predicted rates of fibrosis using the LKB model (see figure 4) and the Niemierko model (see figure 5). Using the Pearson chi-square test with 5 degree of freedom, the $\chi^2$ was 0.053 (p=0.95) for the LKB model and $\chi^2$ was 0.058 (p=0.95) for the Niemierko model suggesting a good fit for both models.
Figure 4: Lyman Kutcher Burman Model - The probability of moderate-severe breast fibrosis versus biological equivalent dose using $\alpha/\beta$ of 3 Gy (BED$_3$). The solid line is based on the best-fit parameters (BED$_3$ = 132 Gy and $m = 0.35$) and the dashed lines are upper and lower 95% CI. The summative toxicity data of the three dose fractionations ± boost at five years from the START pilot trial are plotted.
Figure 5: Niemierko Model - The probability of moderate-severe breast fibrosis versus biological equivalent dose using $\alpha/\beta$ of 3 Gy ($BED_3$). The solid line is based on the best-fit parameters ($BED_3 = 136.4$ Gy and $\gamma_{50} = 0.9$) and the dashed lines are upper and lower 95%CI. The summative toxicity data of the three dose fractionations ± boost at five years from the START pilot trial are plotted.

Three previous studies have estimated the NTCP parameters for breast fibrosis and these results are summarised in table 4. The Borger et al.\textsuperscript{14} model was based on 404 patients treated with WBI (50 Gy in 25 fractions over 5 weeks) followed by low dose rate Ir192 tumour bed boost (15-25 Gy). BEUD was calculated using $\alpha/\beta$ of 2 Gy and repair half-time of 1.5 hours. The positions of the implants were reconstructed on radiographs and dose distributions calculated. The best-fit model parameters in the study were found to be $TD_{50} = 72$ Gy and $n = 0.16 \pm 0.04$. The model parameters were
estimated from patients with brachytherapy boost alone and it is not clear how to compare parameters generated from brachytherapy to external beam techniques due to inherent differences in the dose distributions and possibly different radiobiological effects. Hence patients with brachytherapy boost were excluded in the current study. Avanzo et al.\textsuperscript{78} estimated the best fit parameters for the model using average values of dosimetric parameters (prescription dose, fraction dose, median follow up and dose-volume data) from three WBI studies without boost and four external beam PBI studies. Three of the PBI studies used twice daily fractionation, and BEUD calculations included a repair half-time of 4.4 hours in the model. As the median follow up of the PBI studies was short (1.3-4.2 years), a latency function correction was also included. The parameters were estimated using the weighted least square fitting method, with the number of patients in each dataset used as weighting. The parameters found for moderate and severe breast fibrosis from the model were $\text{BEUD50}= 105.8$, $n=0.15$ and $m=0.22$.

Alexander et al.\textsuperscript{79} reported that the volume parameter exhibited a strong effect on breast fibrosis. This study included summative data of 806 patients from the START pilot trial\textsuperscript{16}, 590 patients from a German study\textsuperscript{80} and 150 post-mastectomy patients treated in the 1960’s\textsuperscript{81}. All patients received WBI and no partial volume data was available for the fitting analysis. The dose-volume data were generated using an anthropomorphic phantom and parameters were estimated for a relative seriality model and Lyman model. The study suggested a parallel structure for breast tissue with a strong volume effect for breast fibrosis ($n=0.78$). However, these results cannot be generalised for several reasons:

i. The study did not account for the tumour bed boost doses (additional RT dose) in the models.
ii. Different toxicity outcome measures are used in the studies. The START pilot and German study assessed breast fibrosis based on clinical examination, whereas the post-mastectomy study used photographic assessment.

iii. The planning techniques for the post-mastectomy study based on 1960s data are outmoded by present standards. Different NTCP parameters may be expected for breast fibrosis after BCS and tissue fibrosis after mastectomy.

iv. The study corrected for latency time (START pilot and German) based on the results of the historic post-mastectomy series.
<table>
<thead>
<tr>
<th>Number of patients</th>
<th>BEUD$_3$(50)</th>
<th>$\gamma^{50}$</th>
<th>m</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>404</td>
<td>NTD$_{50}$=72 Gy</td>
<td>-</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($\alpha/\beta$=2Gy)</td>
<td>($t_{1/2}$=1.5hrs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1546</td>
<td>104 Gy</td>
<td>104 Gy</td>
<td>1.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.27</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(s=0.12)</td>
</tr>
<tr>
<td>2562</td>
<td>105.8 Gy</td>
<td>107.2 Gy</td>
<td>-</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td>5856</td>
<td>132 Gy</td>
<td>136.4 Gy</td>
<td>0.9</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.011</td>
</tr>
</tbody>
</table>

* these studies used summative dosimetric and toxicity data

NTD: Normalised total dose
BEUD$_3$(50): Biologically equivalent uniform dose using $\alpha/\beta$ of 3Gy
$\gamma^{50}$/m: slope of the dose response curve
n: volume parameter
t$_{1/2}$: repair half-time
s: describes the serial/parallel architecture of the organ. A large value indicates a serial structure and a small value indicates a parallel structure
Table 4: Summarised results of the best fit NTCP parameters for moderate-severe breast fibrosis

Discussion

A better understanding of the dose-volume effect for breast tissue is timely as many patients now receive non-uniform breast irradiation in form of a variety of modern techniques: accelerated PBI, simultaneous integrated boost and risk adapted radiotherapy\textsuperscript{56,59,82,83}. The EORTC 22881-10882 trial breast fibrosis nomogram showed a strong association between radiotherapy dose and risk of fibrosis, with large boost volumes as a prognostic factor based on univariate analysis only\textsuperscript{26}. The purpose of this study was specifically to investigate the volume effect by developing a NTCP model based on the data. This was approached by pooling individual data from two large prospective trials (5856 patients) that offered robust information on radiotherapy dose, boost volume and late toxicity.

Using the MLE method, the volume parameter ‘n’ was close to zero for both the LKB model and the Niemierko model analyses. This finding suggests that for moderate and severe fibrosis, the breast tissue behaves as a serial organ and that the maximum RT dose is the best predictor of complication. The summative data of 1410 patients from an independent dataset with six radiotherapy dose levels was found to have a good fit to both the LKB and Niemierko models (figures 4 and 5).

Two other published studies have suggested a weak volume effect. However, one was based on a tumour bed boost using brachytherapy and other was based on summative patient data. The Alexander et al. study\textsuperscript{79} indicating a large volume effect has major limitations. To our knowledge, this is the largest dose-volume study for breast fibrosis using individual patient data. Parameter correlation leads to uncertainty in the
estimation of those parameter, independent of the size and diversity of the dataset\textsuperscript{84}. An effective method to decrease the uncertainty is fixing one or more of the model parameters. Hence the $\alpha/\beta$ was fixed as 3Gy in this study based on the previously published literature\textsuperscript{16}. There is no evidence to suggest the superiority of one model (LKB or Niemierko) over another\textsuperscript{85}. However, similar values of the estimated parameters from the two models strengthen the results of this study.

There are several possible reasons to explain the difficulty in demonstrating a volume effect for breast fibrosis, and can be considered limitations to the current study. Breast fibrosis may represent a focal effect, with the maximum radiotherapy dose as the most predictive factor of that focal effect. It is also possible that current scoring methods for breast fibrosis are not sensitive to the volume effect. Breast fibrosis is often graded as mild to severe based on the severity; however the scoring system does not take into account the extent of fibrosis e.g. a small discrete region of fibrosis and a widespread region of fibrosis are potentially scored alike. It has been suggested that NTCP parameters are influenced by the severity of the measured toxicity\textsuperscript{86}. For rectum, Rancati et al. estimated the best fit ‘$n$’ parameter was 0.23 for $\geq$ grade 2 rectal bleeding, which decreased to 0.06 when only severe rectal bleeding (grade 3) was considered\textsuperscript{86}. It is plausible that a volume effect for breast tissue may have been seen for mild fibrosis, but this endpoint was considered to be of less clinical significance and therefore not assessed.

Other toxicity endpoints like photographic assessment of breast shrinkage may also be more sensitive to the volume effect as it represents an affect across the whole organ effect, is more objective and scored independent of surgical changes. The current study only focused on breast fibrosis measured using photographic assessment, as patient reported scoring was not available for the majority of the patients included in the study.
Conclusions

Modelled NTCP parameters suggest that for moderate and severe fibrosis, the breast tissue behaves as a serial organ and that the maximum RT dose is the best predictor of complication. The derived model, predicts close to zero effect of volume of irradiated tissue on the risk of toxicity. Evidence of a volume effect reported in the literature, warrants further investigation (Chapter 2). Further work will use IMPORT high toxicity and dosimetry data to test model parameters using similar methodology to that reported here.
Chapter 4 – A multi-institutional investigation of image guided radiotherapy for breast cancer

This chapter addresses the main research objective of the study: the difference in spatial accuracy between IGRT and standard imaging. Differences in daily set-up errors measured using IGRT and standard imaging were assessed and overall accuracy was defined as the difference in population systematic set-up error obtained using IGRT compared to standard imaging. This chapter also addresses two secondary objectives, the decrease in safety margin provided by IGRT and the time required to perform IGRT and standard imaging verification. Differences between centres and imaging modalities are also investigated.

Introduction

The breast radiotherapy process can be divided into three steps:

1. Simulation – A planning computer tomography (CT) scan is performed while the patient is in the position that will be used during radiotherapy treatment delivery. The patient is immobilised using breast board/vacuum bag and pre-defined tattoo marks are placed on the skin surface.

2. Radiotherapy planning – The planning CT scan is used to identify and contour the target volumes (tumour bed and whole breast) and organs at risk (lung, heart, contralateral breast). Radiotherapy beams are designed to optimally cover the target volumes and spare the organs at risk.
3. Treatment verification and delivery – Radiotherapy treatment is delivered over a series of sessions (called fractions), to allow preferential DNA repair to take place in normal cells but not the cancerous cells. Before each radiotherapy fraction, the patient position is reproduced using laser light beams; pre-defined tattoo marks on the patient’s skin are aligned with the lasers. Verification images are taken while the patient is on the treatment couch to confirm correct positioning and treatment is delivered.

A difference in patient position between the planning CT scan and treatment session can lead to geographical miss of the target, potentially increasing the risk of cancer recurrence. Due to the uncertainty in patient and target position with each fraction (subsequently called as set-up errors), a planning target volume (PTV) margin is routinely added around the target volume. This PTV margin not only accounts for the daily interfraction and intrafraction motion (figure 6), but also beam penumbra and other geometrical uncertainties associated with the radiotherapy equipment. Interfraction motion includes differences in patients positioning between radiotherapy fractions. Intrafraction motion includes movement that occurs during each radiotherapy fraction, for example respiratory motion and tissue deformation.
Pre-defined skin tattoo marks and laser beams are currently used to position patients for breast radiotherapy. Though simple to use, the set-up errors using this technique are large. Studies have reported that positional error using surface markers could range from 1-30 mm\textsuperscript{87-89}. These large positional errors mean that a relatively large PTV margins has to be used (commonly 10 mm).

Due to the addition of a large PTV margin, a considerable volume of the healthy surrounding tissue is unnecessarily irradiated to treatment dose, increasing the risk of radiation related adverse events. It also limits our ability to safely escalate the radiation dose to the target. IGRT technique can be used to reduce both interfractional and intrafractional errors and potentially reduce the PTV margins.

Figure 6: Set-up errors and PTV margin in radiotherapy
Megavoltage portal imaging method (2D-MV) is the current standard imaging verification technique for breast radiotherapy. The breast radiation treatment is usually carried out using lateral and medial tangential beams and these high energy (megavoltage) treatment beams are used to generate portal images (PI). The position of the ribs and lung on PI are compared to a digitally reconstructed radiograph (DRR) generated from planning CT images to identify the day to day variation in patient positioning (figure 7).

![Image](image_url)

**Figure 7:** Standard verification technique compares digitally reconstructed radiograph and mega-voltage portal image

Parameters including central lung distance (CLD), defined as distance between the posterior field edge and the interior chest wall at the central axis and cranio-caudal distance (CCD), defined as distance between skin and the caudal beam edge are compared between the PI and DRR, to calculate positional errors in both the transverse and longitudinal direction (figure 8). If the positional errors exceed pre-
defined limits (commonly set as 5mm), the patient is re-positioned before the next treatment is delivered.

![Central lung distance (CLD) and Cranio-caudal distance (CCD)](image)

Figure 8: Measurement of central lung distance (CLD) in black and cranio-caudal distance (CCD) in blue on a DRR

Though simple and effective, the portal images (PI) provide information about patients position based on bony anatomy, and not the breast tissue. In addition, the tumour bed (area at highest risk of cancer recurrence) cannot be directly visualised on the PI. The chest wall is used as a surrogate for the breast and the tumour bed.

In recent years, studies have shown that bony anatomy (chest wall) is a poor surrogate for both the tumour bed and the whole breast. Hasan et al.\textsuperscript{9} study of 27 patients treated with accelerated partial breast irradiation indicated that (a) whole breast can move
independent of bony anatomy and (b) tumour bed can also move independent of the whole breast.

Due to our inability to directly visualise the tumour bed for positional verification and correct for intrafraction motion, a PTV margin of 10 mm is commonly added to the tumour bed, to generate a planning target volume (PTV) for photon tumour bed boost\textsuperscript{90}.

Due to the additional PTV margin around the tumour bed, a large volume of normal breast tissue is treated to a high radiation dose. This can potentially increase the risk of late breast tissue toxicity. Apart from ipsilateral breast, an increase in PTV margin will also increase the radiation dose to contralateral breast, heart and ipsilateral lung. If we could safely reduce set-up errors, PTV margins around the tumour bed can also be safely reduced. This is desirable to reduce the risk of late breast and other normal tissue toxicity.

The British Association of Surgical Oncology (BASO) have recommended that all patients undergoing breast conserving surgery should have surgical clips on the wall of the tumour bed\textsuperscript{91}. Clips are currently used as fiducial markers, for the accurate localisation the tumour bed\textsuperscript{92}. In addition, clips have been shown to be a better surrogate for the tumour bed compared to bony anatomy and used for IGRT\textsuperscript{6,9,93} In this report we call this approach clip-based image guided radiotherapy or clip-based IGRT.

The use of surgical clips as a surrogate for the tumour bed was evaluated in 28 patients by Weed et al.\textsuperscript{6}. Each patient underwent two planning CT scans on separate days. The tumour bed and clips were identified as regions of interest (ROI). The scans were then fused based on bony anatomy and the displacement of the tumour bed was
compared to the displacement of the clips over time. The study found that the
displacement of clips tracked the displacement of the excision cavity during radiation
therapy. An average displacement error of 3mm was seen between the two ROIs,
which were attributed to the finite thickness of the CT slices and use of limited
number of clips. Hasan et al.\textsuperscript{9} also demonstrated that surgical clips are a better
surrogate for tumour bed compared to bony anatomy and breast surface. Twenty
seven patients underwent two CT scans in treatment position, one initial planning CT
scan and a second scan at an average of 27 days after the first scan. The centre of
mass (COM) of the lumpectomy cavity was determined on both CT scans for each
patient. Localisation of the tumour bed was performed using CT registration of the
following: bony anatomy, COM of surgical clips embedded in the excision cavity and
breast surface. The distance between COMs using the three registration process were
compared ($\Delta$COM\textsubscript{bony anatomy}, $\Delta$COM\textsubscript{clips} and $\Delta$COM\textsubscript{breast surface}). It was observed that
localisation of the tumour bed using surgical clips is most accurate compared to
localisation using bony anatomy and breast surface. Topolnjak et al.\textsuperscript{93} compared the
residual error (surrogate error) between excision cavity and surgical clips placed in
the excision cavity to determine if surgical clips are a good surrogate for the tumour
bed and quantify the stability of the clips position. Twenty one breast cancer patients
were treated with 28 fractions and cone beam CT (CBCT) scans were regularly
acquired for set-up correction protocol. The CBCT scans were registered to the
planning CT scan using grey value registration of the excision cavity and chamfer
matching of the clips. The study showed that surgical clips are a good surrogate for
excision cavity with small residual errors of 0.7-1.3 mm.

It is now known that surgical clips are a better surrogate of the surgical cavity (tumour
bed), as compared to bony anatomy. A number of studies have evaluated the
feasibility of using these surgical clips for image guided breast radiotherapy\textsuperscript{94-96}. 
The IMPORT HIGH trial group used gold fiducial markers (small metallic seeds that can be sutured on to the cavity wall) as a tumour bed surrogate for IGRT to estimate safe PTV margins around the breast tumour bed\textsuperscript{94}. Treatment verification and daily on-line correction were performed on 42 patients with 2D-MV (high energy) portal image or kV (low energy) planar image or cone beam CT. The study concluded that using extended no action level (e-NAL) or daily on-line correction strategy (discussed later), the tumour bed PTV margins can be safely reduced to 5mm. Leonard et al.\textsuperscript{95} also demonstrated the feasibility of gold seed fiducial markers for marker-based IGRT using orthogonal and lateral MV portal films in 20 patients.

2D-MV portal imaging verification method using bony anatomy is relatively simple and easy to use whereas additional verification time and resources are required for marker-based IGRT technique. The benefit of clip-based IGRT over portal imaging needs to be quantified by comparing the PTV margins and verification time for both techniques. In addition, most of the feasibility studies of fiducial marker-based IGRT were based on small number of patients, using gold seeds as fiducial markers. The use of gold seeds as fiducial markers is quite expensive (~£200/patient), considering that breast radiotherapy constitutes a large part of radiotherapy department work load. Titanium clips can be used as an alternative fiducial marker (~£1/patient), though due to their low density, they cannot be visualised on 2D-MV portal image. Several different imaging modalities can be used for titanium surgical clip-based IGRT: kV planar images, kV Cone beam CT and Mega Voltage CT (TomoTherapy). It is currently unclear if the PTV margin will depend on the type of IGRT imaging modality used.
Materials and Methods

The sample size calculation was carried out by the ICR-CTSU. The study statistician (JH) was responsible for overseeing all statistical analyses. In this study, the primary research objective was the accuracy of clip-based IGRT compared to standard imaging. This study proposed to use accuracy to determine two secondary research objectives of this study, volume of normal tissue irradiated and the probability of adverse effects (fibrosis) using the following steps:

- Accuracy is calculated using the mean set-up error for each patient (patient systematic error). The mean set-up error for each patient is the mean of the 15 set-up errors measured at each fraction. The accuracy of standard imaging compared to clip-based IGRT will determine the additional safety margin.

- The technique-specific safety margin is calculated from the distribution of the patients’ overall set-up errors. The size of the margin is approximately two and a half times the standard deviation of the mean set-up errors for all patients (population systematic error).\(^{97}\)

- The volume of normal tissue irradiated when standard imaging is used is then calculated by simply adding this margin to the treatment volume used for IMPORT high (which uses clip-based IGRT).

- The probability of adverse effects (fibrosis) for clip-based IGRT and standard imaging will be determined from the volumes of tissue irradiated with each of these methods and using the relationship between incidence of fibrosis and volume determined using radiobiological modelling.

This study was designed to generate two sets of data: the mean set-up error from n patients receiving curative breast radiotherapy using clip-based IGRT and using standard imaging.
A difference in the standard deviation (or variance) of these data sets will result in a difference in safety margin and the volumes of normal tissue irradiated. Thus, the sample size calculation was based on finding a significant difference in the variance of these two data sets. To perform the sample size calculation the standard deviation of set-up errors for IGRT and standard imaging the correlation between the data-sets was estimated using evidence from the literature:

- **Standard deviation:** From a small study of 20 patients, Topolnjak et al.\(^98\) found that, if standard imaging were used daily, the standard deviation of the set-up errors was between 2.7mm and 3.8mm. In a similar study of 10 patients by Kim et al.\(^96\) the standard deviation of set-up errors was measured to be between 0.9 mm and 1.4 mm when daily clip-based IGRT is used. Kim et al. estimate that this range of values rises from 2.2mm to 2.6mm when other factors such as deformation of the breast are taken into account. To calculate the sample size, based on these studies the aim was to detect differences in standard deviations corresponding to a decrease from 3mm for standard imaging to 2mm for clip-based IGRT.

- **Correlation:** Because no similar studies have previously been performed directly comparing clip-based IGRT and standard imaging the correlation between the two data sets is unknown. Work by Penninkhof et al.\(^99\) shows that set-up errors measured in the same patient using two imaging techniques to image bony anatomy are highly correlated (~0.85). It was expected that the correlation between set-up errors for clip-based IGRT and standard imaging is high (> 0.5) as they will be measured in the same patient but not as high as for two techniques measuring bony anatomy (i.e. < 0.85).

The sample size required was determined for high correlation, 0.7 and very low correlation, 0.1. Using computer simulations based on Fisher's test\(^100\), the number of patients required to detect a 1mm difference in the standard deviations from 2mm to
3mm assuming correlation=0.7, power=80%, alpha=0.05 was determined. Using the same analysis but assuming very low correlation (correlation = 0.1), with 250 patients it was found that it was possible to detect the same difference (2mm v. 3mm) in standard deviations (power=80% and alpha=0.05).

This sample size calculation was based upon estimates from studies using small numbers of patients and a definitive value for the correlation between set-up errors measured using the two techniques was not available. We based the study on the requirement for a larger cohort of patients, 250 which allows for smaller correlations. It was proposed that an Independent Data Monitoring Committee confidentially review the data after the first 100 patients, and advise on the final sample size.

All patients participating in the national Phase 3 IMPORT HIGH trial have surgical clips inserted into the walls of the TB and are receiving clip based IGRT as routine. The daily verification image data for IGRT were used to calculate the set up error with clip-based technique. These imaging data were also used to calculate the set up error if bony anatomy was used for verification (ignoring the information from the clips). All patients had previously consented for their imaging data to be used for research purposes. As the imaging data was retrospectively analysed, it had no direct impact on the study population.

Imaging data was collected from five different centres participating in the IMPORT HIGH trial: Addenbrookes Hospital, Royal Marsden Hospital (RMH), Ipswich Hospital, Cheltenham Hospital and Clatterbridge Hospital. The 5 centres were chosen because they were early implementers and high recruiters of the IMPORT high trial. Combined, these 5 centres used all 3 imaging techniques used in the IMPORT trial which represented current national practice. Ipswich, Cheltenham and Clatterbridge
centres used KV planar images (2D-kV) and daily on-line image verification protocol, RMH used Cone beam CT (kV-CBCT) with an e-NAL verification protocol, and Addenbrookes used Mega Voltage CT (MV-CT) (TomoTherapy) with daily on-line image verification protocol for treatment verification and positional correction (figure 9). The details of different image verification protocols (IVPs) are discussed later.
<table>
<thead>
<tr>
<th>BONY ANATOMY VERIFICATION</th>
<th>CLIP BASED VERIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>kV-Cone Beam CT</td>
<td>MVCT (TomoTherapy)</td>
</tr>
</tbody>
</table>

| 2D-kV Planar               |                         |
Bony anatomy (BA) based set-up errors were measured using automatic bony anatomy registration software for kV-CBCT (Synergy, Elekta Ltd, Crawley, UK) and manually for MV-CT and 2D-kV. Then clip-based IGRT set-up errors were measured by manually adjusting alignment of images from their BA-matched position. BA set-up errors ($S_{BA}$) and clip set-up errors ($S_{clip}$) in the lateral (LR), superior-inferior (SI) and anterior-posterior (AP) directions were recorded. The time taken to perform bony anatomy ($T_{BA}$) and clips ($T_{clips}$) based image assessment (a secondary objective of this study) were also recorded. Only images with sufficient information on bony anatomy were used. Intra- and inter-observer errors were assessed using nine images from three patients and a minimum of two observers.

For RMH patients, e-NAL corrections applied for actual treatment were removed from the measured set-up errors so that the effects of various IVPs could be studied.
Radiotherapy positional errors are classified into systematic errors and random errors. Systematic errors occur if the mean irradiation geometry in a fractionated treatment differs from the geometry of the treatment plan. Fraction to fraction variation around the mean deviation is called as random error. Systematic error can shift the cumulative dose distribution relative to the target and contributes more towards the PTV margin as compared to random error which blurs the dose distribution.

The PTV margin calculation is based on the population systematic error ($\Sigma$) and the population random error ($\sigma$). $^{97}$

$$\text{PTV margin} = 2.5 \Sigma + 0.7 \sigma \quad (8)$$

For a given population, systematic error ($\Sigma$) is the mean of the standard deviation (SD) of all patients’ mean set-up errors and random error ($\sigma$) is root mean square of the all patients’ SD of daily errors.

For this project, verification images of the study population were used to measure the distance of bony anatomy (BA) and clips (IGRT) from a reference position to determine bony set-up error ($S_{BA}$) and clip set-up error ($S_{clips}$). The additional PTV margin required if standard bony anatomy verification technique is used over clip based IGRT was calculated using the difference in the distance between bony anatomy and clip position. For each patient Delta error, $S_{\text{DIFF}} = S_{BA} - S_{clips}$ was generated. The mean and SD of the $S_{\text{DIFF}}, S_{BA}$ and $S_{clips}$ for the study population was used to generate the systematic error (delta $\Sigma$) and random error (delta $\sigma$) for the margin formula.
$S_{\text{DIFF}}$ for the first 112 patients was collected and analysed to calculate the required sample size. The calculations were based on the 95% confidence interval (CI) that will give the required precision of 0.05 cm on the PTV margin estimate.

Individual patient and population mean error (M), systematic ($\sum$) and random ($\sigma$) errors were calculated for bony anatomy (standard imaging) and clips (clip-based IGRT) error data. Bland-Altman analysis, least squares linear regression and calculation of the Coefficient of determination ($R^2$) between $S_{\text{BA}}$ and $S_{\text{clips}}$ were also performed.

As the tumour bed has previous received background radiation during whole breast irradiation, the margin formula was modified by reducing the contribution of the $\sigma$ error.

$$\text{PTV margin} = 2.5 \sum + 0.3 \sigma \tag{9}$$

The PTV margins required for safe treatment may not only depend on the method of verification used (standard imaging using bony anatomy or clip-based IGRT), but also on the type of verification protocol used. Four different verification protocols were investigated to calculate the PTV margins:

1. No correction protocol – No imaging is undertaken and patient is positioned using laser based set-up.
2. No action level (NAL) – The systematic error is calculated after 3 treatment fractions and systematic set-up error is corrected for all subsequent fractions, regardless of the magnitude of the error$^{97}$.
3. Extended-no action level (e-NAL) - The first stage of the protocol follows NAL strategy with additional once weekly verification and correction$^{101}$. 
4. Daily correction protocol - Patient position is verified daily against the planning CT scan and corrected (if required).

\[ S_{BA} \] and \[ S_{clips} \] data was not available for all 15 treatment fractions. In order to evaluate the effect of IVPs for a 15 fraction treatment, a simulation of set-up errors was performed. For each patient, if \( N \) was the total number of images analysed, in cases with \( N < 15 \), a normal distribution with mean and standard deviation equal to the patient’s real set-up data was sampled \( N-15 \) times. Combined real and simulated patient set-up data were used to simulate the IVPs using Matlab (Mathworks, Natick, MA, USA). The smallest number of images available per patient was \( N=5 \). To test if 5 images was adequate to describe a patient’s set-up data, the mean and standard deviations of set-up errors of 28 patients with \( N = 15 \) were determined, for all 15 set-up errors and 5 set-up errors (fractions 1, 2, 3, 7 and 11). When using 5 images compared to 15 images, mean difference in patient’s mean and standard deviation of set-up errors was 0.006 cm and 0.013 cm, respectively.

Having obtained 15 measures of set-up error for each patient, these data were used to simulate the effect of different IVPs on set-up errors, and hence PTV margins. Simulated IVPs included: on-line BA (OL\(_{BA}\)), on-line clip (OL\(_{clips}\)), e-NAL BA (e-NAL\(_{BA}\)) and e-NAL clip (eNAL\(_{clips}\)). Post IVP simulation, any remaining systematic and random errors were calculated for the patient population. Set-up error simulation and error calculation was repeated 1000 times for each IVP. Error values from repeat simulations were averaged to give more precise results giving less than 0.1% uncertainty (1 SD) from random sampling.

This study used clips as a surrogate for the tumour bed. Surrogate systematic and random errors of 1.1mm were added in quadrature to the set-up errors\(^{96}\), to account for the uncertainty introduced by the localisation of clips rather than the tumour bed.
As discussed before, 2D–MV portal imaging is the current standard treatment verification method for breast radiotherapy. No portal imaging data was collected as part of the study. The 2D-MV set-up error data ($S_{2D-MV}$) were derived from $S_{BA}$ using the method previously proposed by Topolnjak et al.\textsuperscript{98}

\[
S_{2D-MV} = \beta S_{BA} + \alpha + \text{rand} \times \eta 
\]  \hspace{1cm} (11)

Where parameters $\beta$ (slope) and $\alpha$ (intercept) were determined from regression analysis, $\eta$ is the standard deviation of the differences between 2D-MV and kV-CBCT set-up errors measured by Topolnjak et al.\textsuperscript{98} and \textit{rand} is a random number sampled from a normal distribution.

Tangential portal imaging will not provide a measure of set-up error in all three ordinal directions. It was assumed that 3D set-up errors are available from 2D-MV imaging, which is possible if anterior and lateral portal images are used\textsuperscript{102}.

In this study, for the LR and AP directions, $\beta = 0.82$, $\alpha = 0.66$ mm and $\eta = 0.18$mm were used. For the SI direction $\beta = 0.43$, $\alpha = -0.28$ mm and $\eta = 0.32$mm were used\textsuperscript{98}. Bony anatomy based IVP simulation was repeated using $S_{2D-MV}$.

\textit{Statistical Analyses:} Data were tested for heterogeneity between imaging techniques and radiotherapy centres. All data were tested for normality using the Shapiro-Wilk test. Shapiro-Wilk tests indicated some data were not normally distributed. Overall patient mean $S_{DIFF}$ was tested for significant difference from zero using a one-sample Student’s t-test. For $S_{DIFF}$, the difference in absolute $S_{DIFF}$ between centres and imaging
techniques was tested using Kruskal-Wallis followed by sensitivity analysis. Differences in overall mean patient systematic error $M$, population systematic error $\Sigma$ and population random error $\sigma$ between modalities, centres and imaging protocols were tested. For $M$, difference from zero was calculated using one-sample Student’s t-test (all data) or Wilcoxon signed-rank test (per centre). For $\Sigma$, Non-Parametric Levene’s test (NPLT) was used to test difference in the variance of mean patient errors. For $\sigma$, Kruskal-Wallis test was used to test for differences in patient’s random errors. For all tests, data were considered to be significantly different if $p < 0.05$. Sensitivity analysis was performed by removing data from one centre at a time and repeating tests using Holm-Bonferroni correction. The time required for positional verification between IGRT and standard imaging was compared using Wilcoxon Signed Ranks Test.

Results

Study sample size calculation: Using computer simulations based on Fisher's test$^{100}$, with 128 patients, it was possible to detect a 1mm difference in the standard deviations from 2mm to 3mm assuming correlation=0.7, power=80%, alpha=0.05; using the same analysis but assuming very low correlation (correlation = 0.1), with 250 patients it was possible to detect the same difference (2mm v.s. 3mm) in standard deviations (power=80% and alpha=0.05).

Bony set-up error ($S_{BA}$) and clip set-up error ($S_{clips}$) of 112 patients from three different centres were initially collected for the review of the sample size.
The additional margin required if standard imaging verification is used instead of clip-based IGRT was calculated from $S_{\text{DIFF}}$ in left-right (LR), superior-inferior (SI) and anterior-posterior (AP) directions. These results are summarised for the three centres in table 5.

<table>
<thead>
<tr>
<th>Centre</th>
<th>Mean (cm)</th>
<th>Variance (cm)</th>
<th>Margin (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LR</td>
<td>SI</td>
<td>AP</td>
</tr>
<tr>
<td>ALL</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.03</td>
</tr>
<tr>
<td>CCC</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>ADD</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>RMH</td>
<td>-0.03</td>
<td>-0.06</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

The variances of $S_{\text{DIFF}}$ in the 3 directions for all centres were:

Table 5: Mean and Variance of $S_{\text{DIFF}}$ for all 112 patients and individual centres. (Highlighted data were used in the sample size calculation)

$$S_{\text{DIFF}}^2_{\text{LR}} = 0.039, \ S_{\text{DIFF}}^2_{\text{SI}} = 0.043 \ \text{and} \ S_{\text{DIFF}}^2_{\text{AP}} = 0.044 \quad (12)$$

Taking the largest variance $S_{\text{DIFF}}^2_{\text{AP}} = 0.044$, the 95% confidence interval (CI) was calculated from the chi-square distribution table using the following:

$$\text{Lower limit} = (n-1) \ S_{\text{DIFF}}^2 / \chi^2_L \quad \text{to} \quad \text{Upper limit} = (n-1) \ S_{\text{DIFF}}^2 / \chi^2_U \quad (13)$$

Where $\chi^2_U$ = upper 2.5% point of $\chi^2$ distribution for 111 degrees of freedom = 83.735

And $\chi^2_L$ = lower 2.5% point of $\chi^2$ distribution for 111 degrees of freedom = 142.049
Hence 95% CI for variance is 0.344 (111 x 0.044/142.049) to 0.0583 (111 x 0.044/83.735) and the 95% CI for standard deviation = 0.1855 (√0.344) to 0.2415 (√0.0583). Based on this calculation, there was 95% confidence that the margin lies in the region of 0.4637cm to 0.6037cm (SD x 2.5). This is an overall width of 0.140cm.

As the overall precision required was 0.05 (overall width of 0.1cm), the above formula was applied for different sample sizes and produced an estimate that a sample size of 200 patients would be required for this study.

The bony anatomy set-up error (S_{BA}) and clip set-up error (S_{clips}) for 218 patients from five different centres was collected. Each centre uses a different imaging technique as summarised below in table 6.
<table>
<thead>
<tr>
<th>Hospital</th>
<th>Centre</th>
<th>Imaging modality</th>
<th>No of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Royal Marsden Hospital (RMH)</td>
<td>A</td>
<td>Cone Beam CT (CBCT)</td>
<td>79</td>
</tr>
<tr>
<td>Addenbrookes (ADD)</td>
<td>B</td>
<td>MV-CT (TomoTherapy)</td>
<td>40</td>
</tr>
<tr>
<td>Clatterbridge Cancer centre (CCC)</td>
<td>C</td>
<td>2D-kV Planar images</td>
<td>39</td>
</tr>
<tr>
<td>Cheltenham (CHE)</td>
<td>D</td>
<td>2D-kV Planar images</td>
<td>30</td>
</tr>
<tr>
<td>Ipswich (IPS)</td>
<td>E</td>
<td>2D-kV Planar images</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 6: Summary of imaging technique and patient accrual for each centre

The intraobserver and interobserver errors were less than 1.4mm for all three imaging modalities. No significant difference was seen in inter-observer errors between the five centres (ANOVA, p=0.34). The overall mean error was found to be significantly different from zero for Addenbrookes (centre B) and Ipswich (centre E). At Addenbrookes, this is due to couch sag associated with Tomotherapy\(^\text{103}\). The Ipswich centre is investigating the cause of their non-zero mean error.

The systematic errors (\(\Sigma\)) and random errors (\(\sigma\)) using bony anatomy verification and titanium clips based verification were mostly 2-4 mm across all centres. Individual centre and overall errors are summarised in table 7. Individual patient systematic error using bony anatomy and surgical clips are compared in figure 10 using Bland-Altman analysis. The bias and limits of agreement (±1.96SD) between BA and clip
systematic errors were 0.0±0.21 cm, 0.0±0.26 cm and 0.1±0.22 cm in the LR, SI and AP directions, respectively. Using linear regression analysis to compare bony anatomy and surgical clips systematic errors, the coefficient of determination ($R^2$) were 0.57, 0.42 and 0.82 in the LR, SI and AP directions respectively, suggesting that bony anatomy based verification underestimates the patient systematic error by up to 23% compared to clips based verification. The difference in set-up errors using bony anatomy and clips ($\Sigma_{DIFF}$ and $\sigma_{DIFF}$) for individual centres and all patients are summarised in table 8.

The time required to perform bony anatomy ($T_{BA}$) and clips ($T_{clips}$) based image assessment are also summarised in table 8. If all centres and all imaging techniques were considered there was no significant difference between ($T_{BA}$) and clips ($T_{clips}$) (Wilcoxon Signed-Rank, $p = 0.36$). The ranges in times were 8 s to 240 s for clip-based IGRT image assessment and 8 s to 178 s for bony anatomy image assessment.

The time required for image assessment varied with the type of imaging modality used. For all centres using 2D-kV modality, $T_{BA}$ was greater than $T_{clips}$ (Wilcoxon Signed-Rank Test, $p<0.001$). In contrast, centre “A” using kV-CBCT found median time $T_{clips} > T_{BA}$, 92 seconds versus 26 seconds respectively (Wilcoxon Signed-Rank Test, $p< 0.001$). No significant time difference was found for MVCT (Wilcoxon Signed-Rank Test, $p = 0.92$). The time to perform both clip and bony anatomy based image assessment varied significantly between the 3 centres using 2DkV imaging (Kruskal-Wallis, $p<0.001$). Using sensitivity analysis, it was found that centre C required significantly shorter time to match both clips and bones, compared to centres D and E.

There was a small but statistically significant difference in the difference between standard imaging and clip-based IGRT set-up errors (delta errors) between centres
(Kruskal-Wallis, p<0.05). The smallest delta error, $S_{\text{DIFF}}$, was seen for centre B using MVCT and the largest delta error was seen for centre C using 2D-kV (Table 8). Using the Kruskal-Wallis test, a significant difference between delta errors was seen between different centres in the LR and SI direction (p<0.05). No significant difference between in delta error was seen among centres using 2D-kV imaging modality. Non-Parametric Levene’s test and Bartlett’s box test also indicated non-homogeneity of variance among centres (table 9). After excluding data of patients with MV-CT imaging, the variance of delta error were similar between centres.
<table>
<thead>
<tr>
<th>Centre</th>
<th>No. Patients</th>
<th>Total Number of images</th>
<th>Bony Anatomy Random Error $\sigma_{BA}$ (cm)</th>
<th>Bony Anatomy Systematic Error $\Sigma_{BA}$ (cm)</th>
<th>Clips Random Error $\sigma_{clip}$ (cm)</th>
<th>Clips Systematic Error $\Sigma_{clip}$ (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>218</td>
<td>1574</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>A (kVCBCT)</td>
<td>79</td>
<td>504</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>B (MVCT)</td>
<td>40</td>
<td>200</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>C (2D-kV)</td>
<td>39</td>
<td>510</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>D (2D-kV)</td>
<td>30</td>
<td>180</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>E (2D-kV)</td>
<td>30</td>
<td>180</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

*Table 7: Systematic and random errors using bony anatomy and clip verification for each centre and for all centres combined*
Figure 10: Bland-Altman plots of average of bony anatomy and clips mean set-up error versus difference between mean clip set-up error mean bony anatomy set-up error in the LR, SI and AP directions. Solid line indicates mean difference between mean clip and mean bony anatomy set-up errors (the bias) and the dotted lines represent the limits of agreement (±1.96SD).
<table>
<thead>
<tr>
<th>Centre</th>
<th>Delta Error ($S_{DIFF}$)</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Absolute Delta (range)</td>
<td>Median (range)</td>
</tr>
<tr>
<td></td>
<td>(cm)</td>
<td>(seconds)</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>SI</td>
</tr>
<tr>
<td>ALL</td>
<td>0.20 (0,1.7)</td>
<td>0.26 (0,3.2)</td>
</tr>
<tr>
<td>A (kV-CBCT)</td>
<td>0.19 (0,0.7)</td>
<td>0.24 (0,3.2)</td>
</tr>
<tr>
<td>B (MVCT)</td>
<td>0.14 (0,0.7)</td>
<td>0.12 (0,1.2)</td>
</tr>
<tr>
<td>C (2DkV)</td>
<td>0.23 (0,1.7)</td>
<td>0.29 (0,2.4)</td>
</tr>
<tr>
<td>D (2DkV)</td>
<td>0.21 (0,1.3)</td>
<td>0.32 (0,1.3)</td>
</tr>
<tr>
<td>E (2DkV)</td>
<td>0.20 (0,1.5)</td>
<td>0.31 (0,1.4)</td>
</tr>
</tbody>
</table>

Table 8: Delta errors (difference between bony anatomy and clips, $S_{DIFF}$) in the LR, SI and AP directions and the magnitude of their 3D vector. Time required for image matching with both techniques has also been summarised.
The PTV margins should be estimated using a sample group which is representative of the whole population. In view of significant heterogeneity using MV-CT, set-up data of centre B (n=40 patients) were not included in the simulations of PTV margin estimation. Based on the set-up data of 178 patients, the overall width of 95% confidence interval on the PTV margins is ~0.107 cm giving a precision of ± 0.05 cm.

The mean (M), systematic (∑) and random (σ) residual error were calculated using the following image verification protocols (IVP):

a. No correction protocol - NO IMAGING
b. Extended- no action level using bony anatomy- e-NAL_{BA}
c. Extended- no action level using clips - e-NAL_{clips}
d. Daily correction protocol (On line correction) using bony anatomy - OL_{BA}
e. Daily correction protocol (On line correction) using clips – OL_{clips}

The results are summarised in table 10. When using 5 images compared to 15 images, mean difference in patient mean and standard deviation of set-up errors was 0.006 cm and 0.013 cm, respectively.
In all cases, the variation (1 SD) in residual errors due to random sampling was less than 0.01 mm. Residual systematic and random errors were smaller for clip-based verification as compared to BA verification, irrespective of the IVP method.
<table>
<thead>
<tr>
<th>IVP</th>
<th>M(cm)</th>
<th>σ(cm)</th>
<th>Σ(cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LR</td>
<td>SI</td>
<td>AP</td>
</tr>
<tr>
<td>No Imaging</td>
<td>-0.06</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>OL_{clips} (2D-kV or kV-CBCT)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>eNAL_{clip} (2D-kV or kV-CBCT)</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>OL_{BA} (2D-kV or kV-CBCT)</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.03</td>
</tr>
<tr>
<td>eNAL_{BA} (2D-kV or kV-CBCT)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>OL_{BA} (2D-MV)</td>
<td>-0.64</td>
<td>0.27</td>
<td>-0.60</td>
</tr>
<tr>
<td>eNAL_{BA} (2D-MV)</td>
<td>-0.45</td>
<td>0.21</td>
<td>-0.43</td>
</tr>
</tbody>
</table>

*Table 10: Mean, Random and Systematic errors using different image verification protocols*
The study showed that for kV based imaging modalities, on-line IVP produces smaller errors than e-NAL for both clips and bones. Using, on-line BA based verification, systematic errors were larger than those for off-line clip based verification (e-NAL_{clips}), by up to 0.11 cm (Non-Parametric Levene’s Test (p-values <0.001). Off-line BA verification (e-NAL_{BA}) increased systematic error further by ~ 0.05cm (Non-Parametric Levene’s Test p-values <0.001).

2D-MV based image verification increased systematic error for both on-line and off-line protocols by an average of ~0.3cm (Non-Parametric Levene’s Test, p<0.001). For off-line imaging of BA (e-NAL_{BA}), the difference between kV and 2D-MV was significant in the SI direction only. For 2D-MV imaging of BA, there were no significant difference in systematic errors between on-line and off-line IVPs (p=0.12).

The overall mean error (M) for 2D-MV IVPs was non-zero (table 10). This is likely due to the use of simulation technique to generate 2D-MV set-up errors and reflects the relationship between kV-CBCT and 2D-MV set-up errors.

The estimated PTV margin using different IVPs are given in figure 11. Based on this study, a tumour bed boost PTV margin of 1 cm is required if no imaging modality is used. If standard bony verification technique is used (2D-MV), a PTV margin of 0.8cm is required. This can be reduced to 0.6 to 0.7 cm if 2D-kV/ kV-CBCT based bony anatomy verification is used. The use of clip-based verification (clip based IGRT) allows a 0.5cm boost PTV margin for both on-line and e-NAL protocols.
Figure 11. Tumour bed PTV margins required for the different imaging verification protocols considered in this study. Margins are given for the LR, SI and AP directions.

Discussion

This large multicentre multimodality study has compared the set-up errors of bony anatomy and clip based verification. It demonstrates that PTV boost margin of 5 mm are adequate if clip-based IGRT (2D-kV and kV-CBCT) is used both for on-line and off-line IVP. However, if standard portal imaging (2D-MV) is used, an increase in PTV margin of ~3mm is necessary. The major strengths of the study include its large patient cohort, use of different imaging modalities and direct comparison of bony anatomy verification against clip based verification.

In this study, the measured population random and systematic errors (for the TB) were within 2 to 4 mm, suggesting that a PTV margin of 10 mm is required if no IVP is used. Similar results have been reported by Topolnjak at el. who reported kV-CBCT measured systematic errors of 0.31, 0.38 and 0.25 cm in the LR, SI and AP direction respectively based on 20 patients. This current study found that the BA based verification underestimate patients’ systematic error as compared to clip-based IGRT. Other authors have reported differences between set-up errors measured bony and clip using small patients’ cohorts. Gierga et al. used 2D-kV in 12 patients and reported a median 3D delta error of 0.54 cm,
upper and lower quartile values were 0.75 cm and 0.41 cm respectively. Fatunase et al.\textsuperscript{105} reported a mean 3D delta error of 0.6 cm using kV-CBCT in 10 patients. In the current study, 1379 images were analysed and the mean, median, and upper and lower quartile 3D vector difference between bones and clips (delta) were 0.35, 0.18, 0.06 cm and 0.41 cm respectively (table 8). This implies that the PTV margins of 5mm (as used in the IMPORT HIGH study) may be insufficient if BA is used as a surrogate for the tumour bed and a larger PTV margin of ~8mm is required if 2D-MV based bony anatomy verification (online and offline protocols) is used.

Similar results have been reported by Penninkhof and colleagues\textsuperscript{99}. Two orthogonal planar kV images and one 2D-MV portal image were acquired for 80 patients throughout their radiotherapy. Surgical clips based registration was performed on all kV images and set-up errors (systematic and random) were estimated for no correction protocol, NAL protocol and e-NAL protocol. The 2D-MV portal images were independently registered with the DRR using lung contour and caudal side of the external breast contour for estimating 2D-MV set-up errors.

Time required for clip-based IGRT image assessment compared to standard imaging image assessment was imaging technique dependent. Clip-based IGRT verification was quicker than bony anatomy verification when using the 2DkV imaging. For kV-CBCT, bony anatomy verification was quickest, this is most likely because automatic bone registration was used. Clip-based IGRT verification using MVCT took the greatest amount of time (mean >2 minutes). This may be due to poor visualisation of clips on MVCT images.

Centres’ C, D and E had significantly different times for both clip-based IGRT and standard imaging. The differences may be explained by differences in observers; at centre C images were matched by a senior radiographer (AB), and at centre D by a physicist (EH) and a senior radiographer (RP) and at centre E by EH only.

Overall, differences in median times between bony anatomy and clip verification for each modality were small, the greatest difference being 76 seconds, for kV-CBCT. IMPORT image data was obtained retrospectively and therefore no times for image data acquisition were available. This was a limitation of our study, however, bony anatomy (T\textsubscript{BA}) and clips (T
clips) based verification data are useful for any future cost-benefit analysis, which is beyond the scope of our study.

Conclusion
The work described in this chapter addressed the main objective of this programme of work, to compare the spatial accuracy of breast radiotherapy using clip-based IGRT and standard imaging during curative radiotherapy for early breast cancer. The study concluded that accuracy of clip-based IGRT was greater than standard imaging for breast boost to the tumour bed. The use of three common imaging protocols (on-line, NAL and e-NAL correction protocols) with clip-based IGRT improved accuracy by between 2 to 4mm, compared to standard imaging. Using no imaging protocol, the systematic set-up errors for tumour bed were 0.26 cm, 0.25 cm and 0.34 cm in LR, SI and AP direction respectively. Using standard imaging (2D-MV portal images) with the e-NAL correction protocol, the systematic set-up errors for tumour bed were 0.23cm, 0.24 cm and 0.28 cm in LR, SI and AP direction respectively.

Two secondary objectives were also addressed: the decrease in safety margin given by clip-based IGRT and the time required to perform clip-based IGRT and standard image assessment. Using IGRT safety (PTV) margins were decreased compared to standard imaging. The study concluded that using clip based registration and the correction protocol, PTV a margin of ≤5 mm for the tumour bed is adequate. Time required for clip-based IGRT verification compared to bony anatomy (standard imaging) verification was technique dependent.
Chapter 5 – The effect of patient and treatment characteristics on set-up accuracy

The work presented in this chapter uses data presented in chapter 4 and is directly related to the primary research objective. The study investigated differences in set-up accuracy using standard imaging between different patient groups, to determine if some patients may benefit more from IGRT than others.

Introduction

Treatment set-up errors, and hence PTV margins, may be influenced by characteristics of the patient and the treatment. Examples of such characteristics include: breast size, tumour bed position and surgical closing technique. Also, it is possible that different patient groups may require different PTV margins which depend on the type of set-up used e.g. laser on skin marks or bony anatomy imaging. Currently uniform TB PTV margins are used across the whole patient population. If the type of imaging, and/or patient and treatment characteristics do influence the size of treatment set up errors then uniform margins may be sub-optimal. If margins are too large this results in the unnecessary irradiation of normal tissues and conversely, smaller PTV margins may lead to the risk of geographical miss of the TB.

The aims of the study described in this chapter are to test whether a set of patient and treatment variables influence set-up errors and to explore the feasibility of individualised TB PTV margins in breast boost radiotherapy.

Materials and methods

Data from 218 patients, from the cohort described in Chapter Four, were used in this work. These data consisted of images plus a set of characteristics hypothesised by 5 clinical oncologists (MM, AK, RJ, CEC and JY) and one breast surgeon (AT) to have an effect on set up errors. These characteristics formed three groups: patient related, surgery related and treatment related.

The image data from the whole cohort (n = 218) were used to calculate treatment set-up errors based on (i) a laser based set-up (no imaging) and (ii) a bony anatomy based set-up
(standard imaging). The population systematic errors ($\Sigma_{\text{laser}}$, $\Sigma_{\text{BA}}$) were calculated from the variance of the individual patients’ systematic set-up errors.

The patient, surgical and treatment characteristics are summarised in Table 11. Tumour bed locations were categorised into regions in the axial and sagittal plane as shown in Figure 12. Breast volume was obtained from the radiotherapy planning CT. Data on surgery: apposed (closed) or unapposed (open) cavity, and seroma were obtained from the surgical notes. A single radiation oncologist rated seroma visibility as not visible/subtle or easily visible and determined the presence of one or more clips placed at the posterior fascia and number of clips placed in the excision cavity.

Figure 12: Schematic diagram to show (a) TB location viewed on axial CT slice (1(blue) =medial, 2(pink) =chest wall, 3(green) =anterior and 4(yellow) =lateral) and (b) TB location in the SI (superior-inferior) direction viewed on sagittal CT slice (1 = superior, 2 = middle and 3 = inferior).
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. of patients with data in each group</th>
<th>Total number of patients with data</th>
<th>Median value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient related:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TB Axial Location (1/2/3/4) (See Figure 12.)</td>
<td>30/96/33/59</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td>TB SI Location (1/2/3) (See Figure 12.)</td>
<td>107/90/21</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td>Breast volume (above median/below median)</td>
<td>109/109</td>
<td>218</td>
<td>855 cm$^3$</td>
</tr>
<tr>
<td><strong>Surgery related:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seroma Visibility (not visible/ easily visible)</td>
<td>158/60</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td>Surgical Closing technique (closed/open)</td>
<td>113/88</td>
<td>201</td>
<td></td>
</tr>
<tr>
<td>No. Clips (above median/below median)</td>
<td>109/109</td>
<td>218</td>
<td>6</td>
</tr>
<tr>
<td>Clip in Posterior Fascia (no/yes)</td>
<td>40/178</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td><strong>Radiotherapy:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time Surgery to Chemotherapy (CT) (days)</td>
<td>101/102</td>
<td>203</td>
<td>133</td>
</tr>
<tr>
<td>Time CT to Radiotherapy (RT) (days)</td>
<td>102/102</td>
<td>204</td>
<td>20</td>
</tr>
<tr>
<td>Trial Arm (control or test)</td>
<td>72/146</td>
<td>218</td>
<td></td>
</tr>
</tbody>
</table>

*Table 11: Patient and treatment characteristics. Characteristics have been categorised according to the information they provide. Median values are given for continuous characteristics.*
Table 11 shows how the data were divided within each patient, surgical or treatment related characteristic. For continuous characteristics the data were dichotomised above and below the median value. For each of the characteristics, differences between population systematic errors between the groups were tested.

Statistical analyses: Data were tested for normality using Shapiro-Wilk test. Fisher’s F-test, or the Levene’s test were used to test for the significance of any difference in the variance of population systematic errors for both laser and bony anatomy set-up ($\Sigma_{\text{laser}}$, $\Sigma_{\text{bone}}$). Where the data were non-normal the Non-Parametric Levene’s Test was used. Significance testing was performed with and without adjustment for multi-testing (Holm-Bonferroni method). Associations between characteristics which gave significantly different systematic errors were investigated using Wilcoxon and Kruskal-Wallis tests.

Results

A total of 1574 images were analysed to provide the set-up data. There was a minimum of 5 images per patient.

Table 12 presents results for the laser based set-up – only results where differences were statistically significant are given. Of the ten characteristics investigated 3 showed a statistical significant difference between the laser based systematic set-up errors $\Sigma_{\text{laser}}$. One of these was a patient characteristic (breast volume) and the other two were surgical characteristics (closing technique and seroma visibility). On application of Holm-Bonferroni correction, only seroma visibility gave a statistical difference between patient groups (in the AP and SI directions). The largest difference in $\Sigma_{\text{laser}}$, of 1 mm, was for the breast volume; all other differences were < 1mm. Differences in $\Sigma_{\text{laser}}$ for breast volume were significant only in the superior-inferior (SI) direction. No association between breast size, tumour bed position, seroma visibility and surgical closing technique was found.
Table 13 gives the results for the bony anatomy based set-up, again only for characteristics where differences were statistically significant. There were only 2 characteristics where this was the case, both were patient related: breast volume and tumour bed axial position. Both characteristics gave significant differences when Holm-Bonferroni correction was applied, although TB axial location no longer affected set-up error in the LR direction. All differences in population systematic error, $\Sigma_{\text{bone}}$ were < 1mm. Again, breast volume differences were significant in the SI direction and neither of the other directions.

No radiotherapy characteristics, clip placement or number, or tumour bed superior-inferior position were related to population systematic set-up errors in this study.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group 1</th>
<th>$\Sigma_{\text{laser}}$ (cm)</th>
<th>Group 2</th>
<th>$\Sigma_{\text{laser}}$ (cm)</th>
<th>p-value</th>
<th>Direction and magnitude of difference in $\Sigma$ (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Breast volume</em></td>
<td>&lt; 855 cm$^3$</td>
<td>0.25</td>
<td>$\geq 855$ cm$^3$</td>
<td>0.32</td>
<td>0.03</td>
<td>SI 0.07</td>
</tr>
<tr>
<td><em>Seroma Visibility</em></td>
<td>not visible/subtle</td>
<td>0.28</td>
<td>easily visible</td>
<td>0.35</td>
<td>0.02</td>
<td>LR 0.07</td>
</tr>
<tr>
<td></td>
<td>not visible/subtle</td>
<td>0.26</td>
<td>easily visible</td>
<td>0.32</td>
<td>0.002</td>
<td>SI 0.06</td>
</tr>
<tr>
<td></td>
<td>not visible/subtle</td>
<td>0.31</td>
<td>easily visible</td>
<td>0.41</td>
<td>0.005</td>
<td>AP 0.10</td>
</tr>
<tr>
<td><em>Surgical Closing technique</em></td>
<td>closed</td>
<td>0.27</td>
<td>open</td>
<td>0.33</td>
<td>0.02</td>
<td>LR 0.06</td>
</tr>
<tr>
<td></td>
<td>closed</td>
<td>0.25</td>
<td>open</td>
<td>0.32</td>
<td>0.04</td>
<td>SI 0.07</td>
</tr>
</tbody>
</table>

*Table 12:* Systematic ($\Sigma_{\text{laser}}$) laser set-up (no imaging) errors for patients grouped using patient and treatment related characteristics. For $\Sigma_{\text{laser}},$ P-values from the non-parametric Levene’s test are given. Values are only shown for characteristic that gave a significant difference between patient groups ($p<0.05,$ without Holm-Bonferroni adjustment). The directions in which differences occur are indicated (SI denotes superior-inferior, LR denotes left-right and AP denotes anterior-posterior).
Table 13: Systematic standard imaging set-up errors \( (\Sigma_{BA}) \) for patients groups determined using patient and treatment related characteristics. \( P \)-values for non-parametric Levene’s test are given. Data only given for characteristic that gave a significant difference in systematic bony anatomy verification error between patient groups \( (p<0.05\text{, without Holm-Bonferroni adjustment}) \). The direction in which errors were different is indicated.

### Discussion

The purpose of this work was to identify whether a set of patient and treatment variables influence set-up errors, and to explore the feasibility of individualised TB PTV margins in breast boost radiotherapy using no imaging or standard imaging. The study has shown that two patient characteristics (breast volume and TB axial location) and two surgical characteristics (closing technique and seroma visibility) affect laser and bony anatomy based set-up accuracy. No radiotherapy characteristics were found to have a significant effect on set-up errors.

Population systematic errors for both laser and bony anatomy set-up \( (\Sigma_{laser} \text{ and } \Sigma_{BA}) \) were greater (by 2mm and 1 mm) for patients with a breast volume greater than 850 cm\(^3\) but this was only statistically significant for the SI direction of movement. This may be because breast tissue moves more independently of bony anatomy and skin based tattoos in larger breast women. Hasan et al.\(^9\) had previously reported a weak correlation between the mean patient set-up error, measured using upon bony anatomy, and breast volume \( (n=27\text{ and } p = \)
0.02). They did not investigate, however, the association of the population systematic set-up error with breast volume as we have done in this work.

If patients are set-up at the time of treatment with lasers matched to skin marks, then those patients with easily visible seroma and open surgical cavities have a statistically significantly increased population systematic error, although this is small in magnitude (0.6 to 1.0 mm depending of the direction of movement). It is possible that this increase is due to changes in the location of the clips between the planning CT scan and treatment due to shrinking seroma or clip migration. If this were the case, then $\Sigma_{BA}$ would also be affected by seroma visibility and surgical closing technique. A greater value for $\Sigma_{BA}$ was not observed which indicates that the observed differences in $\Sigma_{laser}$ are probably not due to changes in clip location.

If patients are set-up using standard imaging, we found that population systematic errors were influenced by the tumour bed location. Patients with medially located tumour beds had smaller $\Sigma_{BA}$ in the AP direction, while patients with laterally located TBs had larger $\Sigma_{BA}$ in LR direction by 0.6 and 0.7 mm respectively. It is likely that there is little movement of medial breast tissue compared to bony anatomy and greater movement of lateral breast tissue which may explain these results. This is also supported by Hasan et al. who found correlation of 3D bony anatomy verification errors with distance from the chest wall ($p = 0.003$). Similarly, Topolnjak et al. showed that the distance of the TB from the chest wall was correlated with the difference between TB set-up error for chest wall and the breast surface ($r = 0.476$, $p = 0.034$).

All differences in population systematic error were small < 1mm, however, these systematic errors make the greatest contribution to tumour bed PTV margins (PTV margin = 2.5$\Sigma$ + 0.3$\sigma$). The largest difference in systematic errors (1 mm) was observed with a laser based set-up in the AP direction between patients with easily visible or not visible seroma ($\Sigma_{laser} = 0.31$ cm for smaller breasted patients compared to $\Sigma_{laser} = 0.41$ for larger breasted patients). This gives a difference in PTV margin of 2.5 mm and indicates larger margins for patient with large seroma may be appropriate for laser-based set-up. All others changes in margins were estimated to be 2 mm or less.
This study used univariate analysis to identify several variables which may be used to group patients with smaller or larger systematic errors. A limitation of this study is that no multivariate analysis has been employed. However, no standard multivariate statistical model was identified which was suitable to test the interaction of multiple variances. Care should be taken when interpreting p-values, presented in tables 12 and 13, as these have not been adjusted for multiple testing. Using, one method to control false positive results, Holm-Bonferroni correction, differences between patients grouped using seroma visibility and TB axial position remain significant. A further limitation of our study may be the use of only one observer to grade seroma visibility. In Lee et al.\textsuperscript{107}, seroma visibility in 20 patients was scored by radiation oncologists and radiographers using the Clarity Visualisation Score (CVS) which grades seroma visibility on a scale of 1 to 5. There was a 0.2 difference in the median grade between the two groups (3.8 v. 3.6). Amongst radiation oncologists, all grades agreed with median CVS agreed within 1 grade except in 1 of 20 cases. These variations amongst observers are small and it is expected would be smaller still if only two ranks are used, as is the case in our study. Consequently, we expect any observer error to be small. Furthermore, to minimise intra-observer error, the radiation oncologist (MM) scoring seroma used a pre-defined protocol and has previously outlined/scored seroma visibility on nearly 800 patients (Mukesh et al.\textsuperscript{108}).

The consequences of small changes in PTV margins have been investigated and are described in Chapter 6. Reductions in margins may reduce the dose to normal tissues such as heart and lungs. Darby et al.\textsuperscript{109} have recently reported evidence a zero threshold for cardiac toxicity so even small changes in mean heart dose are of importance given the large numbers of patients receiving radiotherapy for breast cancer.

**Conclusions**

Patient and surgical related characteristics have limited affect on population systematic errors derived from laser based (no imaging) and bony anatomy (standard imaging) set-up methods. Four groups have been identified who may benefit modestly from reduced PTV margins: women with breast volume of $< 850 \text{ cm}^3$, those with invisible/subtle seroma, closed cavities or medial tumour bed locations.
Chapter 6 – The Impact of Image Guidance on Dose Distributions in Breast Boost Radiotherapy

This work described in this chapter investigates the decrease in normal tissue irradiated to 95% of the breast boost dose if standard imaging is used, compared to clip-based IGRT, a secondary research objective of this study. It also investigates the effect of clip-based IGRT on dose to the heart and lungs, and the effect of IMPORT high trial arm on the volume of tissue spared.

Introduction

Whole breast irradiation (WBI) following breast conserving surgery is a standard treatment for patients with breast cancer. As part of their radiotherapy, patients at high risk of recurrence receive a boost dose to the region around the tumour bed\textsuperscript{10}. Evidence shows that accurate localisation of the tumour bed can only be achieved if internal markers are used to indicate its position on imaging, particularly CT images\textsuperscript{11,12}. This approach to improve localisation allows conformal photon dose distributions to be used to deliver the tumour bed boost dose whilst minimising dose to normal tissues. This further enables studies of dose escalation using sequential or integrated boost techniques\textsuperscript{57,113}.

Accurate patient set-up and in-treatment verification is essential for the delivery of conformal radiotherapy and higher boost doses. As discussed in the two previous chapters, standard verification imaging for breast radiotherapy uses bony anatomy and often the outline contour of the breast to match electronic portal images acquired at megavoltage energies (MV) to pre-treatment images – often digitally reconstructed radiographs. This is widely available and requires a surrogate for the tumour bed, such as the patient’s ribs, as neither the tumour bed nor implanted surrogate markers are visible on MV images. Gold markers are visible on MV images and have been demonstrated in this setting, but are not widely used surrogates\textsuperscript{94}. Kilovoltage energy imaging is necessary to visualise surgical clips, hence determining tumour bed positions most accurately and bringing the potential to decrease tumour bed Planning Target Volumes (PTV).

Previous studies have shown that 5 mm tumour bed PTV margins of 5 mm can be achieved\textsuperscript{94,114} using IGRT. The purpose of the work in this chapter was to evaluate the
dosimetric impact in terms of doses to breast tissue and organs at risk as a result of the use of IGRT and the reduced margins it allows in breast boost radiotherapy.

Materials and Methods

The CT data sets, used for treatment planning, for patients treated in the IMPORT HIGH trial were selected sequentially form an alphabetical ordered list. The patients were treated between July 2009 and December 2011. As discussed above, the patients had surgical clips implanted close to the tumour bed at during breast conserving surgery (BCS). Two target volumes were defined. Firstly a Clinical Target Volume (CTV) was defined for the tumour bed (CTV_TB). This encompassed the surgical clips, plus any seroma and architectural distortion. Secondly a target volume for the whole breast (WB_TV) was defined from the extent of the treatment fields for the whole breast. This excluded the lung and ribcage and tissue within 5 mm of the surface of the skin. On the treatment plans, organs at risk were delineated. These were: the ipsilateral lung, contralateral lung, heart and contralateral breast.

The CTV_TB structure was expanded to create two Planning Target Volumes (PTV_TB) for each data set as shown in figure 13: the first using a 5 mm margin as required by IMPORT HIGH. We have previously described\textsuperscript{94,114} how the use of image guidance based on imaging the positions of markers implanted in the tumour bed, coupled to set-up correction strategies reduces population set up errors to the extent that a 5 mm tumour bed PTV margin may be achieved. The second PTV_TB margin was 8 mm and was derived from the analysis presented in chapter 4 for bony anatomy based set-up measurement. Systematic errors in patient set up were determined based on measurements from images of the first 3 fractions and a correction applied on fraction 4. This approach enabled a comparison between standard bony anatomy based verification and the IGRT approach.

30 patients were planned using a sequential, conformal photon boost to the tumour bed and 30 using the simultaneous integrated boost (SIB) technique. The sequential boost technique delivered a phase 1 WBRT dose of 40 Gy in 15 fractions followed by 16 Gy in 8 fractions to the tumour bed boost volume only for phase 2. The concomitant boost technique involved delivering 15 fractions with a total dose 36 Gy to the whole breast using tangential fields; plus 40 Gy to the partial breast volume and an escalated dose to the tumour bed via co-planar
conformal fields. The escalated dose was 48Gy or 53Gy (15 patients each) depending on randomisation\textsuperscript{57} (see figure 2). The criteria for plan assessment and constraints on the organ at doses used for the IMPORT HIGH trial were used to guide the planning (Table 14). Plans were generated using the Philips Pinnacle\textsuperscript{3} (Philips Medical System, Netherlands) treatment planning system (v8.0 and v9.0). The forward treatment planned method reported by Donovan et al.\textsuperscript{115} was used, with the collapsed cone convolution dose calculation algorithm on a 0.25 cm × 0.25 cm × 0.25 cm calculation grid. The beam quality used was 6MV in most cases.

Figure 13: section through a patient’s treatment plan showing CTV\_TB (red), PTV\_TB = 5 mm (yellow) and PTV\_TB = 8 mm (pink).

Statistical Analysis: For the analysis of the treatment plans produced to determine the impact of IGRT on the breast boost plans, the main metric used was the volume of tissue receiving 95% of the tumour bed dose. Data were also collected on the doses to the lungs, heart and contralateral breast from the plan assessment criteria listed in Table 14. Mean heart and lung doses were also collected. The data were tested for normality. The Wilcoxon signed ranks test was used to test for statistical significance of the differences in the various metrics between the plans with 5 mm and 8 mm PTV\_TB margins. Data were dichotomised by tumour bed laterality and the Mann-Whitney test statistic was used to determine the significance of any differences observed.
This work is novel and there were no similar studies in the literature on which to base estimates for sample size calculations. The heart is one of the most important organs at risk in breast radiotherapy. The results from planning the first ten cases were used to estimate a sample size of 58 cases, which gave 90% power to determine a difference of 0.2Gy at a significance level of 0.05. The additional cases were included to allow for any unforeseen problems with the data.
### Sequential Boost

<table>
<thead>
<tr>
<th>Location</th>
<th>Minimum Dose</th>
<th>Median Dose</th>
<th>Maximum Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole breast</td>
<td>&gt; 90% volume &gt; 36 Gy</td>
<td>40 to 44 Gy</td>
<td>&lt; 5% volume &gt; 56 Gy</td>
</tr>
<tr>
<td>Tumour bed PTV</td>
<td>&gt; 95% volume &gt; 53.2 Gy</td>
<td>55.5 to 56.5 Gy</td>
<td>&lt; 5% volume &gt; 60 Gy</td>
</tr>
</tbody>
</table>

### Concomitant Boost

<table>
<thead>
<tr>
<th>Location</th>
<th>Minimum Dose</th>
<th>Median Dose</th>
<th>Maximum Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole breast</td>
<td>&gt; 90% volume &gt; 32.4 Gy</td>
<td>34 to 37 Gy</td>
<td>&lt; 5% volume &gt; 40 Gy</td>
</tr>
<tr>
<td>Partial breast PTV</td>
<td>&gt; 90% volume &gt; 36 Gy</td>
<td>40 to 44 Gy</td>
<td></td>
</tr>
<tr>
<td>Tumour bed PTV</td>
<td>&gt; 95% volume &gt; 45.6 Gy or 50.4 Gy</td>
<td>47.5 to 48.5 Gy or 52.5 to 53.5 Gy</td>
<td>&gt; 3% volume &gt; 51.4 Gy or 56.7 Gy</td>
</tr>
</tbody>
</table>

### Organs at Risk

<table>
<thead>
<tr>
<th>Organ</th>
<th>Dose (Gy)</th>
<th>Maximum Allowed Volume (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ipsilateral Lung</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>Contralateral Lung</td>
<td>2.5</td>
<td>15</td>
</tr>
<tr>
<td>Heart</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Contralateral Breast</td>
<td>Mean Dose &lt; 0.5 Gy</td>
<td>Permitted Maximum Mean Dose 1.5 Gy</td>
</tr>
</tbody>
</table>

Table 14: Radiotherapy treatment planning constraints for IMPORT HIGH. Bold indicates mandatory constraints. Where two dose levels are given they are for the 48Gy or 53Gy test arm doses.

**Results**

Of the patients recruited to this study, 35 had left breast disease and 25 had right breast disease. The median CTV_TB volume was 10.2 cm³ (range 2.4 – 205.0 cm³). There was no statistical significant difference in the CTV_TB or PTV_TB volumes (grouped into 5mm and
8 mm margins) between the sequential and concomitant boost plans, or the concomitant boost plans at boost doses of 48 Gy or 53 Gy.

Table 15 summarises the volumes of breast tissue and the percentage of whole breast volume receiving 95% of the dose prescribed to the tumour bed. There was a statistically significant difference (p < 0.01) between the volumes of breast tissue receiving a high dose for the two types of plan, with the volumes larger in the sequential boost plans. The magnitude of the volume changes between a PTV_TB of 5 mm and a PTV_TB of 8 mm was not different between: i) sequential and concomitant boost plans; ii) between left and right breast plans; and iii) between the 48Gy and 53Gy boost doses. The difference data for sequential and concomitant boost treatments were combined and the median decrease in the high dose volume, for IGRT was found to be 29 cm³ (range 11 to 193 cm³). This equates to an additional 3.3% (median value) up to a maximum of 11.8% of the whole breast volume spared high dose irradiation from these boost techniques, if clip-based IGRT is used.

All dose metrics for the organs at risk increased with the use of the 8 mm margin for standard verification, compared to the IGRT margins. This was as anticipated and a modest effect (table 16). Of the various metrics, only mean heart dose and V₁₃Gy for the heart had a statistically significant relationship with tumour bed laterality (p < 0.01); with higher values in the left breast group.

In the case of the IGRT margins of 5 mm, 56 of the 60 cases met all the treatment planning criteria (table 14). The minimum dose coverage of the tumour bed was between 91% and 95% in the other four cases (2 sequential and 2 concomitant boost). These were all left breast cases and had the tumour bed in close proximity to the chest wall, hence the PTV_TB extended into lung and the heart (see figure 14). Thus a compromise was accepted between target coverage and heart dose for the clinical treatment. Also, in one of the two concomitant boost cases that failed the planning criteria, the maximum volume limit for the highest dose to whole breast was exceeded (9.7% compared with 5%). This patient had a large CTV_TB of 49 cm³ whereas the median CTV_TB volume for the patients in the study was 9 cm³, hence the difficulty in obtaining a dose distribution that meet all of the treatment planning requirements.
As the planning objectives used in IMPORT HIGH were based on a PTV_TB margin of 5 mm, it was likely that increasing this margin to 8 mm would cause more plans to fail the criteria. This was found to be the case with four sequential boost plans and 10 concomitant boost plans breaching mandatory planning constraints. In all the sequential boost cases and eight of the concomitant cases, the PTV_TB coverage was below 95% but above 91%. In three of the concomitant boost cases, the maximum dose criterion in the whole breast volume was exceeded by 2% in two cases and 5% in the third. In nine of the concomitant boost cases the median dose constraint to the partial breast volume of 40 to 44 Gy was exceeded by between 0.5Gy and 3.7 Gy.

<table>
<thead>
<tr>
<th></th>
<th>PTV_TB = 5mm</th>
<th>PTV_TB = 8mm</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High dose Volume</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential boost</td>
<td>91 (30 – 863)</td>
<td>125 (42 – 1005)</td>
<td>33 (11 - 193)</td>
</tr>
<tr>
<td>Concomitant boost</td>
<td>60 (19 - 228)</td>
<td>87 (30 – 260)</td>
<td>23 (11 - 66)</td>
</tr>
<tr>
<td>Both combined</td>
<td>69 (19 - 863)</td>
<td>100 (30 – 1005)</td>
<td>29 (11 - 193)</td>
</tr>
<tr>
<td><strong>Percentage of whole</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>breast volume (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential boost</td>
<td>10 (4 – 35)</td>
<td>14 (5 - 41)</td>
<td>4 (2 – 12)</td>
</tr>
<tr>
<td>Concomitant boost</td>
<td>8 (2 – 19)</td>
<td>11 (4 – 24)</td>
<td>3 (1 – 6)</td>
</tr>
<tr>
<td>Both combined</td>
<td>3.0 (1 – 6)</td>
<td>4 (2 – 12)</td>
<td>3 (1 – 12)</td>
</tr>
</tbody>
</table>

Table 15: Volumes of the breast receiving 95% of prescribed dose from plans based on 5 mm and 8 mm PTV_TB margins. Data are given as median (range) and presented in absolute volume (cc) and as a percentage of the whole breast volume. Differences between 95% volumes for the two PTV_TB were statistically significant ($p<0.01$) and are given in column 4.
Figure 14: A) shows an original PTV_TB with 5 mm margin and in close proximity to the lung. When PTV_TB is increased by a further 3 mm it expands into the lung B) and to achieve coverage requires an increase in the width of the tangential fields by 4 mm C) which in turn increasing the dose to heart.
<table>
<thead>
<tr>
<th></th>
<th>PTV_TB = 5 mm</th>
<th>PTV_TB = 8 mm</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ipsilateral lung $V_{18\text{Gy}}$ (%)</td>
<td>9.6 (1.9 - 27.6)</td>
<td>10.0 (2.3 - 27.8)</td>
<td>0.3 (-0.9 - 5.0)</td>
</tr>
<tr>
<td>Ipsilateral lung mean dose (Gy)</td>
<td>5.6 (2.6 - 11.3)</td>
<td>6.1 (2.8 - 11.5)</td>
<td>0.3 (-0.7 - 2.7)</td>
</tr>
<tr>
<td><strong>Contralateral lung $V_{2.5\text{Gy}}$ (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential boost</td>
<td>0.0 (0.0 - 12.2)</td>
<td>0.1 (0.0 - 13.9)</td>
<td>0.0 (-3.3 - 7.5)</td>
</tr>
<tr>
<td>Concomitant boost</td>
<td>1.6 (0.0 - 13.4)</td>
<td>3.2 (0.0 - 17.4)</td>
<td>1.0 (-2.5 - 16.1)</td>
</tr>
<tr>
<td>Contralateral lung mean dose (Gy)</td>
<td>0.4 (0.1 - 1.2)</td>
<td>0.5 (0.1 - 3.3)</td>
<td>0.1 (-0.2 - 3.1)</td>
</tr>
<tr>
<td>Contralateral breast mean dose (Gy)</td>
<td>5.0 (0.0 - 1.8)</td>
<td>5.0 (0.0 - 1.4)</td>
<td>0.1 (-1.3 - 0.4)</td>
</tr>
<tr>
<td><strong>Heart mean dose (Gy)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right breast cases</td>
<td>1.2 (0.4 - 2.2)</td>
<td>1.4 (0.5 - 2.6)</td>
<td>0.2 (-0.3 - 1.6)</td>
</tr>
<tr>
<td>Left breast cases</td>
<td>1.9 (0.6 - 5.1)</td>
<td>2.1 (0.6 - 6.0)</td>
<td>0.2 (-0.2 - 1.0)</td>
</tr>
<tr>
<td><strong>Heart $V_{1\text{Gy}}$ (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right breast cases</td>
<td>0.0 (0.0 - 0.0)</td>
<td>0.0 (0.0 - 0.2)</td>
<td>0.0 (0.0 - 0.2)</td>
</tr>
<tr>
<td>Left breast cases</td>
<td>0.2 (0.0 - 5.5)</td>
<td>0.4 (0.0 - 6.3)</td>
<td>0.1 (-0.1 - 2.4)</td>
</tr>
</tbody>
</table>

*Table 16: Dosimetric data given as median (range) for each of the assessment criteria. All differences were found to be statistically significant ($p < 0.01$).*

**Discussion**

The work presented in this chapter evaluated the impact of IGRT on normal tissue doses in breast radiotherapy. We have seen in the two previous chapters that IGRT enables smaller PTV margins due to the reduction of set-up errors across the patient population. The reduced
margins achievable with the use of internal markers (and associated image guidance) led to a reduction of 29 cm$^3$ (range 11 – 193 cm$^3$) in the volume of breast tissue receiving a high dose. This is a consequence of the ability to reduce the PTV_TB margin from 8 mm to 5 mm.

In chapter two we discussed the evidence for a dose/volume relationship for normal tissue toxicity in breast tissue. This is still an open question and the dose/volume constraints needed are still a subject of research$^{116}$. Hence it is unclear what outcome effect is expected at the dose levels and volumes reported in this chapter. In the EORTC study of Bartelink et al.$^2$ they reported that an WBRT dose of 50 Gy followed by a boost dose of 16 Gy, the 10 year risk of fibrosis increased by approximately 15% from 13.2% (for the no boost group) to 28.1% (for the boost group). Patients in this study were treated with a PTV margin of 15 mm compared to the much smaller values of 5 and 8 mm discussed here. The sequential boost prescription used in this study was 16 Gy in 2 Gy fractions and hence is expected to lead to a lower rate of fibrosis that than in the EORTC study, due to the smaller high-dose volumes.

The larger PTV_TB margin had modest impact on the calculated doses to organs at risk for both types of boost plans: sequential and concomitant. The majority of the therapeutic dose was delivered using standard tangential fields, which maintained organ at risk sparing in these complex situations for all PTV_TB margins. The recent cardiac risk study of Darby et al.$^{109}$ suggests that small changes in heart dose are important as a consequence of the linear relationship between mean heart dose and Major Coronary Events (MCE). Hence given the tens of thousands of women treated each year with radiotherapy for breast cancer, a modest reduction in heart dose may impact significantly on the rate of MCE in the survivor population.

One finding of this study was the increased difficulty in meeting the planning requirements for the boost which an increase in the PTV_TB margin of only 3 mm. This had most impact for the concomitant boost plans where one in three patients failed at least one of the dosimetry criteria.

The compromised median dose to the partial breast in nine cases is particular relevant for the IMPORT HIGH trial which requires discrimination between the three dose levels of the whole breast, partial breast and tumour bed. This work shows that image guidance is
necessary to achieve this level of dose discrimination: an additional benefit to reduced normal tissue doses.

Conclusions

The reduction in normal tissue irradiated when using clip-based IGRT was modest (29 cm$^3$). For breast radiotherapy methods involving a complex boost technique, image guidance is important as it allows the dose levels to be sufficiently discriminated. Its use allows some reduction in the dose to breast, heart and lung for both sequential boost and concomitant boost approaches.
Chapter 7 – Discussion and conclusions

The previous five chapters have summarised the key outcomes of this Efficacy and Mechanism programme. The chapters have been grouped by the key stages of the programme. Chapter 2 and 3 discuss the evidence for a dose volume effect in breast radiotherapy following BCS. Chapters 4, 5 and 6 discuss the evaluation of IGRT in this setting.

The critical review in chapter 2 explored the evidence in the literature for a dose volume effect in normal breast tissue. The review found differing results from studies addressing the relationship between irradiated breast volume and late breast tissue complications. For example, Borger et al reported strong evidence for a volume effect.14 They found that for every 100 cm$^3$ increase in the boost volume, the risk of fibrosis increased by a factor of four and that a two-fold increase in boost volume results in an 11% reduction in tolerance dose (NTD$_{50}$). Borger’s study used low dose rate iridium brachytherapy implants, which produce a very high dose region within the implant and rapid fall off of dose outside the implant. The other studies discussed in chapter 2 include partial breast irradiation (PBI) and intra-operative radiotherapy (IORT) trials and matched case series that compared PBI with whole breast irradiation (WBI). Whilst these other studies generally suggested some evidence of a relationship between volume irradiated to high dose and normal tissue complication probability (NTCP), they did not quantify the volume effect. The brachytherapy and intra-operative dose distribution can differ from the external beam radiotherapy (teletherapy) and therefore, it is unclear whether these results can be extrapolated to external beam techniques. There are several ongoing external beam radiotherapy breast trials that are designed to provide further data in this direction: IMPORT LOW$^{56,57}$ and the Danish Breast Cancer Cooperative Group trial$^{62}$ are trials which compare PBI with WBI and IMPORT HIGH$^{56,57}$ is investigating the effects of three dose regions throughout the breast using intensity modulated radiotherapy (IMRT) for dose delivery, imaged with clip-based IGRT.

In chapter 3, we addressed a secondary research objective of this study: to estimate the reduced risk of late adverse effects resulting from the smaller tissue volume irradiated, using data generated and published from earlier randomised trials conducted by members of our collaboration.
Individual patient data of 5856 patients from the Cambridge trial\textsuperscript{66,72} and EORTC trial\textsuperscript{66,72} were used for the analysis with moderate-severe breast fibrosis as the radiotherapy toxicity endpoint. Fits to the data using two standard models of NTCP (the LKB\textsuperscript{74} and Niemierko\textsuperscript{75} models) produced a volume parameter ‘n’ close to zero, suggesting that for moderate and severe fibrosis, the breast acts as an organ with serial structure (as discussed in chapter 3). These results were successfully validated on an independent dataset and indicated that for moderate-severe breast fibrosis, the maximum radiotherapy dose is the most important parameter rather than volume of tissue irradiated. Based on this model, a change in volume of normal tissue irradiated will not change the risk of breast fibrosis. Clearly, any model has limitations and the mature results from the clinical trials addressing this question are awaited.

In chapter 4, we addressed the primary research objective of this study, namely the difference in accuracy of clip-based IGRT and standard imaging using bony anatomy. We also addressed two secondary objectives: 1) the reduction in safety planning target volume (PTV) margin, and 2) the time required for clip-based IGRT and standard imaging. We presented results of the analysis of the impact of clip-based IGRT on set-up errors and treatment margins in patients recruited to the IMPORT HIGH study. To our knowledge this is the largest study to evaluate IGRT in the breast radiotherapy setting. This study found that clip-based IGRT was more accurate than standard imaging. The population systematic error was between 2 to 4 mm greater with standard imaging. A key finding of the study was that a PTV boost safety margin of 5 mm is sufficient if clip-based IGRT is employed. The clip-based IGRT approach was based on imaging the positions of titanium clips implanted in the tumour bed at time of breast conserving surgery (BCS). In contrast, standard imaging using bony anatomy required an 8 mm PTV boost safety margin and no imaging (i.e. set-up based on laser-based alignment of the patient surface) required a 10 mm margin. These results indicate that for patients receiving concomitant tumour bed boost, a margin less than 8mm cannot be safely used without clip-based IGRT as there is a risk of geographical miss of the tumour bed being treated within the high dose region. The difference in time required to perform clip-based IGRT and standard image assessment was technique dependent. Clip-based IGRT was quicker than standard imaging when using 2D-KV technique, but not when using cone beam CT imaging.

Chapter 5 explored the factors influencing the primary objective. It tested the hypothesis that some characteristics of the patient and treatment may influence: i) the relationship between
the set-up error measured with IGRT and with bony anatomy imaging (standard imaging), and ii) the relationship between the set-up error measured with IGRT and with laser-based alignment (no imaging). Patients with larger breasts required a larger PTV margin for both standard imaging and no imaging. Seroma visibility and surgery technique both affected no imaging set-up errors, whereas tumour bed position affected standard imaging based set-up errors. The work implies that clip-based IGRT may be of greater benefit compared to standard imaging or no imaging, for some patient groups and that treatment margins can be modified accordingly if clip-based IGRT is not available.

Chapter 6 evaluates the reduction in volume of normal tissue receiving 95% the high boost dose of radiation when clip-based IGRT is used compared with standard imaging, which is a secondary objective of this study. The consequences of the results of the IGRT study on treatment planning were evaluated. The main quantitative finding was that the use of clip-based IGRT allowed the volume of breast tissue irradiated to a high dose to be reduced by 29 cm$^3$ (with a range of 11 to 193 cm$^3$) for the 60 cases studied. The use of smaller PTV margins with clip-based IGRT also allowed a small reduction in the radiotherapy dose to the contralateral breast, heart and lung. The larger margins needed with standard imaging meant that the treatment planning constraints for the dose boost could not be met in some cases.

In conclusion, this research demonstrates the benefits of clip-based IGRT over standard imaging, with a reduction in PTV margins. Margins less than 8 mm cannot be safely used without clip-based IGRT for patients receiving concomitant tumour bed boost as there is a risk of geographical miss of the tumour bed being treated within the high dose region. In principle, these smaller, but accurately placed margins may influence local control rates, but this needs to be evaluated from mature clinical trial data in the future. We have not been able to develop a model that can predict the effect of irradiated volume on breast tissue toxicity, but mature results from the ongoing clinical trials may provide a definitive answer.
Acknowledgements

Contributions of authors

Emma Harris (Research Physicist and Principle Clinical Scientist) study design, physics input and analysis for all aspects, drafts of physics aspects of report
Mukesh Mukesh (Clinical Research Fellow) clinical input and analysis for all aspects, drafts of clinical oncology aspects of report
Rajesh Jena (Clinician Scientist & Consultant in Radiation Oncology) contributed to study design, clinical oncology input to all aspects
Angela Baker (Therapy Research Radiographer) data acquisition and analysis
Harry Bartelink (Professor, Consultant Clinical Oncologist) input to analysis of NTCP parameters for breast fibrosis
Corrine Brooks (Planning Radiographer) contributed to treatment planning aspects
Sandra Collette (Statistician) input to analysis of NTCP parameters for breast fibrosis
June Dean (therapy radiographer) contributed to data acquisition and analysis
Ellen Donovan (Principle Clinical Scientist) Treatment planning and margin modelling
Sally Eagle (Therapy Superintendent Radiotherapy) data acquisition and analysis
John Fenwick (Physics Research Team Leader) contributed to study design, radiobiology
Peter H Graham (Associate Professor Clinical Oncology) input to analysis of NTCP parameters for breast fibrosis
Jo Haviland (Medical Statistician) contributed to study design, study statistician
Anna Kirby (Consultant Clinical Oncologist) contributed to study design, clinical oncology input
Helen Mayles (Head of Clinical Radiotherapy Physics) contributed to study design, management, data acquisition and analysis
RA Mitchell (Trainee Medical Physicist) contributed to treatment planning aspects
Rosalind Perry (Treatment Area Supt Radiographer) contributed to study design, data acquisition and analysis
Philip Poortmans (Consultant Radiation Oncologist) input to analysis of NTCP parameters for breast fibrosis
Andrew Poynter (Head of Radiotherapy Physics) contributed to study design, management, data acquisition and analysis
Glyn Shentall (Head of Radiotherapy Physics) contributed to study design, management, data analysis and manuscript preparation

Jenny Titley (Trial Manager) contributed to study design, Manager of study and interface to IMPORT HIGH

Alistair Thompson (Professor of Surgical Oncology) input on surgical aspects

John Yarnold (Professor and Consultant Clinical Oncologist) contributed to study design, clinical oncology input to all aspects, interface to IMPORT HIGH

Charlotte Coles (Consultant in Clinical Oncology) Joint chief investigator - contributed to study design, management, clinical oncology input to all aspects, interface to IMPORT HIGH

Philip Evans (Professor of Medical Physics and Medical Imaging) Joint chief investigator - contributed to study design, management, physics input to all aspects

Other Acknowledgements

- The IMPORT HIGH trial patients who agreed to their imaging and radiotherapy planning data being used for this study
- Sue Cross and Christine Rawlings from Torbay for their contribution to the development of the project and grant
- Sandra Collette for input to the analysis of NTCP parameters for breast fibrosis
- Alison Round for help with the work at Ipswich
- The members of the IMPORT HIGH and IMPORT LOW trials Trial Management Group for supporting this study as a sub-study of IMPORT HIGH
- The members of the IMPORT HIGH and IMPORT LOW trials Independent Data Monitoring Committee for supporting this study as a sub-study of IMPORT HIGH
- The members of the IMPORT HIGH and IMPORT LOW trials Trial Steering Committee for supporting this study as a sub-study of IMPORT HIGH
- The radiographers, physicists, engineers, clinical oncologists and surgeons at each centre that provided data for this study.
- Dr Ellen Donovan is funded by a National Institute for Health Research/CSO Healthcare Scientist Post-Doctoral Fellowship. This article/paper/report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
**Publications**


**References**


http://www.clinicaltrials.gov/ct2/show/NCT00892814

http://www.clinicaltrials.gov/ct2/show/NCT01247233


