Optimising trial methodologies to maximise trial efficiencies: a case study in breast cancer radiotherapy trials

Indrani Subarna Bhattacharya

The Institute of Cancer Research, London

Submitted for the degree of PhD

August 2019

Declaration of Originality

I declare that the work submitted in this thesis is based on research carried out by the author at The Institute of Cancer Research, London. It is all the author's own work except where otherwise acknowledged.

Abstract

Background: In an era of falling local relapse rates, the risk-benefit ratio of adjuvant breast radiotherapy requires careful consideration, and the collection of normal tissue effect (NTE) data optimised. For patients at lowest risk of local recurrence, the risk of NTE may outweigh the benefits and radiotherapy can be avoided. Where a component of standard of care is being removed, there is an opportunity to improve patient information provision.

Methods: A series of exploratory analyses was conducted within the IMPORT trials to investigate whether patient-reported outcomes (PRO) can be used as primary toxicity endpoints in breast radiotherapy trials as well as determining how PRO change over time and whether baseline characteristics can predict PRO. A patient decision aid (PDA) was developed and tested within PRIMETIME to investigate whether the PDA reduced decisional conflict.

Results: Patients reported NTE more frequently than clinician-reported outcomes (CRO) or photographs. Concordance between PRO and CRO or photographs was poor on an individual patient level. However, the results from the comparison of radiotherapy schedules were consistent between PRO and CRO or photographs. Most NTE reduced over time except for breast shrinkage which increased. Baseline predictors of PRO included younger age and larger breast size. Seroma was not associated with worse NTE, but haematoma and smoking were significant risk factors. Decisional conflict scores were low in PRIMETIME and there were no clinically significant reductions after PDA implementation. Around 50% patients did not use the PDA.

Conclusions: PRO can be used as primary toxicity endpoints in breast radiotherapy trials. Baseline predictors of PRO can contribute to the informed consent process for patients considering breast radiotherapy. PRIMETIME-eligible patients at low risk of recurrence displayed low decisional conflict scores and only half used the additional PDA suggesting that standard patient information was sufficient for this patient group.

Publications

Bhattacharya IS, Bliss JM: Clinical Trials From the Other Side: Lessons Learned by a Clinician Venturing Into a Clinical Trials Unit. Clin Oncol (R Coll Radiol) 31:420-423, 2019

Bhattacharya IS, Haviland JS, Perotti C, Eaton D, Gulliford S, Harris E, Coles CE, Kirwan CC, Bliss JM, Kirby AM; IMPORT Trialists. Is breast seroma after tumour resection associated with patient-reported breast appearance change following radiotherapy? Results from the IMPORT HIGH (CRUK/06/003) trial. Radiother Oncol 136:190-196, 2019

Bhattacharya IS, Haviland JS, Hopwood P, Coles CE, Yarnold JR, Bliss JM, Kirby AM; IMPORT Trialists. Can patient-reported outcomes be used instead of clinician-reported outcomes and photographs as primary endpoints of late normal tissue effects in breast radiotherapy trials? Results from the IMPORT LOW trial. Radiother Oncol 134:220-230, 2019

Bhattacharya IS, Haviland JS, Kirby AM, Kirwan CC, Hopwood P, Yarnold JR, Bliss JM, Coles CE; IMPORT Trialists. Patient-Reported Outcomes Over 5 Years After Whole- or Partial-Breast Radiotherapy: Longitudinal Analysis of the IMPORT LOW (CRUK/06/003) Phase III Randomized Controlled Trial. J Clin Oncol:Jco1800982, 2018

Kirby AM, **Bhattacharya IS**, Wilcox M, Haviland JS. The IMPORT LOW Trial: Collaborative Research Accelerates Practice Change in Breast Radiotherapy. Clin Oncol (R Coll Radiol) 31:5-8, 2019

Batten LM, **Bhattacharya IS**, Moretti L, Haviland JS, Emson MA, Miller SE, Jefford M, MacKenzie M, Wilcox M, Hyslop M, Todd R, Snowdon CF, Bliss JM. Patient advocate involvement in the design and conduct of breast cancer clinical trials requiring the collection of multiple biopsies. Res Involv Engagem 4:22, 2018

Bhattacharya IS, Kirby AM, Bliss JM, Coles CE. Can Interrogation of Tumour Characteristics Lead us to Safely Omit Adjuvant Radiotherapy in Patients with Early Breast Cancer? Clin Oncol (R Coll Radiol) 30:158-165, 2018

Bhattacharya IS, Morden JP, Griffin C, Snowdon C, Brannan R, Bliss JM, Kilburn L. The Application and Feasibility of Using Routine Data Sources for Long-term Cancer Clinical Trial Follow-up. Clin Oncol (R Coll Radiol) 29:796-798, 2017

Coles CE, Griffin CL, Kirby AM, Titley J, Agrawal RK, Alhasso A, **Bhattacharya IS**, Brunt AM, Ciurlionis L, Chan C, Donovan EM, Emson MA, Harnett AN, Haviland JS, Hopwood P, Jefford ML, Kaggwa R, Sawyer EJ, Syndikus I, Tsang YM, Wheatley DA, Wilcox M, Yarnold JR, Bliss JM; IMPORT Trialists. Partial-breast radiotherapy after breast conservation surgery for patients with early breast cancer (UK IMPORT LOW trial): 5-year results from a multicentre, randomised, controlled, phase 3, non-inferiority trial. Lancet 390:1048-1060, 2017

Acknowledgements

I would like to gratefully acknowledge:

Dr Anna Kirby, my primary supervisor for all her guidance and support as well as steering me in the right direction over the last 3 years.

Professor Judith Bliss, my associate supervisor similarly for her guidance, support and the introduction to the world that is clinical trials.

Professor John Yarnold, my back-up supervisor for his mentorship.

Dr Charlotte Coles, chief investigator of the IMPORT and PRIMETIME trials for all her time, manuscript reviewing and enthusiasm.

Mrs Joanne Haviland, senior statistician for her endless help and statistical support.

Mr James Morden, senior statistician and my desk-buddy for his training and patience whilst I got to grips with STATA, James you are so missed by us all.

Dr Penny Hopwood, expert in psychosocial oncology for sharing her knowledge and experience with me.

Dr Sarah Gulliford and Dr Emma Harris for their radiobiology advice and training in using VODCA dosimetry software.

Dr David Eaton and RTTQA for all their help in acquiring CT planning data from the IMPORT HIGH trial.

Members of the ICR-CTSU IMPORT and PRIMETIME teams including Jenny, Ronald, Mark, Ada and Lisa (to name a few) with whom it has been a pleasure to work with and from whom I have learned so much.

The patients who participated in the IMPORT and PRIMETIME trials without whom this thesis would not have been possible.

Miss Freya Martin for taking the time out of her precious medical school summer holidays to proofread this thesis.

Mr Daniel Vinson for helping me with getting to grips with excel and all of his support.

Finally I would like to thank my Mum, Dad and brother for their endless support and for putting up with me.

I also acknowledge CRUK and funding from the CRUK Clinical Trials Fellowship for facilitating this fellowship and PhD.

Table of Contents

Summary of Figures

Summary of Tables

Chapter 1 Introduction

In all fields of medicine, rigorously conducted clinical trials are required to identify optimal treatment approaches for patients and the field of breast cancer is no exception. Data from the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) analysis of over >10 000 patients randomised into trials of breast conserving surgery (BCS) with and without radiotherapy have shown radiotherapy to the conserved breast reduces the risk of recurrence by one-half and breast cancer mortality by one-sixth [1]. The results of these trials have led to adjuvant radiotherapy following BCS becoming a standard of care in patients with early breast cancer (EBC) [1]. Furthermore, trials of hypofractionated whole-breast radiotherapy following BCS have demonstrated that 40 Gray (Gy) delivered in 15 fractions is safe, effective and gentler on normal tissues compared with the conventional 50Gy in 2Gy daily fractions [2,3]. Consequently, the three-week regimen tested in the START-B trial [2] has become the UK standard of care for whole-breast radiotherapy and is used increasingly in many other countries [4].

The primary efficacy endpoint in many breast radiotherapy trials, including those within the above-mentioned EBCTCG meta-analysis, has been local recurrence [1]. However, improvements in breast cancer outcomes mean that local recurrence rates have fallen substantially over recent decades. Earlier cancer detection, improvements in the quality and standardisation of surgery, developments in systemic therapies and radiation techniques have all contributed to the reduced rates of local relapse [5]. Although the reduction in local relapse rates is excellent for patients it has created a challenge in EBC radiotherapy trials. The low local recurrence rate means that clinical trials in EBC radiotherapy require large sample sizes in order to determine clinically relevant differences between treatments. Furthermore, given that many local recurrences occur beyond the 5 year time point [6] trials require prolonged follow-up and the consequent financial and resource burdens on both hospitals and clinical trials units need to be considered.

Radiotherapy is also not without risk. The toxicities from radiotherapy are usually assessed as secondary endpoints in breast radiotherapy trials. Commonly,

patients can develop late normal tissue effects (NTE) affecting the treated breast. These NTE include moderate /severe breast shrinkage, pain, tenderness and hardness [2] which can result in impaired quality of life and psychological distress [7]. Rarely, radiotherapy can also be associated with cardiac toxicity [8] and the development of second cancers [9]. Historically, NTE assessments have been primarily clinician based. However, with an increasing emphasis on patient-centred care, patient reported outcomes (PRO) are adopting a more prominent role as important endpoints in cancer clinical trials [10-12] and specifically in breast radiotherapy trials [13,14]. It may be possible to further optimise the role of PRO and even use PRO as primary NTE endpoints in breast radiotherapy trials. Furthermore, if PRO are to be used as primary endpoints, the manner in which PRO change over time, as well as if baseline factors influence PRO reporting needs to be better understood.

Given the low local relapse rates and possible risks of radiotherapy, there exists the potential for overtreatment of patients and this has led to clinical trials evaluating the concept of de-escalation of radiotherapy treatment for selected groups of patients. These trials include the IMPORT LOW partial breast radiotherapy trial [15] and the PRIMETIME avoidance of radiotherapy study [16]. Clinical trials are conducted when there is uncertainty regarding the optimal treatment option and all patients will have a degree of uncertainty with regards to participating in a clinical trial. However, this uncertainty may be increased for patients in de-escalation studies where a component of standard of care is removed. In these circumstances, there is the opportunity to optimise the decision-making process by improving the information provided to patients, for example, through the introduction of a patient decision aid. The overarching aim of this thesis is to explore strategies to optimise selected trial methodologies in order to maximise the efficiency of breast radiotherapy trials through a series of exploratory analyses alongside the development of a study within an existing trial.

1.1 Theme 1 Optimising collection and evaluation of adverse/ normal tissue effect data in breast radiotherapy trials

1.1.1 Importance of normal tissue effect detection

The low local relapse rates following adjuvant breast radiotherapy [2,15] mean that the absolute gains of radiotherapy are now much smaller, such that patients, clinicians and trialists can give greater consideration to adverse effects of treatment [17]. In order to accurately determine the risks of adverse events, detailed collection of normal tissue effect (NTE) data is required in breast radiotherapy trials. Furthermore, with improvements in breast radiotherapy techniques, including the introduction of intensity-modulated [18] and partial-breast radiotherapy [15], the NTE event rate has fallen substantially. It is therefore important that methods of detecting NTE are sufficiently sensitive.

1.1.2 Methods of detecting normal tissue effects

NTE have been variously assessed in breast radiotherapy trials using patient-reported outcomes (PRO), clinician-reported outcomes (CRO) and photographs [2,18]. PRO are defined as 'any clinical outcome that is reported directly by the patient and can be captured either though self-report or interview as long as the review directly records the patient's responses' [10 19 20]. PRO enable us to record patient perceptions of the impact of their cancer and the consequences of treatment [21]. Historically, standards for PRO inclusion in clinical trials were not available. Revised guidance by the Food and Drug Agency in 2009 resulted in a positive shift in how PRO were reported in clinical trials [22]. Firstly, it was required that when PRO endpoints were used in clinical trials, these were selected and interpreted appropriately based on a priori hypotheses regarding treatment outcome [22]. Secondly, the content validity of the PRO endpoint, which is the extent to which an instrument measures a concept, needs to be clearly explained [22]. Thirdly, the development history of a PRO must be documented to provide evidence that the patient-reported outcome measure (PROM) adequately measures what it claims to measure [22]. This includes questionnaire item design and modification, measurement properties such as reliability, validity and responsiveness, and how well patients understand questionnaire items and answer questions [22]. The appropriateness of the PROM to the patient group

must also be considered [22]. Finally, data collection methods (e.g. written versus electronic) and details regarding the timeline over which the patient is asked to answer questions should also be described [22].

There has been a rapid increase in the use of PROMs in oncology with various tools available [21]. These range from 'multidimensional measures of a patient's global perception of their health to specific tools that assess the severity of symptoms' [21,23,24]. For example, generic PROMs such as the Health Related Quality of Life (EQ5D), measure the 'patient's perceptions and societal values of the impact of disease or treatment' [21]. Generic tools measure emotional, physical and social functioning, but can be strongly influenced by environmental factors [21,24]. Although these generic measures provide population based data which can be used in trial comparisons and enable health economics analyses, they have a low sensitivity to change at an individual patient level [21]. In contrast, disease-specific PROMs which include the European Organisation for Research and Treatment of Cancer (EORTC) general cancer scale QLQ-C30 [25] ask patients about condition-specific problems with disease-specific attributes (including tumour-specific subscales e.g. QLQ-BR23 breast-cancer specific module [26]) as well as more specific symptom scores such as anxiety and depression via the Hospital Anxiety and Depression subscales (HADS) [27], and body image from the Body Image Scale (BIS) [21,28]. These disease-specific PROMs tend to have a higher patient sensitivity on an individual patient level, but less sensitivity on a population level [21,29]. In general, PROMs including both global and disease-specific concepts are used [21].

Examples of PROMs assessed in breast cancer trials include the EORTC general cancer scale QLQ-C30 and QLQ-BR23 breast-cancer specific module [25,26], HADS subscales [27], Body Image Scale [28], Breast Cancer Treatment Outcomes Scale (BCTOS) [7] and the FACT-B questionnaires [30]. A standard set of PROMs in breast cancer primarily based on the EORTC questionnaires has also been proposed by the International Consortium of Health Outcome Measures [31]. However, the PROMs described above are not specifically designed to capture NTE post-radiotherapy. In order to capture NTE post-radiotherapy, a series of questionnaire items were developed by the START trialists [13] to assess specific post-radiotherapy NTE not included in the EORTC

questionnaires. These were named the protocol-specific items (table 1.1) [13]. These questionnaire items were then added to by the patient advocate members of the IMPORT Trial Management Group (IMPORT trials described below). As the same set of trialists have run the START and IMPORT trials (as well as other breast radiotherapy trials including FAST [32] and FAST FORWARD [33]) this has meant that many of the data collection tools and processes have been consistent and have been refined over time building on previous experiences. For example, patient advocates in the IMPORT trial management group suggested adding questions to the protocol-specific questionnaire items regarding bra fitting and nipple position as patients advised that these outcomes were important to them following radiotherapy.

Table 1.1: Summary of protocol-specific questionnaire items

Patients asked to score:
• Change in breast appearance • Breast hardness/ firmness • Reduction in size of breast • Change in skin appearance • Is the position of the nipple of your affected breast different from the other side? • Problem getting a bra to fit • Shoulder stiffness

The protocol-specific questionnaire items are scored on a four-point scale: none, a little, quite a bit, very much (interpreted as none, mild, moderate, marked). In the START trials, patients were asked to score changes in the context of 'their radiotherapy treatment'. However, it is not possible for patients to determine whether changes to their breast were definitely due to radiotherapy, as surgical and systemic therapy effects may have contributed to any changes. In order to address this, in the IMPORT trials, the wording of the questions was changed to ask patients to assess changes in the context of 'any prior breast cancer treatments'.

The issues associated with using multiple PROMs in a clinical trial should be considered. When multiple questionnaire tools are used, they may contain overlapping questions and differences in scoring systems may result in differences in PROMs results within the same study [34]. It is also a burden for the

patient having to complete multiple questionnaire items, especially with overlapping questions. Furthermore, the lack of standardisation of PROMs makes it difficult to interpret PROMs results and compare results across different trials [10].

With respect to CRO, various scoring systems including the Common Terminology Criteria for Adverse Events (CTCAE v3) [35], Radiation Therapy Oncology Group/ European Organisation for Research and Treatment of Cancer Late Radiation Morbidity Scoring Schema (RTOG) [36] and overall cosmetic outcome scores have been used in breast radiotherapy trials [37] [38]. One of the challenges is that different scoring systems grade NTE using different criteria and contexts. For example, skin atrophy, skin pigmentation and breast induration may be assessed using the 'dermatology/skin' section of the CTCAE [35] albeit that the CTCAE is not radiotherapy-specific. The RTOG scoring system, on the other hand, captures NTE in relation to induration and telangiectasia, but does not capture breast shrinkage or oedema [36]. In the START trials, a series of CRO were established including breast shrinkage, breast induration, telangiectasia and breast oedema scored using the contralateral breast as a comparator with a four-point graded scale (none, a little, quite a bit, very much; interpreted as none, mild, moderate, marked) [39]. These CRO have since been used in a number of UK breast radiotherapy trials including the FAST [32] and FAST-FORWARD [33] trials as well as the IMPORT LOW [15] and HIGH [40] trials. As mentioned earlier, these trials have been run by the same group of trialists enabling a refinement of data collected. For example, building on the work from the START and IMPORT LOW trials, questions regarding breast discomfort and breast tenderness on palpation were added to the CRO in the IMPORT HIGH trial [40].

Cosmetic outcome has been assessed both by clinicians as individuals and through panel assessment of photographic data. Global measures of overall cosmesis include the Harvard criteria [37] and the EORTC cosmetic rating for breast cancer [38]. The Harvard criteria incorporate 1) fibrosis and retraction of the breast, 2) skin changes and 3) a matchline effect (defined as a localised area of fibrosis and skin change at the matchline between adjacent radiation fields) creating an overall cosmesis score assessed on a 4-point scale of 'excellent,

good, fair or poor' [37]. The EORTC cosmetic rating for breast cancer expands on the Harvard criteria to include appearance of the surgical scar, breast size, breast shape, skin colour, location of the areola and nipple and shape of the areola and nipple, with a similar scoring system [38]. However, it should be acknowledged that cosmesis is an aesthetic judgement and, as such, it could be argued that it is only valid when scored by patients rather than external observers [41,42]. An alternative to 'overall cosmesis' is the assessment of change in breast appearance using a panel consensus scoring method which was established in a pilot study in the START trials [43] and has since been used in several breast radiotherapy trials including the FAST [32], FAST-FORWARD [33] and the IMPORT [15 40] trials. Frontal photographs of both breasts are taken post-surgery and pre-radiotherapy (baseline) and are repeated at specific time-points post-radiotherapy. Change in appearance of the ipsilateral breast is compared with the baseline photograph and scored on a 3-point graded scale (none/minimal=0, mild=1, marked=2) based on a change in breast size, shrinkage and distortion. Software programmes have also been developed which combine various measurements of both breasts with objective scores of skin colour and scar appearance resulting in an overall cosmetic score [44-46].

The optimal NTE data collection method is unclear and there is no gold standard. The methodology of each assessment type differs, with each method asking a different question. For example, patients are asked to assess changes in their treated breast since their breast cancer treatment, whereas clinicians compare the patient's treated and contralateral breasts at a specific hospital visit. Also, patients score changes in the treated breast in the context of previous breast cancer treatment whereas clinicians score NTE based on their prior experience of seeing NTE from a range of patients or photographs, and they are specifically looking to identify radiotherapy-related NTE. Furthermore, when clinicians score photographs they have no knowledge of the patient's history, whereas a patient would be aware of any post-treatment complications they experienced. In addition, the scales used for scoring the different assessment methods may vary.

Irrespective of differences between the methods, the priorities for breast radiotherapy trials are that the method used to detect NTE should be able to

differentiate between randomised treatment groups (if a difference exists), and that the information obtained should be clinically relevant to patients. The protocol-specific PROMs used in the START trials were able to differentiate between small total dose differences in normofractionated versus hypofractionated regimens [13]. As mentioned above, these protocol-specific items were then added to by the IMPORT trialists and patient advocates. In particular, the two questions added were 1) Do you have a problem getting a bra to fit? and 2) Is the position of the nipple of your affected breast different from the other side? Both questions are highly relevant and have practical outcomes for the patient. The IMPORT LOW trial (described below) also found that certain protocol-specific PROMs were able to differentiate dose/volume regimens [15]. As well as being able to differentiate between small differences in dose-fractionation regimens and between dose/volume regimens, PRO also provide the patients' perceptions of the impact of their cancer and the consequences of treatment [21] within the context of the question asked. This raises the question as to whether PRO can be used as the primary NTE endpoints in future breast radiotherapy trials.

The START and Cambridge IMRT breast radiotherapy trials have investigated concordance between PRO, CRO and photographic assessments [47,48]. Both trials found concordance between PRO and CRO and photographic assessment on an individual patient basis was generally poor. This is not surprising as the methodology of each assessment type differs as described earlier [47,48]. The START trials showed that patients reported more NTE compared with clinicians and photographs. However, in the Cambridge IMRT trial clinicians reported more NTE than patients [47,48]. Conducting analyses to further investigate whether PRO can be used as primary NTE endpoints may build on this existing divergent literature.

The IMPORT LOW (ISRCTN12852634) trial provides a vehicle in which to investigate whether PRO can be used as primary NTE endpoints. IMPORT LOW is a multicentre, randomised, controlled, phase III trial investigating whether partial-breast radiotherapy is non-inferior to whole-breast radiotherapy in women found to have low-risk breast cancer following BCS (IMPORT LOW Appendix). Patients were randomly assigned (1:1:1) to receive 40Gy whole-breast

radiotherapy (control), 36Gy whole-breast radiotherapy and 40Gy to the partial-breast (reduced-dose group), or 40Gy to the partial-breast only (partial-breast group) in 15 daily treatment fractions. The IMPORT LOW trial included a comprehensive collection of NTE data with PRO and photographic sub-studies conducted in a subset of patients and CRO in all patients. IMPORT LOW reached its accrual of 2018 patients in 2010. The trial had reached 5 years of follow-up and was due to report the primary endpoint analysis at the beginning of IB's fellowship. This enabled IB to use the data which was collected from IMPORT LOW in a series of exploratory analyses.

1.1.3 PRO as primary NTE endpoints in breast radiotherapy trials

The NTE data in IMPORT LOW will be used to determine whether PRO can be used as primary NTE endpoints in breast radiotherapy trials. A series of analyses will be conducted to investigate this. Firstly, the degree of concordance on an individual patient level between PRO and CRO or photographs will be assessed. Secondly, whether results for the randomised comparisons obtained from PRO are consistent with those using CRO or photographs will be examined. Thirdly, whether there is any influence of baseline characteristics (including patient, tumour or treatment factors) on concordance will be investigated. **These analyses and their relevance to whether or not PRO can be used as primary endpoints in breast radiotherapy trials will be discussed in Chapter 2.**

1.2 Theme 2 Characteristics of PRO and the factors that influence them

1.2.1 Trajectory of PRO over time and factors influencing PRO

If PRO are to be used as primary NTE endpoints we must also understand the trajectory of how PRO change over time following radiotherapy. Previous whole-breast radiotherapy trials demonstrate that in general many NTE following radiotherapy reduce over time [13] [14]. For example, the START-B trial reported reduction in breast symptoms assessed using the QLQ-BR23 subscale over 5 years following radiotherapy in both standard and hypofractionation groups [13]. Similarly, the Cambridge IMRT trial [14] reported a reduction in patient-reported

adverse effects (AE) over the same period. On the other hand, the Cambridge IMRT trial reported a non-significant increase in breast shrinkage over time [14]. In the START-B trial, patients reported significantly less breast shrinkage in the hypofractionation group compared with the control group [13].

Whether baseline characteristics including patient, tumour and treatment factors can predict patient-reported NTE over time is also important. The Cambridge IMRT trial showed that younger age was associated with increased rates of patient-reported skin changes and breast hardness [14]. They also found that larger breast size was a predictor of patient-reported NTE. Furthermore, they reported that poor baseline surgical cosmesis predicted for increased skin changes and breast hardness. It has also been reported that pre-treatment psychological status may affect perception of cosmetic outcome from BCS and radiotherapy [49].

Using PRO data from the IMPORT LOW trial, a longitudinal analysis will be conducted with the main objective of determining whether breast cancer treatment-related adverse effects (including both NTE and effects pertaining to the patient's body image) improve, persist or worsen over time and, in addition, whether baseline patient, tumour and treatment specific factors influence patterns of patient AE reporting during the 5-year period. **The longitudinal analysis of IMPORT LOW PRO and baseline predictors of PRO reporting are reported in Chapter 3.**

1.2.2 Association of breast seroma and patient-reported NTE following breast radiotherapy

Building on the identification of baseline patient, tumour and treatment factors which predict NTE, the specific association of NTE with post-operative tumour bed seroma may also be investigated. Seroma formation describes the collection of serous fluid within a cavity and has been reported following BCS. Surgical technique may influence seroma formation. Seroma may be more likely to form when the excision cavity is left open compared with primary closure of the excision cavity either by direct suturing of cavity walls together, local

glandular mobilisation or therapeutic mammoplasty. In patients who develop a seroma in an open cavity, fibrosis and retraction of tissue surrounding the excision cavity (following seroma reabsorption) could result in a noticeable defect [50].

A seroma prevalence of 37% and 57% was reported in the Cambridge IMRT [51] and FAST [52] trials respectively. Seroma has been associated with increased rates of post-operative infection and haematoma, and is an independent risk factor for NTE following radiotherapy [51]. An association between seroma and NTE has been reported in the RAPID [53] and Cambridge IMRT trials [51]. The mechanisms by which seroma may lead to NTE following radiotherapy are unknown. As well as fibrosis and retraction of the seroma cavity being possible contributing factors [50], seroma leading to larger volumes receiving radiotherapy boost doses should also be considered. In the EORTC 'boost versus no boost' trial there was an increased risk of fibrosis in those patients receiving a boost [54] and this risk was further increased in patients with a seroma. However, this was significant on univariate analysis only.

The majority of these trials used clinician assessments of NTE and/ or serial photographs. However, the association between the presence of seroma and *patient*-reported NTE following breast radiotherapy has not been investigated to date. As previously discussed, PRO provide the patient's experience of the NTE following breast radiotherapy. Also, it is known from the START trials that patients reported more NTE than clinicians or panel assessment of photographs [47]. Therefore, if PROMs are not used NTE could potentially be underestimated.

It was initially planned to investigate the association between seroma and patient-reported NTE within the context of the IMPORT LOW partial-breast radiotherapy trial. However, data regarding smoking was not collected in IMPORT LOW as it was not known to be a risk-factor for NTE at the time of trial set-up. Smoking is an important variable to adjust for in an analysis investigating the association between seroma and NTE (as it is a known predictor of NTE) [53]. In contrast, smoking data were collected in the IMPORT HIGH trial which was set-up after IMPORT LOW. Also, the 3 year IMPORT HIGH toxicity data were available at the time of this analysis. As such, it was planned to investigate the

association between seroma and patient-reported NTE within the IMPORT HIGH trial.

The IMPORT HIGH trial provides an opportunity for the exploratory analysis of whether seroma is associated with patient-reported NTE. IMPORT HIGH is a randomised, multi-centre, phase III trial, testing dose escalated simultaneous integrated boost (SIB) against sequential boost each delivered by IMRT in patients with EBC and a higher risk of local relapse (IMPORT HIGH appendix). Patients were randomly assigned (1:1:1) between 40Gy/15F to whole breast (WB) + 16Gy/8F sequential photon boost to tumour bed (40+16Gy), 36Gy/15F to WB, 40Gy to partial breast + 48Gy (48Gy) or + 53Gy (53Gy) in 15F SIB to tumour bed. CT planning scan data for all patients recruited into IMPORT HIGH were collected by the Radiotherapy Quality Assurance Team (RTTQA). Similar to IMPORT LOW, NTE data in IMPORT HIGH were collected using PRO, CRO and photographs. By using the CT planning scan data obtained by RTTQA as well as the PRO data, associations between breast seroma and patient-reported NTE can be investigated within IMPORT HIGH, whilst taking into account other baseline factors which may be associated with NTE. **The association between breast seroma and patient-reported NTE will be investigated within IMPORT HIGH and discussed in Chapter 4.**

1.3 Theme 3 Improving information delivery in trials of de-escalation of breast radiotherapy to reduce patient uncertainty

1.3.1 Rationale behind “avoidance of radiotherapy” trials

Given the reduction in local relapse rates over recent decades (described in the opening section of this thesis), the risk-benefit ratio of adjuvant radiotherapy has changed. Although the relative benefit of breast radiotherapy remains the same as when the trials on which the EBCTCG meta-analysis was based were conducted, the absolute benefit is much smaller by virtue of the decreased local relapse rate. In women at the lowest risk of local relapse the absolute benefits of radiotherapy will be so small that the increased risk of NTE cannot be justified.

The challenge is now to identify women at the very lowest risk of relapse and demonstrate, in this group, that radiotherapy can be safely omitted. In order to address this changing risk-benefit ratio and reduce overtreatment of patients there has been worldwide interest in evaluating the concept of de-escalation of radiotherapy.

1.3.2 Challenges of conducting de-escalation studies

De-escalation of treatment studies present a number of challenges. In particular, they can be challenging to set-up, conduct and recruit to. Patients may perceive that 'more is better' and clinicians may practice 'better safe than sorry' [55]. It has been found that patients often have quantitative misperceptions regarding adjuvant treatment, overestimating the risk of a negative outcome without treatment and overestimating the positive effect of treatment [56 57]. The 10-year analysis of the CALGB 9943 trial which randomised women aged 70 or over with stage 1 ER positive and tumour size ≤2cm to receive BCS and tamoxifen with or without radiotherapy, showed local recurrence rates were 2% (95% confidence interval 1–4%) and 9% (95% confidence interval 6–13%) for those who did and did not receive radiotherapy respectively. Despite the increase in local recurrence in patients not receiving radiotherapy, there was no improvement in overall survival or breast cancer specific deaths in those who received radiotherapy (67%; 95% confidence interval 62–72%) versus those who did not (66%; 95% confidence interval 61–71%) [58]. Also, several RCTs have demonstrated that the difference between the absolute benefits of radiotherapy plus tamoxifen compared with tamoxifen alone decreases substantially with increasing patient age [58-60]. Nonetheless, despite the fact that local recurrences can be salvaged with no detriment to survival, some patients may still prefer to receive radiotherapy. Patients may wish to reduce anxiety around the possibility of a local relapse and the disruption that management of that recurrence would entail, regardless of the small absolute magnitude of the risk reduction.

It is important that clinicians clearly communicate the risks and benefits of treatments to patients. In general, it is recommended that absolute risk rather than relative risk be presented [61]. Absolute risk refers to how likely an event will be in one group of patients (which is relevant for the patient in question). In

contrast, relative risk describes how more or less likely an event will be between two groups of patients [61]. Also, patients may be able to more accurately perceive risk when numerical values are used. Natural frequency formats (i.e. numerical values expressed as event rates in groups with and without the intervention) should be used where possible. Expressing probabilities as an event rate out of 100 or 1000 patients can improve patient understanding [62]. Conveying information regarding patient prognosis and side-effects of treatment clearly and effectively is essential to enable patients to make informed choices regarding adjuvant treatment options. Greater patient advocate involvement in the design of information materials for de-escalation trials may enable a more patient-centred approach and perhaps improve the quality of information provided to patients as part of these studies [63].

There may also be considerable financial pressures with regard to 'avoidance of treatment' studies. In countries with privatised medical healthcare systems there may be financial benefits for clinicians to opt for 'treatment' over 'avoidance of treatment' [55]. In the UK's National Health Service, hospitals are paid per fraction of radiotherapy delivered. Given these financial arrangements, it is important that UK trialists and clinicians engage with commissioners to ensure that 'de-escalation of treatment' is not seen to translate into loss of earnings. Commissioners need to be encouraged to support studies that ultimately result in much greater health service savings in terms of finance and toxicity.

1.3.3 De-escalation clinical trials design

The trial design used in treatment de-escalation studies requires careful consideration. Although randomised controlled trials (RCT) are the 'gold standard' trial design, RCT testing treatment versus no treatment can be a challenge to recruit to as patients and clinicians may have strong preferences regarding treatments. The PRIME study initially randomised women to receive or omit radiotherapy following BCS [64]. Patient accrual was challenging particularly as patients did not want to be randomised and the trial design was amended allowing non-randomised patients who requested no radiotherapy to be followed up within a cohort design thereby improving recruitment [65].

On the other hand, the use of a prospective cohort design in the context of radiotherapy de-escalation concentrates specifically on the need for radiotherapy in a population considered to be at a very low risk of recurrence. In this trial design, the observed event rate in one group of patients is compared with a pre-specified cut off whereas, for a RCT, the event rates of two groups of patients are compared. Although the sample size required is dependent on the size of difference we want to detect, one of the benefits of the prospective cohort design compared with an RCT is that it does not require as large a sample size of patients as only one group of patients is required rather than two, which may facilitate a more rapid accrual given fewer patients are needed.

1.3.4 Improved risk stratification and current biomarker directed de-escalation of radiotherapy studies

In order to identify patients at low risk of local relapse, basic clinico-pathological parameters including T1/N0 stage, oestrogen receptor-positivity (ER), low grade (1/2) and older patient age may broadly define a group of patients with an anticipated low 5 year local relapse rate without radiotherapy. The addition of modern molecular diagnostics including gene profiling and immunohistochemistry (IHC) may refine the estimation of relapse risk for individual patients.

A number of prospective biomarker-directed studies exploring the de-escalation of radiotherapy in EBC are currently recruiting in various countries. These include the PRIMETIME [16], LUMINA [66], IDEA [67] and PRECISION [68] studies, all of which have used a biomarker-directed prospective cohort design, whilst the EXPERT [69] trial has adopted a biomarker-directed RCT design [65] (table 1.2). These studies aim to generate evidence supporting de-escalation of adjuvant radiotherapy in populations of patients with such low risks of local relapse that the risks of radiotherapy outweigh the benefits. One of the challenges of these biomarker-directed de-escalation studies is that there is no international consensus regarding the level of local recurrence that would be acceptable to

clinicians and patients with de-escalation of radiotherapy. The risks and benefits of radiotherapy need to be weighed up for each patient to achieve an individualised treatment decision such that international consensus on this issue is unlikely. In the PRIMETIME study a threshold of an ipsilateral breast disease rate of ≤4% at 5 years for selective de-escalation of radiotherapy was set primarily by patient advocates in collaboration with breast cancer clinicians and trialists [16]. However, for the individual patient, the level of local recurrence that the patient is willing to accept will vary between individuals and for some this cut-off may be perceived to be high.

Table 1.2: Summary of biomarker directed 'avoidance of radiotherapy' studies

Study name (date opened)	Country of Origin	Study design	Eligibility Criteria Age	Margin requirement following breast conserving surgery	Eligibility Criteria T1, G1-2, ER/PR+ve HER2-ve, N-ve	Eligibility Criteria Additional	Anticipated ipsilateral recurrence rate	Expected recruitment (patient number)
PRIMETIME (May 2017)	United Kingdom	Prospective cohort	≥60*	≥1 mm microscopic, circumferential margins of normal tissue from invasive	✓	Ki-67 to determine IHC4+C	≤4% at 5 years	1500
LUMINA (July 2013)	Canada	Prospective cohort	>55	≥1mm microscopically clear resection margins for invasive disease and DCIS or no residual disease on re-excision	✓	IHC including ER/PR/HER2, Ki-67 to determine luminal A subtype	< 5% at 5 years < 10% at 10 years	500
IDEA (March 2015)	United States	Prospective cohort/ Single group assignment	50-69	Margins of excision ≥2mm	✓ also included G3	Oncotype-DX RS ≤ 18	<6% at 5 years	200
PRECISION (May 2016)	United States	Phase 2 prospective cohort	50-75	negative margins ("no ink on tumor") or re-excision showing no residual disease in the re-excision specimen	✓	PAM-50 (luminal A subtype, low-risk ROR)	<5% at 5 years	690
EXPERT (August 2017)	Australia and New Zealand	Randomised controlled trial	≥50	microscopically negative margins for invasive carcinoma and any associated DCIS (no cancer cells adjacent to any inked edge/surface of specimen) or re-excision showing no residual disease	✓	PAM-50 (luminal A subtype, ROR ≤60)	≤4% at 5 years	1170

** Younger patients are eligible if they are post-menopausal and have co-morbidities that imply a high risk of radiotherapy toxicity*

1.3.5 Effective and preference-sensitive care

Treatment options or decisions, also referred to as healthcare services, can be defined as either 'effective' or 'preference-sensitive' [70,71]. Effective care describes a clinical scenario where the best possible treatment option is clear to both patients and clinicians as the evidence of the benefits and risks are known and the benefits clearly outweigh the risks. In contrast, preference-sensitive care describes the scenario where the best possible treatment option is unknown due to either 1) inadequate evidence or 2) the decision being dependent on the patient's values i.e. the importance attributed by the patient to the benefits versus risks of the treatment in question.

In general, research studies are conducted in the setting of 'preference-sensitive' care where there is uncertainty for both patients and clinicians regarding the optimal treatment option. This uncertainty may be increased for patients in a de-escalation of treatment study where a component of standard treatment is omitted. The state of uncertainty regarding a course of action is known as decisional conflict.

1.3.6 Patient uncertainty and decisional conflict

Decisional conflict is characterised by uncertainty in selecting the best option due to 1) the uncertainty about benefits and risks, 2) the need to consider personal values about potential benefits versus risks and 3) anticipated regret over the positive aspects of rejected options [72] [73] [74]. Decisional conflict has been described as 'the simultaneous opposing tendencies within the individual to accept and reject a given course of action' [75].

Verbalising uncertainty is the primary characteristic of decisional conflict [72]. Other characteristics include voicing concerns about undesired outcomes, hesitation between choices, delayed decision making, questioning personal values and beliefs when attempting decision making, pre-occupation with the decision and finally demonstrating signs and symptoms of distress or tension [72] [76,77].

There are two main sources of decisional conflict [72]. Firstly, there is the 'inherent difficulty' of the choice regarding a treatment which has both risks and benefits.

Secondly, there are a number of modifiable factors which make an 'inherently difficult choice even more difficult' [77] [78] [79-84]. Some of these modifiable factors include the lack of knowledge regarding treatment options and potential outcomes, unrealistic expectations of treatment options, unclear values, unclear perceptions of others, social pressures, lack of support, skills or self-confidence in decision making and lack of other resources. It is possible to address many of these modifiable factors using interventions in the patient decision making process. One of these interventions is the patient decision aid.

1.3.7 Use of patient decision aids to manage patient uncertainty

Patient decision aids (PDA) have been defined as 'interventions designed to help people make specific and deliberative choices among options by providing information about the options and outcomes that is relevant to a person's health status' [85] [86] [87]. PDA help patients to understand the risks and benefits of treatment options, consider the value they place on the risk-benefit ratio and participate actively with clinicians in deciding treatment options. The International Patient Decision Aid Standards (IPDAS) guidelines state that a PDA should have a systematic development process, provision of information about options and probabilities, clarification of values, disclosure of conflicts of interest, a balanced presentation of options and use of plain language and information based on current evidence [88]. PDA can take on a number of formats including booklets, counselling, visual aids, videotapes, audiotapes and software programmes accessible in clinics or via the internet [89]. Benefits and risks of treatment are often conveyed in graphical format using diagrams [89]. However, the optimal PDA format remains unknown [90].

A Cochrane review of PDA for people facing health treatment or screening decisions [87] found high quality evidence that PDA improved patient knowledge regarding treatment options (as scored using study-specific questionnaires) as well as reducing their decisional conflict. There is also evidence that PDA encourage patients to take a more active role in decision making and improve the accuracy of patients' perceptions of risk. A meta-analysis of cancer-related decision aids for patients entering RCTs demonstrated that patients receiving decision aids had reduced decisional conflict [89].

The IBIS II trial investigated the use of a decision aid in a RCT of an aromatase inhibitor in two patient groups; patients at high risk of breast cancer (prevention group) and patients with Ductal Carcinoma In Situ [DCIS] (treatment group). The authors found there was no difference in their primary outcome of decisional conflict. However, patients who received the decision aid in the 'treatment' group had higher knowledge post-decision compared with patients who did not receive the decision aid. In the 'prevention' group patients who received the decision aid had lower decisional regret at follow-up compared with those who did not receive the decision aid [91]. A study involving women aged 70 or above with stage 1 breast cancer considering radiotherapy after lumpectomy found, after using a decision aid, that patients had a statistically significant reduction in decisional conflict, increased clarity over the benefits and risks of treatment, and improved general treatment knowledge [92].

1.3.8 Use of a patient decision aid to reduce decisional conflict regarding entry to a de-escalation of treatment study

Whether a PDA can reduce decisional conflict in patients considering treatment de-escalation can be investigated using a 'Study Within A Trial' or SWAT concept. The SWAT approach enables us to assess different ways of designing, conducting, analysing and evaluating studies through the conduct of research within research [93]. It allows researchers to investigate aspects of trial recruitment and resource provision for participants embedded within a larger trial, enabling important research questions to be answered in an efficient manner i.e. multiple questions answered from one group of patients and trialists. 'SWATs' were established as a concept by the All Ireland Hub for Trials Methodology Research in collaboration with the Medical Research Council Network of Hubs in the United Kingdom [94].

There is the opportunity to investigate whether a PDA can reduce decisional conflict within the PRIMETIME study. PRIMETIME is a biomarker directed prospective cohort study aiming to identify a group of breast cancer patients who can safely avoid adjuvant breast radiotherapy following BCS (PRIMETIME appendix). The biomarker IHC4+C (incorporating Ki-67) is used to determine the patient's recurrence risk. Patients found to be at very low risk are directed to

avoid radiotherapy, and patients at low, intermediate or high risk are directed to receive radiotherapy. The PRIMETIME SWAT is named the PRIMETIME Information Giving Study. **Whether the introduction of a PDA in addition to standard patient information can reduce decisional conflict in the PRIMETIME study is discussed in Chapters 5 and 6. Chapter 5 describes the method development for the PRIMETIME SWAT, and Chapter 6 describes the results of the PRIMETIME SWAT.**

1.4 Theme 4 Clinical trials from the other side: Lessons learned by a clinician venturing into a clinical trials unit

1.4.1 The CRUK Clinical trials fellowship

In the opening line of this thesis introduction I stated, 'In all fields of medicine, rigorously conducted clinical trials are required to identify optimal treatment approaches for patients'. As the number and complexity of clinical trials grows, extra clinician involvement and time is needed [95] to ensure appropriate trial conduct. However, there is a workforce crisis in clinical oncology, with a failure to recruit sufficient trainees to meet future consultant demands [96]. Furthermore, new consultants have reported that the clinical oncology training programme did not adequately prepare them for the research aspect of consultant posts [97]. There are also increasing resource demands in the NHS which make running clinical trials at sites more and more challenging.

Effective collaboration between clinicians recruiting to clinical trials and clinical trials units (CTU) is important to improve understanding of trials methodology and streamline the clinical trials development and implementation process. In order to improve this collaboration, Cancer Research UK (CRUK) have developed 'Clinical Trials Fellowships' which embed clinicians in training in CTU for 1-3 years, enabling Fellows to develop the skills required to successfully deliver clinical trials. As well as giving the Fellow the opportunity to develop these skills, the Fellow can provide clinical expertise and develop sub-studies as part of a wider effort to ensure trials deliver maximal outputs. IB was an early recipient of this scheme and was based in The Institute of Cancer Research's Clinical

Trials and Statistics Unit (ICR-CTSU) working within the team undertaking a portfolio of breast radiotherapy trials. **In Chapter 7, I discuss my experience and the important lessons learned during the CRUK Clinical Trials Fellowship.**

1.5 Data sources accessed during the CRUK fellowship and thesis

During the CRUK clinical trials fellowship in the ICR-CTSU, IB was able to work with data from various breast radiotherapy trials each in different stages of development and follow-up. These trials provided the vehicles for a number of exploratory analyses. For example, when IB first started in June 2016, the IMPORT trials were in the follow-up phase. The IMPORT LOW primary endpoint was being analysed and the 5 year patient-reported outcome (PRO) data was available. This enabled the exploratory PRO analyses detailed in Chapters 2 and 3, which complement the results of the IMPORT LOW primary endpoint analysis.

It was also planned that the analysis of whether seroma is associated with patient-reported NTE would be conducted within IMPORT LOW. However, given that smoking data were not available for IMPORT LOW, this was not the correct vehicle in which to conduct this analysis. In contrast, smoking data was available for IMPORT HIGH and so the seroma analysis was conducted in IMPORT HIGH. The timing of this exploratory analysis in IMPORT HIGH was adjusted in order to complement the reporting of the 3 year toxicity endpoint which was reported in December 2018. For all the exploratory analyses within the IMPORT trials it was important that these analyses were only conducted and data released **after** the main study endpoint data were publicly available so as not to compromise any aspect of reporting from the trial.

It is acknowledged that the exploratory analyses were conducted using pre-collected data from the IMPORT trials and IB was not involved in the concept development or data collection for the IMPORT trials whereas the development of the PRIMETIME SWAT investigating whether a PDA reduces decisional

conflict enabled IB to generate her own data. Although the PDA could have been tested within patient focus groups, the PRIMETIME study was due to open to recruitment soon after IB begun her fellowship providing an opportunity to test the PDA within a de-escalation study. Also, if IB was to test the PDA in a clinical trial, it would only have been feasible to do so using the SWAT concept as it would not have been possible to develop, set-up and recruit to a new trial in 3 years.

1.6 Framework of thesis

Chapter 2 investigates whether patient-reported outcomes (PRO) can be used as primary NTE endpoints in breast radiotherapy trials using data from the IMPORT LOW trial. Building on this, Chapter 3 determines how PRO change over time and whether baseline characteristics can predict patient-reported NTE, again using data from IMPORT LOW. Chapter 4 takes a different slant by using CT planning scan and dosimetric data to investigate whether seroma is associated with patient-reported NTE using data from the IMPORT HIGH trial. Whilst chapters 2-4 focus on exploratory analyses of NTE data within the IMPORT trials, Chapters 5 and 6 discuss whether the introduction of a PDA can reduce decisional conflict within the context of a de-escalation of radiotherapy study, the PRIMETIME study. In particular Chapter 5 focusses of the method development of the PRIMETIME SWAT whilst Chapter 6 focusses on the implementation and results of the PRIMETIME SWAT. Finally Chapter 7 describes IB's experience and the lessons learned during the CRUK fellowship whilst based at the ICR-CTSU.

Chapter 2 Can patient-reported outcomes be used as primary endpoints of late normal tissue effects in breast radiotherapy trials?

2.1 Abstract

Background: In an era of low local relapse rates after adjuvant breast radiotherapy, risks of late normal-tissue effects (NTE) need to be balanced against risk of relapse. NTE are assessed using patient-reported outcomes (PRO), clinician-reported outcomes (CRO) and photographs. This analysis investigates whether PRO can be used as primary NTE endpoints in breast radiotherapy trials.

Methods: Analyses were conducted within IMPORT LOW (ISRCTN12852634) using data at 2 and 5 years post-randomisation. NTE were recorded by CRO, photographs and PRO. Measures of agreement tested concordance, risk ratios for radiotherapy groups were compared, and influence of baseline characteristics on concordance investigated.

Results: In 1095 patients who consented to PRO and photographs, PRO were available at 2 and/or 5 years for 976 patients, of whom 909 had CRO and 844 had photographs. Few patients had moderate/marked NTE, irrespective of method used (e.g. 19% patients and 9% clinicians reported breast shrinkage at year 5). Patients reported NTE more frequently than assessed from CRO or photographs ($p<0.001$ for most NTE). Concordance between assessments was poor on an individual patient level; e.g. for year 5 breast shrinkage, % agreement=48% and weighted kappa=0.17. Risk ratios comparing radiotherapy schedules were consistent between PRO and CRO or photographs.

Conclusions: Few patients had moderate/marked NTE irrespective of the method used. Patients reported NTE more frequently than CRO and photographs, therefore NTE may be underestimated if PRO are not used. Despite poor concordance between methods, effect sizes from PRO were consistent with CRO and photographs, suggesting PRO can be used as primary NTE endpoints in breast radiotherapy trials.

2.2 Introduction

In the current era of low local relapse rates after adjuvant breast radiotherapy [2,15], the risks of radiotherapy-related late normal tissue effects (NTE) need to be carefully balanced against the benefits of treatment, requiring detailed collection of NTE data in breast radiotherapy trials. Furthermore, with improvements in breast radiotherapy techniques, including the introduction of intensity-modulated [18] and partial-breast radiotherapy [15], the NTE event rate has also fallen substantially. Consequently, measuring NTE is becoming increasingly challenging.

NTE have been variously assessed in breast radiotherapy trials using clinician-reported outcomes (CRO), photographs and patient-reported outcomes (PRO) [2,18]. The optimal NTE data collection method is unclear and there is no gold standard. The methodology of each assessment type differs. For example, patients may be asked to assess changes in their treated breast since their breast cancer treatment, whereas clinicians compare the patient's treated and contralateral breasts. Also, the scales used for scoring the different assessments vary.

Irrespective of differences between the methods, the priorities for breast radiotherapy trials are that the method used to detect NTE should be able to differentiate between randomised treatment groups (if a difference exists), and that the information obtained is clinically relevant to patients. Data from breast radiotherapy trials demonstrate that PRO are able to differentiate between dose/volume regimens [15] and between small dose differences in hypofractionated regimens [2,13]. PRO also provide the patients' perceptions of the impact of their cancer and the consequences of treatment [21] within the framework of the question asked. This analysis investigates within the context of the IMPORT LOW partial-breast radiotherapy trial, 1) the degree of concordance on an individual patient level between PRO and CRO or photographs, 2) whether results for the randomised comparisons obtained from PRO are consistent with those using CRO or photographs and 3) the influence of baseline characteristics on concordance, with the overall aim of assessing

whether PRO could be used as primary NTE endpoints in future breast radiotherapy trials.

2.3 Methods

2.3.1 Patient population

IMPORT LOW (ISRCTN12852634) is a multicentre randomised phase III non-inferiority trial comparing safety and efficacy of standard whole-breast radiotherapy with two experimental schedules (reduced-dose and partial-breast radiotherapy) in women with low-risk breast cancer after breast conserving surgery (IMPORT LOW appendix) [15].

IMPORT LOW included a comprehensive and systematic investigation of NTE including CRO in all participants, and PRO and photographs in a subset of patients [15]. All centres were invited to participate in the PRO and photographic sub-studies (until sufficient accrual was achieved). All patients at these centres were invited to participate in the sub-studies until the designated sample size for each sub-study was obtained.

2.3.2 Procedures

Patients who consented to the PRO sub-study completed the EORTC QLQ-C30 core questionnaire and QLQ-BR23 breast-specific module [25,26] and 10-item Body Image Scale (BIS) [28], all of which asked patients to consider their symptoms during the past week. The Hospital Anxiety and Depression Scale (HADS) [27] and protocol-specific questionnaire items relating to ‘change in breast appearance’, ‘breast hardness/firmness’, ‘reduction in size of breast’, ‘change in skin appearance’, ‘is the position of the nipple of your affected breast different from the other side’, ‘problem getting a bra to fit’ and ‘shoulder stiffness’ which may have resulted from any prior breast cancer treatments [13] were also completed. All items (with the exception of HADS) were scored on a four-point scale: none, a little, quite a bit, very much (interpreted as none, mild, moderate, marked). Questionnaires were completed at baseline (pre-radiotherapy) and 6 months, 1, 2 and 5 years after radiotherapy. Patients completed PRO sub-study questionnaires alone with no help from clinicians.

For patients participating in the photographic sub-study, photographs were taken at baseline (post-surgery but pre-radiotherapy), year 2 and year 5. Change in photographic breast appearance of the ipsilateral breast was assessed at 2 and 5 years compared with the baseline photograph. Breast size and surgical deficit were scored from the baseline photographs on a 3-point scale (small, medium, large). At 2 and 5 years after radiotherapy, breast appearance change (none/ mild/ marked) was scored on a pair of photographs (one with the patients' hands on the hips and one with hands raised) in comparison with the baseline photograph. A panel of observers blinded to patient identity, treatment allocation, and radiotherapy centre scored the photographs, the methodology having been validated in the START pilot trial [43].

CRO including breast shrinkage, breast induration, telangiectasia and breast oedema were scored using the contralateral breast as a comparator with a four-point graded scale (none, a little, quite a bit, very much; interpreted as none, mild, moderate, marked) at 1, 2 and 5 years following radiotherapy in all patients. The CRO items were established and validated in the START trials [39]. Clinicians were not blinded to treatment group.

2.3.3 Statistical analysis

PRO were paired with the relevant CRO or photograph at 2 and 5 years for the analyses (table 2.1).

Table 2.1: Patient reported outcome measures of specific late NTE in the breast and the corresponding clinician and photographic assessment

Patient Reported Outcome Measure *(resulting from prior breast cancer treatment)*	**Clinician Assessment** *(treated breast compared with contralateral breast)*	**Photographic Assessment** *(change in appearance compared with baseline photograph)*
Has your affected breast become smaller?	breast shrinkage	-
Has your affected breast become harder/firmer to the touch?	breast induration*	-
Was the area of your affected breast swollen?	breast oedema	-
Have you had a problem getting a bra to fit?	breast shrinkage	-
Has the overall appearance of your affected breast changed compared with the other side?	-	Overall change in breast appearance
Is the position of the nipple of your affected breast different from the other side?	-	Overall change in breast appearance

**maximum score in and outside the tumour bed was recorded*

The “quite a bit” and “very much” categories were combined for PRO and CRO as few NTE were scored as “very much”. This resulted in a 3-point scale corresponding to none, a little (mild), quite a bit/very much (moderate/ marked). This also enabled direct comparison with photos, also scored on a 3-point scale.

Agreement between the data ascertainment methods on an individual patient level was assessed using percentage agreement (with 95% confidence interval), weighted kappa statistic (with 95% confidence interval) and Bowker’s test of symmetry [98]. Guidelines for interpreting the value of weighted kappa in terms of the strength of agreement are <0.20: poor; 0.21-0.40: fair; 0.41-0.6: moderate; 0.61-0.8: good; 0.81-1.00: very good [99]. A significance level of ≤ 0.005 was used to account for multiple testing in all analyses.

Risk ratios comparing each test radiotherapy schedule with the control group were calculated for each NTE endpoint at year 5 and presented in forest plots for the different assessment methods. Results for breast oedema were not included in this comparison as so few events were reported using PRO and CRO at year 5.

The influence of baseline patient characteristics on concordance was investigated using stratified analyses, and formally assessed in logistic regression models defining a binary outcome as 1=concordant (same scores for PRO and CRO/ photographs) versus 0=discordant (different scores). Baseline factors found to be statistically significantly associated with concordance on univariate analysis were tested together on multivariate analysis. Baseline characteristics tested included age, treatment group, breast size and surgical deficit (assessed from baseline photographs), HADS anxiety and depression subscale scores and body image scores.

All analyses were carried out using STATA version 14 based on a database snapshot taken on June 15th 2016 (as per the primary endpoint analysis).

2.4 Results

2018 patients were recruited to IMPORT LOW from 71 centres. 2 patients requested exclusion from analysis. In the 41 centres participating in the PRO sub-study, 1265/1333 (95%) patients consented to PRO, and 1318/1466 (90%) patients consented to the photographic sub-study from 37 participating centres. 1095 patients consented to both sub-studies (figure 2.1).

In 1095 patients who consented to both, PRO were available at 2 and/ or 5 years for 976 patients of whom 909 had CRO and 844 had photographs. PRO, CRO and photographs were available for 651 and 518 patients at year 2 (figure 2.1) and year 5 respectively (figure 2.1). Separate analyses were conducted in patients with PRO and CRO, and PRO and photographs, at year 2 and year 5. Data regarding baseline characteristics were published in the IMPORT LOW primary endpoint analysis [15]. With respect to PRO questionnaire return rates, there was a high proportion of completed questionnaires (based on the number

returned versus expected, excluding patients who had died or withdrawn from the study) at both time points (~80%) [100]. Further details regarding questionnaire return rates are available in Chapter 3 which focusses primarily on the PRO in IMPORT LOW over 5 years.

2.4.1 Overall prevalence of NTE

The overall prevalence of patients with NTE was low, with most scored as none or mild by all three data ascertainment methods (table 2.2). Few patients had NTE scored as moderate or marked. NTE which were commonly reported included breast shrinkage, induration and breast appearance change. At year 5, 19% patients and 9% clinicians reported moderate/ marked breast shrinkage. With respect to breast induration, 7% patients and 5% clinicians reported moderate/ marked changes. For breast appearance change, 18% patients reported moderate/ marked changes and photographic assessment reported marked changes in 4%.

2.4.2 Reporting of NTE by patients versus either CRO or photographs

Patients reported a higher prevalence of breast changes than CRO and photographs for all NTE assessed, except for more clinically-reported mild breast shrinkage compared with patient-reported bra fitting at both time-points (figure 2.2 and 2.3). Patients and clinicians reported similar prevalences of breast oedema, with very few events at 2 and 5 years. Concordance between PRO and CRO or photographs of corresponding NTE on an individual patient basis was generally poor (table 2.2).

For breast shrinkage at year 5, patients reported effects more frequently than clinicians (figure 2.3); percentage agreement was 48% and concordance was poor as evidenced by the low weighted kappa (0.17, table 2.2). Bowker's test of symmetry was also highly significant ($p<0.001$) indicating discordance, with patients reporting effects more frequently than clinicians (table 2.2). With regard to 5 year breast appearance change, patients reported more NTE than scored on photographs (Bowker's test of symmetry <0.001, table 2.2). Agreement was poor (35%), as was concordance (weighted kappa 0.09, table 2.2).

In contrast, for breast induration at year 5, PRO and CRO appeared better aligned with similar levels of effects reported by both (figure 2.3) and a higher % agreement (61%, table 2.2), but concordance remained poor (weighted kappa 0.12, table 2.2). In addition, Bowker's test for symmetry was no longer significant (p=0.025), implying similar effects reported by PRO and CRO (table 2.2).

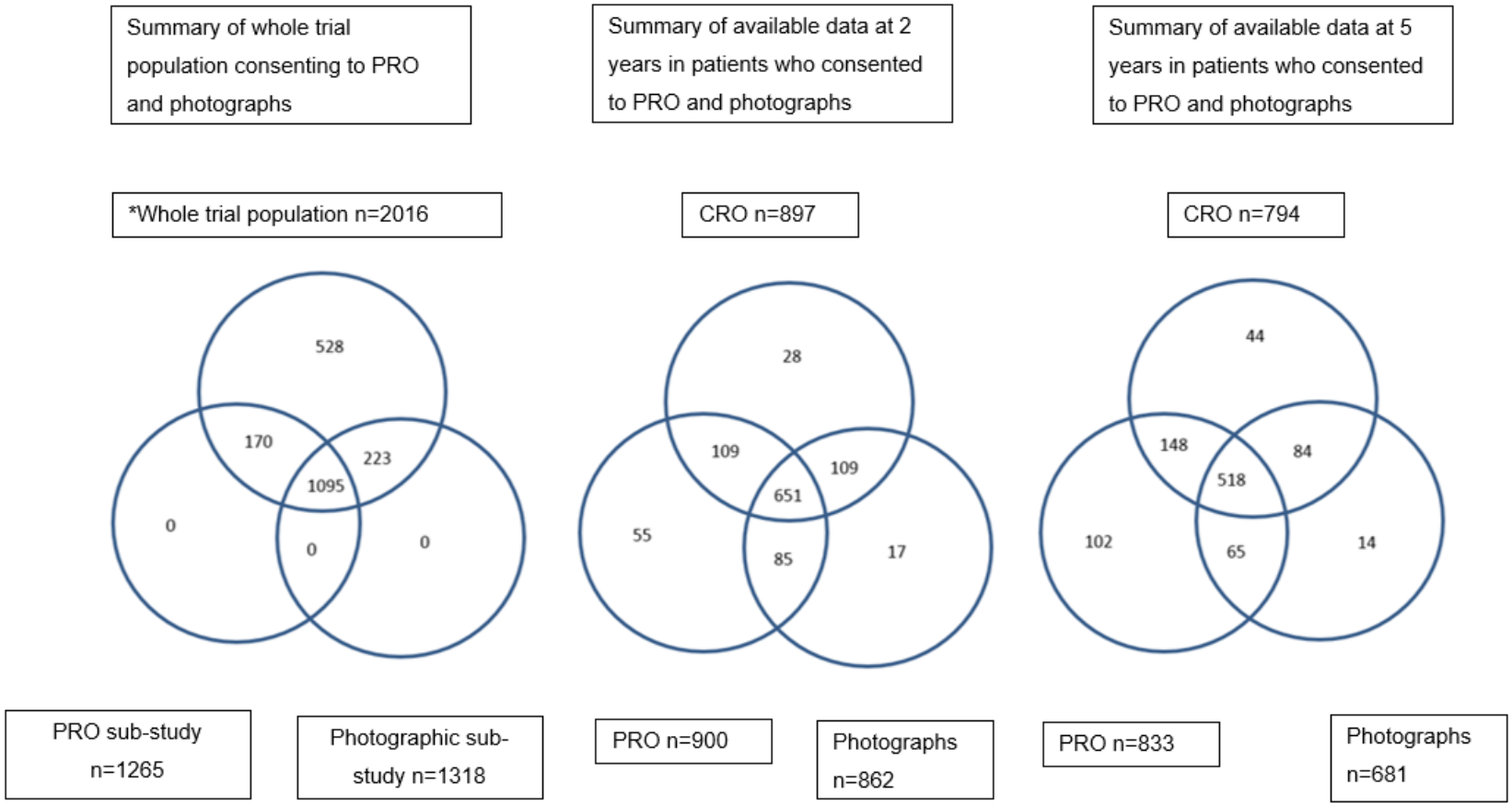
2.4.3 Comparison of radiotherapy schedules using PRO, CRO and photographs

On comparison of the risk ratios for the radiotherapy schedules, similar effect sizes were seen for breast shrinkage and breast appearance change when the analogous question was asked of the patient, or ascertained from either CRO or photographs (figure 2.4). There was some evidence of differing effect sizes between the assessment methods for breast induration, but the confidence intervals overlapped (figure 2.4).

2.4.4 Associations between baseline characteristics and concordance

On stratified analyses, there was little evidence that concordance varied according to baseline characteristics at 2 or 5 years (table 2.3 & 2.4). Some baseline factors were significantly associated with concordance of PRO and either CRO or photographs for certain NTE in logistic regression models, but predominantly on univariate analysis only and not across both time-points (table 2.5). For example, larger surgical deficit was associated with discordance of breast shrinkage at year 5 only [OR 0.32 (95%CI 0.16-0.65)] (table 2.5).

Figure 2.1: Summary of whole trial population consenting to PRO and photographs, and data available at 2 and 5 years



Two patients withdrew consent for any of their data to be used in the analysis

Table 2.2: Concordance between PRO and CRO and photographic assessments of specific NTE at 2 and 5 years in IMPORT LOW

Patient Reported Outcome	Clinician Reported Outcome/ Photograph			% agreement (95% confidence interval)	Weighted Kappa (95% confidence interval)	Bowker's test of symmetry, p value
	None	A little	Quite a bit/very much			
Breast smaller/ shrinkage – 2yrs						
None	276	63	6	400/860;	0.16 (0.16-0.20)	<0.001
A Little	250	105	23	46.50%		
Quite a bit/very much	62	56	19	(43.1-49.9%)		
Breast smaller/ shrinkage – 5yrs						
None	221	60	11	358/751;	0.17 (0.14-0.19)	<0.001
A Little	170	115	31	47.70%		
Quite a bit/very much	75	46	22	(44.0-51.3%)		
Breast harder/ induration – 2yrs						
None	432	87	15	493/860;	0.11 (0.09-0.12)	<0.001
A Little	202	51	15	57.30%		
Quite a bit/very much	34	14	10	(53.9-60.7%)		
Breast harder/ induration – 5yrs						
None	398	93	21	457/751;	0.12 (0.09-0.19)	0.025
A Little	126	53	10	60.90%		
Quite a bit/very much	32	12	6	(57.3-64.4%)		
Breast swollen /oedema – 2yrs						
None	741	44	5	750/854;	0.15 (0.10-0.18)	0.99
A Little	43	9	3	87.80%		
Quite a bit/very much	6	3	0	(85.4-89.9%)		
Breast swollen /oedema – 5yrs						
None	670	24	1	673/743;	0.05 (0.01-0.11)	0.06
A Little	39	3	1	90.60%		
Quite a bit/very much	5	0	0	(88.2-92.6%)		
PRO-Bra fitting/ CRO shrinkage - 2yrs						
None	464	161	22	504/860;	0.11 (0.06-0.13)	<0.001
A Little	98	33	18	58.60%		
Quite a bit/very much	26	31	7	(55.2-61.9)		
PRO-Bra fitting/ CRO shrinkage - 5yrs						
None	356	145	29	421/752;	0.15 (0.08-0.16)	<0.001
A Little	81	52	22	56.00%		
Quite a bit/very much	30	24	13	(52.4-59.6)		
Overall change in appearance* – 2yrs						
None	158	9	3	193/731;	0.03 (0.01-0.03)	<0.001
A Little	406	29	4	26.40%		
Quite a bit/very much	97	19	6	(23.2-29.8%)		
Overall change in appearance* – 5yrs						
None	138	15	2	199/571;	0.09 (0.05-0.14)	<0.001
A Little	262	48	6	34.90%		
Quite a bit/very much	60	27	13	(30.9-38.9%)		
Nipple position/change in appearance*-2yrs						
None	412	30	4	430/728;	0.04 (0.03-0.05)	<0.001
A Little	191	17	8	59.10%		
Quite a bit/very much	56	9	1	(55.4-62.7%)		
Nipple position/change in appearance*-5yrs						
None	279	48	10	314/569;	0.08 (0.03-0.11)	<0.001
A Little	142	28	4	55.20%		
Quite a bit/very much	37	14	7	(51.0-59.3%)		

**change in appearance assessed on photograph*

Figure 2.2: Comparison of year 2 PRO, CRO and photographic assessments of specific late NTE in IMPORT LOW

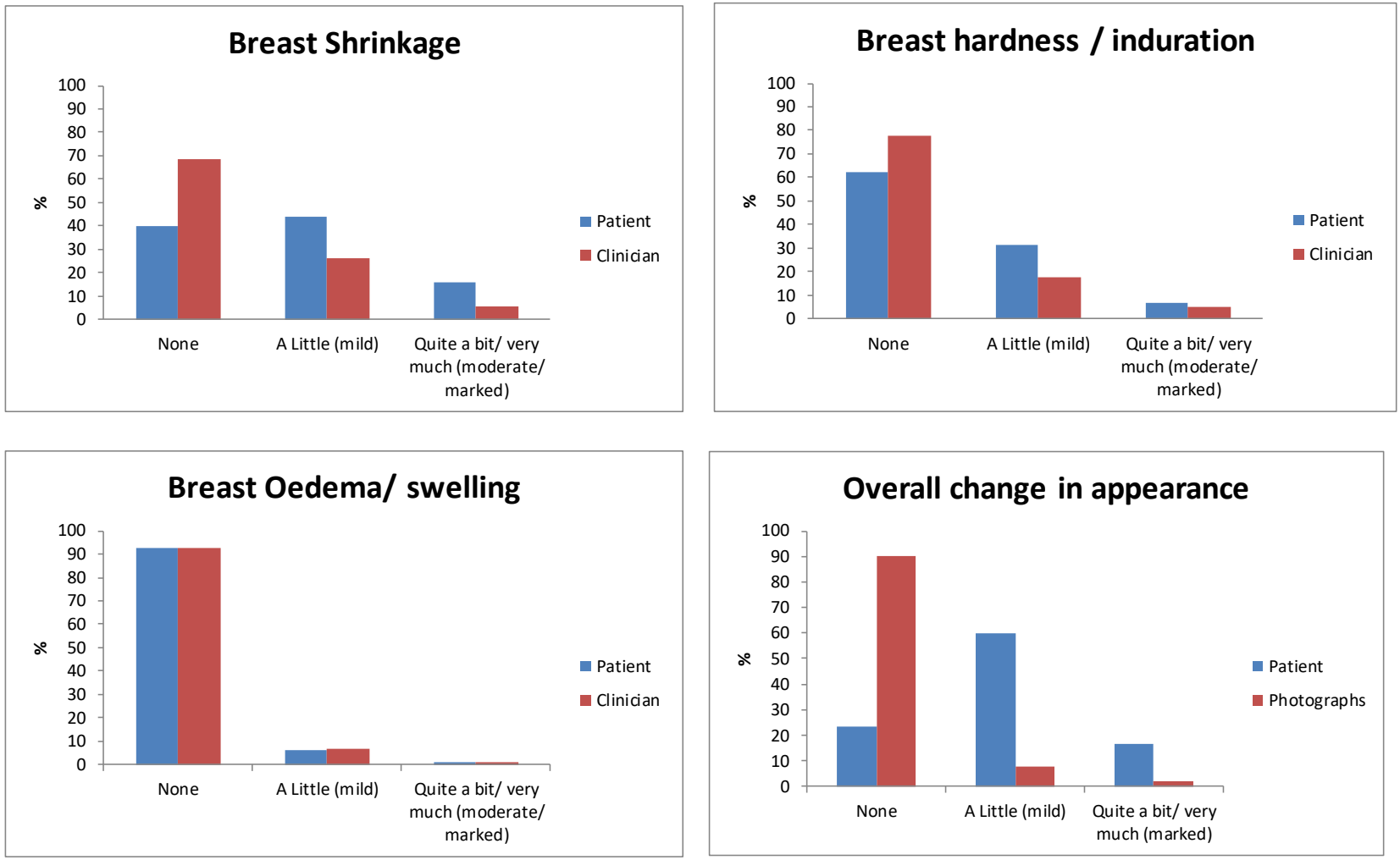
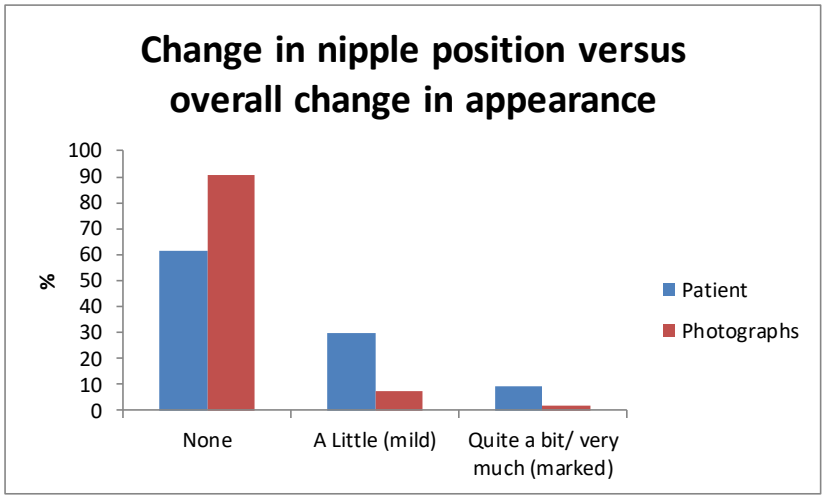


Figure 2.2 continued



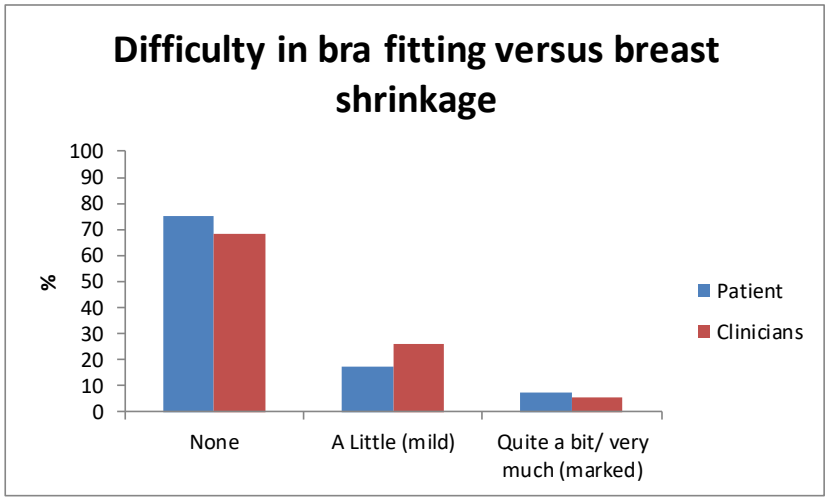


Figure 2.3: Comparison of year 5 PRO, CRO and photographic assessments of specific late NTE in IMPORT LOW

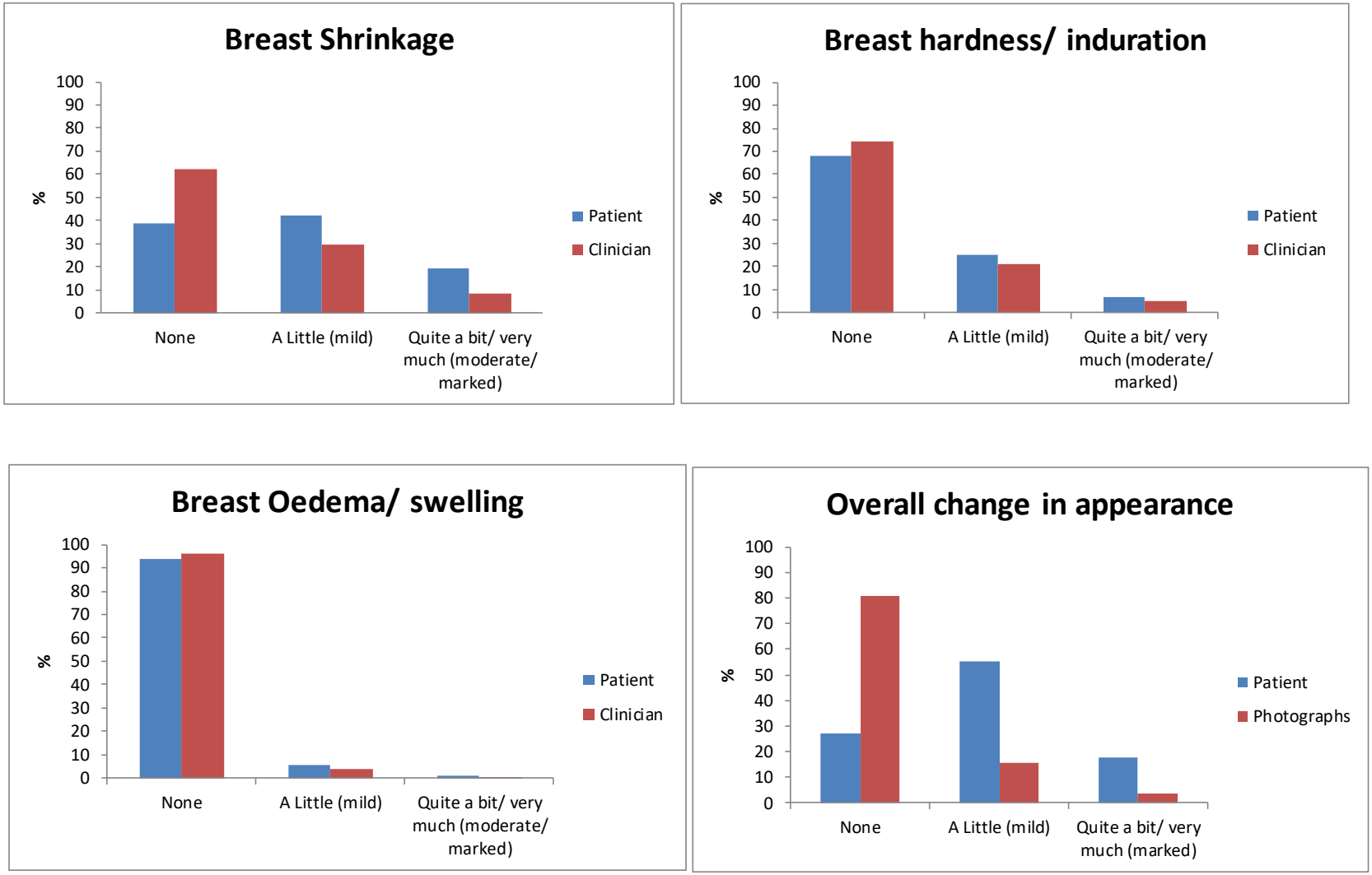
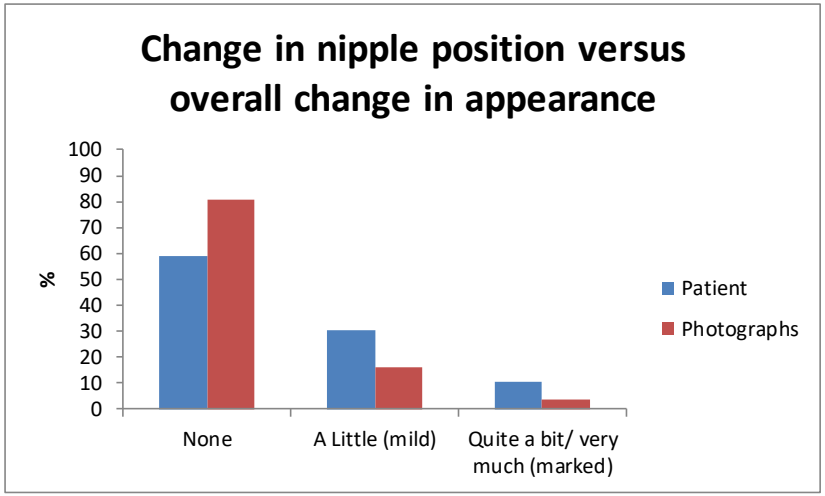


Figure 2.3 continued



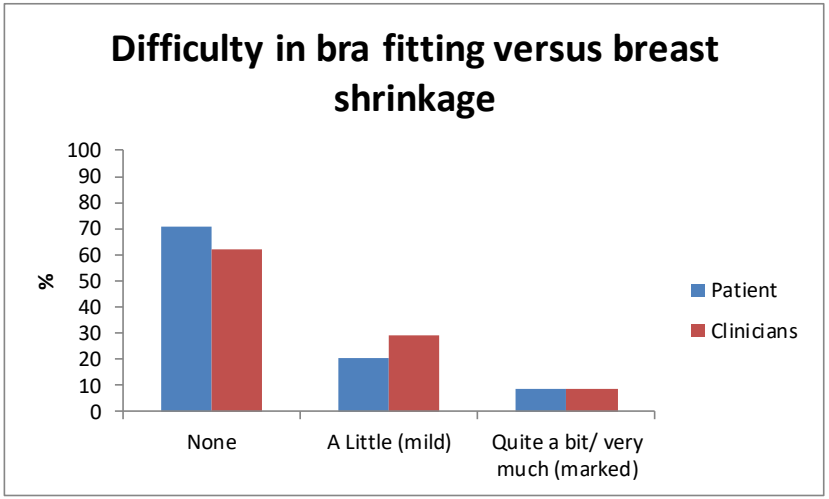


Figure 2.4: Comparison of the estimates of effect sizes for the randomised radiotherapy groups between PRO and CRO/ photographs at 5 years

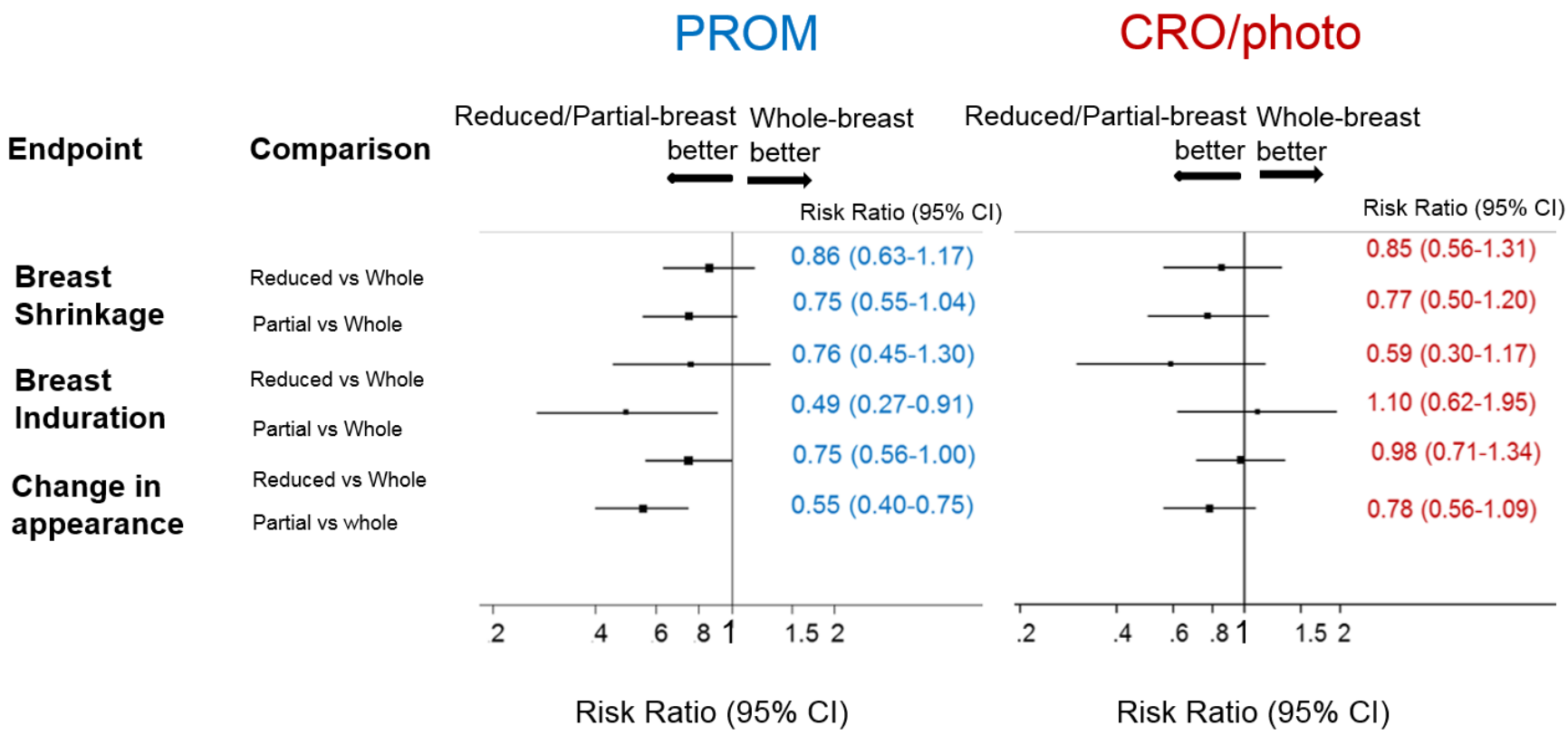


Table 2.3: Concordance between PRO and clinician and photographic assessments of specific NTE at year 2 stratified by baseline characteristics in the IMPORT LOW trial

Baseline item	Breast Shrinkage		Breast induration		Breast Swelling		Overall change in appearance		Nipple position		Bra fitting	
	%agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	%agreement (95%CI)	Weighted kappa (95% CI)
Age												
<60 years	43.7(37.9-49.6)	0.12 (0.03-0.17)	51.2(45.3-57.1)	0.06(0.02-0.06)	87.6(83.2-91.2)	0.07 (-0.001-0.11)	24.2(18.9-30.1)	0.03 (0.01-0.04)	61.5(55.0-67.7)	0.06(0.02-0.10)	57.5(51.6-63.3)	0.08 (0.001-0.14)
≥60 years	48.0 (43.8-52.2)	0.18 (0.14-0.21)	60.5 (56.3-64.5)	0.14 (0.09-0.18)	87.9(85.0-90.5)	0.19 (0.17-0.29)	27.5(23.6-31.7)	0.02 (0.02-0.03)	57.9(53.4-62.3)	0.04(-0.001-0.08)	59.2(55.0-63.2)	0.13 (0.10-0.16)
Treatment Group												
Group 1	49.1(43.1-55.1)	0.23 (0.18-0.35)	48.6(42.6-54.6)	0.03(0.01-0.12)	86.7(82.2-90.5)	0.11 (-0.03-0.19)	25.7 (20.3-31.8)	0.03 (0.02-0.06)	57.9 (51.3-64.3)	0.04 (0.03-0.07)	53.8(47.7-59.8)	0.09 (0.02-0.19)
Group 2	45.2 (39.5-51.1)	0.12 (0.07-0.14)	60.1(54.3-65.8)	0.16(0.08-0.26)	86.0(81.5-89.8)	0.12 (0.04-0.20)	26.7(21.3-32.7)	0.03 (0.002-0.04)	57.1(50.7-63.3)	0.05 (-0.009-0.11)	63.2(57.4-68.7)	0.11 (0.06-0.21)
Group 3	45.3(39.4-51.3)	0.14 (0.13-0.19)	63.0(57.1-68.7)	0.14(0.09-0.20)	90.8 (86.8-93.9)	0.25 (0.15-0.25)	26.7(21.3-32.7)	0.02 (0.01-0.04)	62.2(55.8-68.3)	0.03 (0.004-0.12)	58.5(52.6-64.3)	0.12 (0.09-0.17)
Breast Size												
Small	47.6 (41.9-53.3)	0.17 (0.14-0.18)	58.3(52.5-63.8)	0.05(0.03-0.12)	90.5 (86.6-93.5)	-0.05 (-0.06—0.04)	24.2 (19.6-29.2)	-0.003 (-0.02- -0)	55.6 (50.0-61.0)	0.01 (-0.004-0.03)	56.8 (51.1-62.4)	0.07 (0.05-0.13)
Medium	48.8 (42.2-55.2)	0.21 (0.19-0.27)	58.8(52.3-65.0)	0.09(0.06-0.15)	86.6 (81.6-90.7)	0.16 (0.07-0.24)	24.3 (19.2-30.1)	0.02 (-0.01-0.04)	63.5 (57.3-69.4)	0.10 (0.04-0.14)	60.5 (54.0-66.7)	0.15 (0.05-0.25)
Large	53.7 (44.9-62.3)	0.27 (0.27-0.37)	54.0(45.3-62.6)	0.11(0.05-0.16)	80.4 (72.8-86.7)	0.20 (0.12-0.31)	33.1 (25.4-41.5)	0.10 (0.06-0.14)	59.9 (51.3-68.0)	0.08 (0.08-0.15)	55.1 (46.4-63.7)	0.15 (0.13-0.27)
Surgical deficit												
Small	53.4 (48.7-58.1)	0.20 (0.16-0.23)	61.1(56.4-65.6)	0.10(0.07-0.21)	86.9 (83.5-89.9)	0.18 (0.14-0.30)	29.8 (25.7-34.1)	0.004 (-0.02-0.03)	66.9 (62.5-71.1)	0.02 (0.01-0.04)	66.0 (61.5-70.3)	0.15 (0.13-0.21)
Medium	41.9 (34.6-49.5)	0.12 (0.08-0.27)	53.0(45.5-60.5)	0.04(-0.05-0,08)	86.0 (80.0-90.7)	0.003 (-0.08-0.05)	19.1 (13.8-25.5)	0.03 (0.01-0.06)	44.4 (37.1-51.8)	0.02 (0.001-0.08)	38.9 (31.7-46.4)	-0.06(0.15- -0.04)
Large	37.0 (24.3-51.3)	0.17 (0.06-0.24)	44.6(31.3-58.5)	0.02(-0.05-0.25)	92.6 (82.1-97.9)	0.27 (0-0.43)	17.5 (8.7-29.9)	0.03 (0.003-0.09)	43.6 (30.3-57.7)	0.05 (0.04-0.08)	50.0 (35.8-64.2)	0.28 (0.25-0.31)
HADS anxiety												
0-7 (normal)	46.6 (42.8-50.4)	0.15 (0.11-0.16)	59.8(56.0-63.5)	0.12(0.06-0.19)	88.8(86.2-91.1)	0.16 (0.08-0.24)	26.8(23.2-30.6)	0.02 (0.02-0.03)	59.2(55.1-63.2)	0.03 (0.01-0.05)	62.4 (58.5-66.2)	0.18 (0.14-0.22)
8-10 (borderline)	45.7(36.4-55.2)	0.16 (0.05-0.26)	49.6(40.2-59.0)	0.04(-0.03-0.10)	87.0(79.4-92.5)	0.15 (0.11-0.28)	23.7(15.5-33.6)	0.04 (0.01-0.04)	55.9(45.2-66.2)	0.07 (-0.04-0.11)	55.0 (45.2-64.6)	0.04 (-0.008-0.17)
≥11 (case)	48.2(34.7-62.0)	0.30 (0.26-0.48)	44.6(31.3-58.5)	0.06(-0.09-0.30)	80.0(67.0-89.6)	**	26.3(15.5-39.7)	0.006 (-0.04-0.06)	63.2(49.3-75.6)	0.07 (-0.04-0.11)	66.7 (52.9-78.6)	0.25 (0.14-0.27)
HADS depression												
0-7 (normal)	46.5(43.0-50.1)	0.16 (0.16-0.19)	58.3(54.8-61.8)	0.12(0.04-0.14)	88.8(86.4-90.9)	0.16 (0.07-0.29)	26.9(23.6-30.4)	0.03 (0.02-0.04)	58.9(55.0-62.6)	0.04 (0.01-0.06)	59.5(56.0-62.9)	0.11 (0.09-0.14)
8-10 (borderline)	44.4(29.6-60.0)	0.16(0.12-0.28)	44.4(29.6-60.0)	0.03(-0.08-0.19)	81.8(67.3-91.8)	0.20 (0.06-0.28)	23.7(11.4-40.2)	-0.05 (-0.20—0.001)	65.8(48.6-80.4)	0.07 (-0.02-0.25)	51.1(35.8-66.3)	0.08 (0.03-0.19)
≥11 (case)	46.2(19.2-74.9)	0.19 (-0.29-0.22)	53.8(25.1-80.8)	0.02(-0.18-0.08)	61.5(31.6-86.1)	**	0	-0.07 (-0.50-0.37)	44.4(13.7-78.8)	0.08 (0-0.24)	30.8(9.1-61.4)	-0.04 (-0.46-0.02)
BIS												
0-10	47.1 (41.8-52.4)	0.10 (0.08-0.13)	65.3(60.1-70.2)	0.15(0.07-0.17)	91.1 (87.6-93.8)	0.28 (0.23-0.40)	29.7 (24.6-35.2)	-0.02 (-0.04-0.02)	61.3 (55.6-66.9)	-0.03 (-0.05-0.03)	62.2 (57.0-67.3)	0.04 (-0.03-0.11)
≥11	46.1 (41.7-50.6)	0.20 (0.17-0.26)	51.6(47.1-56.1)	0.07(0.02-0.13)	85.5 (82.1-88.5)	0.08 (0.02-0.15)	24.1 (20.2-28.5)	0.05 (0.04-0.06)	57.5 (52.6-62.2)	0.08 (0.04-0.09)	56.0 (51.5-60.4)	0.14 (0.10-0.20)

*** Weighted kappa statistic not done as insufficient patient numbers in categories*

Table 2.4: Concordance between PRO and clinician and photographic assessments of specific NTE at year 5 stratified by baseline characteristics in the IMPORT LOW trial

Baseline item	Breast Shrinkage		Breast induration		Breast Swelling		Overall change in appearance		Nipple position		Bra fitting	
	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)	% agreement (95%CI)	Weighted kappa (95% CI)
Age												
<60 years	51.8(45.5-58.0)	0.27 (0.15-0.32)	55.4(49.1-61.6)	0.14(0.04-0.20)	88.3(83.7-92.0)	0.15 (0.10-0.20)	26.8(21.1-33.0)	0.03 (0.02-0.04)	54.2(46.8-61.4)	0.02 (-0.05-0.05)	58.5 (52.8-64.6)	0.10 (0.01-0.13)
≥60 years	45.5(41.1-50.1)	0.11 (0.07-0.13)	63.7(59.3-67.9)	0.09(0.04-0.17)	91.8(89.0-94.1)	-0.04 (-0.04—0.03)	32.5(28.2-37.0)	0.03 (0.003-0.03)	55.7(50.5-60.7)	0.10 (0.04-0.17)	63.2 (58.9-67.4)	0.20 (0.14-0.26)
Treatment Group												
Group 1	43.2(36.6-49.9)	0.13 (0.11-0.18)	55.0(48.3- 61.6)	0.08(0.01-0.09)	88.1(83.1-92.0)	0.07 (-0.05-0.13)	36.1(29.1-43.6)	0.08 (0.005-0.09)	56.4(48.8-63.7)	0.12(0.06-0.17)	57.7 (51.4-63.9)	0.13 (0.09-0.16)
Group 2	49.0(42.8-55.3)	0.21 (0.15-0.23)	62.3(56.0-68.2)	0.12(-0.04-0.22)	89.8(85.4-93.2)	0.02 (-0.05-0.06)	31.6(25.0-38.7)	0.09 (0.04-0.16)	51.3(44.0-58.6)	0.07(0.06-0.15)	62.3 (56.2-68.0)	0.15 (0.04-0.19)
Group 3	50.2(44.0-56.4)	0.17 (0.08-0.20)	64.5(58.4-70.3)	0.15(0.10-0.17)	93.5(89.8-96.2)	0.07 (-0.04-0.17)	36.8(30.1-43.9)	0.10 (0.07-0.14)	57.8(50.6-64.7)	0.05 (-0.04-0.11)	64.7 (58.6-70.4)	0.21 (0.07-0.24)
Breast Size												
Small	46.6 (40.5-52.8)	0.20 (0.17-0.25)	61.6(55.5-67.4)	0.16(0.11-0.24)	92.5 (88.6-95.3)	-0.03 (-0.06- -0.003)	29.7 (24.0-35.9)	0.04 (0.02-0.07)	54.4 (47.8-60.8)	0.11 (0.07-0.14)	61.6 (55.7-67.2)	0.19 (0.15-0.23)
Medium	44.9 (37.9-52.0)	0.14 (0.07-0.14)	63.2(56.2-69.9)	0.13(0.09-0.33)	90.6 (85.7-94.2)	0.04 (-0.04-0.06)	37.0 (30.3-44.1)	0.11 (0.10-0.14)	55.5 (48.3-62.5)	0.12 (0.09-0.25)	58.7 (51.9-65.2)	0.13 (0.07-0.21)
Large	51.1 (42.3-60.0)	0.17 (0.14-0.20)	59.1(50.2-67.6)	0.06(-)	85.5 (78.3-91.0)	0.09 (-0.01-0.16)	39.5 (30.7-48.9)	0.18 (0.15-0.21)	55.6 (46.1-64.7)	0.02 (-0.19-0.08)	61.6 (52.5 -70.2)	0.20 (0.14-0.34)
Surgical deficit												
Small	52.1 (47.1-57.0)	0.20 (0.15-0.22)	64.5(59.6-69.1)	0.17(0.07-0.20)	90.0 (86.6-92.7)	0.04 (-0.05-0.09)	39.4 (34.3-44.5)	0.10 (0.09-0.14)	61.0 (55.8-66.0)	0.05 (0.02-0.08)	66.7 (62.1-71.2)	0.19 (0.11-0.29)
Medium	39.2 (31.3-47.5)	0.09 (0.03-0.17)	55.4(47.0-63.6)	0.07(-0.06-0.15)	91.2 (85.4-95.2)	0.09 (-0.02-0.17)	26.4 (19.4-34.4)	0.08 (0.06-0.11)	46.9 (38.6-55.3)	0.09 (-0.02-0.21)	44.7 (36.9-52.7)	0.03 (-0.02-0.07)
Large	26.1 (14.3-41.1)	-0.07 (-0.21-0.06)	56.5(41.1-71.1)	0.05(-0.14-0.28)	90.9 (78.3-97.5)	-0.04 (-0.11-0.04)	18.6 (8.4-33.4)	-0.03 (-0.09-0.04)	31.0 (17.6-47.1)	0.02 (-0.16-0.15)	56.5 (41.1-71.1)	0.27 (0.19-0.35)
HADS anxiety												
0-7 (normal)	48.3(44.2-52.4)	0.18 (0.16-0.24)	62.8(58.8-66.7)	0.09(0.04-0.12)	91.8(89.3-93.9)	0.03 (-0.04-0.06)	34.9(30.5-39.4)	0.07 (0.04-0.13)	55.8(51.1-60.5)	0.05(0.03-0.08)	62.4 (58.5-66.2)	0.18 (0.13-0.20)
8-10 (borderline)	45.1(35.2-55.3)	0.09 (0.06-0.19)	55.9(45.7-65.7)	0.14(0.05-0.33)	93.1(86.2-97.2)	0.18 (0-0.31)	32.4(21.8-44.5)	0.13 (0.09-0.16)	46.5(34.5-58.7)	0.05(0.003-0.13)	55.0 (45.2-64.6)	0.04 (-0.02-0.13)
≥11 (case)	46.3(32.6-60.4)	0.24 (0.20-0.38)	48.1(34.0-62.4)	0.17(0.04-0.21)	71.7(57.7-83.2)	-0.005 (-0.08-0.16)	38.1(23.6-54.4)	0.20 (-0.04-0.37)	62.8(46.7-77.0)	0.31(0.14-0.47)	66.7 (52.9-78.6)	0.25 (0.16-0.32)
HADS depression												
0-7 (normal)	47.8(44.1-51.6)	0.17 (0.12-0.21)	61.3(57.5-64.9)	0.10(-0.004-0.11)	91.6(89.3-93.5)	0.05 (-0.05-0.14)	34.5(30.5-38.7)	0.08 (0.06-0.10)	56.1(51.8-60.4)	0.08 (0.04-0.10)	62.6 (58.9-66.1)	0.17 (0.16-0.19)
8-10 (borderline)	47.8(32.5-63.3)	0.22 (0.16-0.46)	53.5(37.7-68.8)	0.23(0.17-0.35)	76.7(61.4-88.2)	0.07 (0-0.11)	38.7(21.8-57.8)	0.21 (0.10-0.29)	41.9(24.5-60.9)	0.05 (-0.19-0.36)	50 (34.6-65.4)	0.07 (-0.05-0.22)
≥11 (case)	16.7(4.2-64.1)	-0.33 (-0.43-0)	83.3(35.9-99.6)	0.40(0.20-0.42)	66.7(22.3-95.7)	0 (0-1.0)	50.0(6.8-93.2)	0.14 (0-0.25)	50.0(6.8-93.2)	**	17.7 (15.7-84.3)	0.11 (0.07-0.18)
BIS												
0-10	53.0 (47.3-58.6)	0.22 (0.14-0.26)	67.2(61.7-72.4)	0.15 (0.07-0.18)	93.9 (90.6-96.3)	-0.03 (-0.03—0.02)	38.8 (32.6-45.3)	0.06 (0.06-0.12)	61.6 (55.1-67.8)	0.10 (0.03-0.19)	65.6 (60.2-70.7)	0.19 (0.16-0.21)
≥11	43.8 (39.1-48.6)	0.14 (0.10-0.15)	56.3(51.5-61.0)	0.09(0.08-0.13)	88.2 (84.8-91.1)	0.07 (0.04-0.21)	32.0 (27.1-37.3)	0.11 (0.08-0.31)	50.6 (45.1-56.1)	0.06 (0.03-0.12)	58.8 (54.2-63.4)	0.11 (0.04-0.14)

** *Weighted kappa statistic not done as insufficient patient numbers in categories*

Table 2.5: Summary of baseline factors associated with concordance between PRO and CRO/photographs using logistic regression models (univariate analysis) at 2 and 5 years in IMPORT LOW

NTE assessed by PRO vs CRO/photo	**Time point**	**Factor associated with concordance** *Odds Ratio (OR) 95% confidence interval (95%CI), p value*
Breast smaller versus shrinkage	2 years 5 years	- Larger surgical deficit: 0.32 (0.16-0.65), p=0.001
Breast hardness/firmness versus induration	2 years 5 years	Treatment group 3: 1.81 (1.29-2.53), p=0.001 -
Breast swelling oedema	2 years 5 years	Larger breast size: 0.43 (0.24-0.76), p=0.004 Case levels of anxiety: 0.23 (0.12-0.44), p<0.001 and borderline depression: 0.30 (0.14-0.65) p=0.002**
Bra fitting versus breast shrinkage	2 years 5 years	Medium surgical deficit: 0.33 (0.23-0.47), p<0.001 Younger age 1.00 (1.01-1.06), p=0.002
Change in appearance versus photographic appearance change*	2 years 5 years	- -
Nipple position affected versus photographic appearance change*	2 years 5 years	Larger surgical deficit: 0.38 (0.22-0.68), p=0.001 Larger surgical deficit: 0.29 (0.14-0.57), p<0.001

**comparison with photographic appearance, ** Anxiety and depression were tested on multivariate analysis and higher levels of anxiety (as measured on HADs) remained significantly associated with discordance for breast oedema [OR 0.31, 95%CI 0.15-0.68, p=0.003]*

2.5 Discussion

This analysis in the context of a randomised trial of partial-breast radiotherapy found few patients had moderate/ marked NTE irrespective of the data ascertainment method used. In general, patients reported NTE more frequently compared with NTE assessed by clinicians or via photographs. Concordance was poor between PRO and either CRO or photographs on an individual patient level. However, results obtained for randomised comparisons between treatment groups were consistent for PRO and either CRO or photographs. There were no clinically significant associations found between baseline characteristics and concordance of NTE.

The low overall prevalence of moderate/ marked NTE, irrespective of the data ascertainment method used, has been reported in a number of adjuvant breast radiotherapy trials [2,15,100]. It is therefore increasingly important, in an era of improving radiotherapy techniques to monitor NTE using sufficiently sensitive methods. Within IMPORT LOW, patients reported NTE more frequently compared with clinicians or photographs; this has been previously documented in the literature [47,101-107]. This suggests NTE may be underestimated if only clinician-reported or photographic outcomes are used. In contrast, the Cambridge IMRT trial [48] found clinicians reported a higher prevalence of breast changes than patients which may be related to the Cambridge study being a single-centre study with assessments conducted by one individual.

Concordance was poor on an individual patient level in IMPORT LOW. This could be explained by, firstly, the methods not being designed to be interchangeable given the different comparators used. Secondly, each method is asking a slightly different question; when patient-reported bra fitting was compared with clinician-reported breast shrinkage, patients were deciding what is a reasonable fit in general, whereas clinicians reported degree of breast shrinkage. Thirdly, each method has its own scoring sub-scale which may be worded and categorised differently. Poor concordance has been consistently reported in the literature to date [47,48,101,102,108-110]. Furthermore, it has been argued that some variation is ‘quite acceptable and comprehensible’ due to the

methodological differences between toxicity scoring by patients and clinicians [111].

Although concordance was poor on an individual patient level, the three methods generated similar estimates of effect sizes in terms of comparisons between the randomised treatments, suggesting it is reasonable to use any method to assess between treatment group effects. These findings are consistent with those from the START trials [47]. Within IMPORT LOW there also appeared to be a higher sensitivity of PRO to treatment volume, although the effect sizes obtained from PRO remained consistent with CRO and photographs. It should be noted that the PRO investigated in this analysis and the START trials were the protocol-specific items, which were specifically developed to capture late radiotherapy effects [13], rather than generic PRO related to general quality of life [25].

With respect to the influence of baseline characteristics on concordance, findings were not consistent across NTE or years of assessment and most associations found were significant on univariate analysis only. It is therefore not possible to draw any firm conclusions from these data. The START [47] and Cambridge IMRT [48] trials found no evidence of associations between baseline factors and concordance of NTE assessment methods.

In relation to which NTE assessment methods to use in future breast radiotherapy trials, each has advantages and disadvantages. Clinicians are able to assess the breast with a 3-D view whereas this is not possible with standard photographs (unless taken from various angles providing an overall composite of the breast, although limited resources may prevent this). However, there is a risk of 'bias reporting', as clinicians cannot be blinded to the allocated radiotherapy treatment. Also, varying thresholds of experience in grading toxicity between clinicians can lead to interobserver variability; there was no formal training protocol for clinicians assessing NTE in IMPORT LOW. Furthermore, changes in UK working practices including earlier discharge of patients back to primary care make hospital-based follow-up challenging [112].

Obtaining photographs is also becoming increasingly challenging. Firstly, despite consenting to participate in a photographic sub-study, patients may not attend to have photographs taken. There is a risk of 'informative censoring'

where patients may choose not to attend for photographs either 1) because they do not think there is a problem with their treated breast or 2) they may have experienced NTE and feel uncomfortable about having photographs, resulting in a self-selected population. Of note there was no evidence of change in attendance for year 5 photographs based on year 2 photograph scores in IMPORT LOW. Additionally, workforce changes including closure of medical photography departments make it harder to schedule photographs. It should be noted that photographs provide the only unbiased comparison of NTE between randomised treatment groups [2,15,18,48] as the panel of clinicians scoring photographs are blinded to treatment allocation. Photographs also provide a permanent record at a fixed time point and can be filed and stored for future use. Scoring can also be validated by repeat scoring from different observers [43]. However, in IMPORT LOW, there was a large discrepancy in rating overall change in breast appearance between photographs and PRO (% agreement = 26% and 35% at year 2 and 5 respectively). Patients more frequently reported NTE at both time-points, suggesting photographs may not capture the changes which are important for patients.

PRO provide an opportunity to understand the patients' own perception of NTE within the framework of questions asked. It is known that patients report NTE more frequently than clinicians [47,101-107] or photographs and therefore, without the use of PRO, the prevalence of NTE may be underestimated. Furthermore, PRO are able to distinguish between treatment groups [2,13,15]. Within the START trials, all three data ascertainment methods were able to differentiate between randomised treatment groups [113,114] whereas in IMPORT LOW it was found that only PRO were able to distinguish between randomised comparisons [15]. This difference in findings is likely related to the NTE event rate being lower in IMPORT LOW than in the START trials. In future breast radiotherapy trials (with expected low NTE rates), PRO may have better capability in differentiating between treatment groups.

However, there are a number of issues related to PRO. Firstly, certain patient groups may not wish to participate in a PRO study, resulting in a trial population unrepresentative of the trial population. Secondly, obtaining complete datasets can be challenging [21] as questionnaires may not be returned and individual

questions may not be completed. Thirdly, there is a risk of bias related to questionnaire return as patients who return questionnaires may have different characteristics to those who don't and may report either more or fewer side-effects. This is discussed further in Chapter 3 where it was found that in IMPORT LOW, women who declined participation in the PRO sub-study were slightly older than those who did consent [100]. There were no significant differences in the majority of baseline characteristics in those who did or did not return questionnaires at 5 years, with the exception of higher baseline HADS anxiety and depression subscale scores in patients who did not return their year 5 questionnaire [100]. Also, patients who reported adverse effects more frequently at year 2 were more likely to return questionnaires at year 5 [100]. The prevalence of NTE at individual time-points may therefore be overestimated. Finally, irrespective of missing data, there is also risk of 'bias reporting', as patients cannot be blinded to treatment group in radiotherapy trials. Although the risk of bias reporting cannot be avoided, strategies can be implemented to reduce missing data. Strategies to reduce missing data include collecting data electronically, such as via smart phone/ email. Reducing numbers of questions in PRO questionnaires to include only the most salient and discriminating questions may also improve return rates. As well as obtaining complete and unbiased data-sets for PRO, improvements in the standardisation of analysis, interpretation and reporting of PRO data in clinical trials are also required to enable cross-comparison of data between trials [115].

We have discussed whether PRO could potentially replace either CRO or photographs as the primary method to assess NTE. Broadly, patients rate their subjective satisfaction with an experience of a range of breast changes, whilst clinicians seek objective adverse treatment effects. Therefore, the differences and agreements found by the methods contribute to the overall trial evaluation from multiple perspectives, affecting both the individual patient and randomised trial population. It is acknowledged CRO are still widely supported and an alternative viewpoint is that both PRO and CROs may be necessary as they measure differing aspects of disease experience and are complementary [116].

The main limitation of this analysis is that the IMPORT LOW trial was not designed to address the specific question of concordance between the data

ascertainment methods therefore methodological issues regarding data ascertainment exist. These include each of the methods asking a slightly different question and using different comparators, with various subscales. The lack of standardisation between the methods may limit comparability between PRO and either CRO or photographs.

2.6 Conclusions

Few patients had moderate/marked NTE irrespective of method used. Patients reported NTE more frequently than CRO and photographs, therefore NTE may be underestimated if PRO are not used. PRO provide a patient-centred approach to collecting NTE data **and** are able to differentiate between randomised treatment schedules. Furthermore, the effect sizes from PRO were consistent with CRO and photographs. Thereby, it is appropriate to use PRO as primary NTE endpoints in breast radiotherapy trials.

Chapter 3 Patient-Reported Outcomes Over 5 Years After Whole- or Partial-Breast Radiotherapy: Longitudinal Analysis of the IMPORT LOW Phase III Randomised Controlled Trial

3.1 Abstract

Background: IMPORT LOW demonstrated non-inferiority of partial-breast and reduced-dose radiotherapy compared with whole-breast for local relapse, and similar or reduced toxicity at 5 years following radiotherapy. Comprehensive patient reported outcomes (PRO) collected at serial time-points are now reported.

Methods: IMPORT LOW (ISRCTN12852634) recruited women aged ≥50 years after breast conserving surgery for low risk invasive breast cancer. Patients were randomly assigned to 40Gy whole-breast radiotherapy (control), 36Gy whole-breast and 40Gy to partial-breast (reduced-dose), or 40Gy to partial-breast only (partial-breast) in 15 daily fractions. EORTC QLQ-C30, QLQ-BR23, Body Image Scale, protocol-specific items and Hospital Anxiety and Depression Scales were administered at baseline, 6 months, 1, 2 and 5 years. Patterns of moderate/marked adverse effects (AE) were assessed using longitudinal regression models and baseline predictors investigated.

Results: Patients from 41/71 participating centres took part in the PRO sub-study and 1265/1333 (95%) patients consented. 557/962 (58%) patients reported no moderate/marked AE at 5 years. Breast appearance change was most prevalent and persisted over time (around 20% at each time-point). Prevalence of breast hardness, pain, oversensitivity, oedema and skin changes reduced over time ($p<0.001$ for each), whereas breast shrinkage increased ($p<0.001$). Analysis by treatment group showed average number of AE per person at each time point was lower in the partial-breast (incidence rate ratio [IRR] 0.77, 95%CI 0.71-0.84, $p<0.001$) and reduced-dose (IRR 0.83, 95%CI 0.76-0.90, $p<0.001$) groups compared with the whole-breast group, and decreased over time in all groups. Younger age, larger breast size/surgical deficit, lymph node positivity, and higher

levels of anxiety/depression were baseline predictors of subsequent AE reporting.

Conclusions: The majority of AE reduced over time. There were fewer AE in the partial and reduced-dose groups and baseline predictors of AE were identified. These findings will facilitate informed discussion and shared-decision making for future patients receiving moderately hypofractionated breast radiotherapy.

3.2 Introduction

Trials of hypofractionated whole-breast radiotherapy following breast conserving surgery have demonstrated that 40Gy in 15 fractions is safe and effective, with patients reporting lower levels of moderate/marked adverse effects compared with 50Gy in 2Gy daily fractions [2,3]. Consequently, the three-week regimen tested in START-B trial [2] has become UK standard of care for whole breast radiotherapy and is used increasingly in many other countries [4]. Subsequently, IMPORT LOW investigated efficacy of partial-breast versus whole-breast irradiation using standard UK hypofractionated radiotherapy [15]. The randomised trial schedules were: 40 Gy whole-breast radiotherapy (control); 36 Gy whole-breast and 40 Gy partial breast (reduced-dose group); and 40 Gy partial breast only (partial-breast group) in 15 daily fractions, using simple intensity-modulated radiotherapy (IMRT).

IMPORT LOW demonstrated non-inferiority of partial-breast and reduced-dose radiotherapy compared with standard whole-breast radiotherapy for local relapse, with similar or fewer late normal-tissue adverse effects at 5 years using clinician assessments, patient reported outcomes (PRO) related predominantly to changes in the breast and serial photographs [15]. These published results demonstrated that at 5 years, patients generally reported fewer moderate/marked adverse effects for skin changes, overall breast appearance change, breast smaller and breast harder/firmer to touch in the partial-breast group compared with the whole-breast group, although the reduction was only statistically significant ($p<0.0001$) for change in breast appearance [15].

This chapter builds on the previous publication with more detailed interrogation of the large and comprehensive IMPORT LOW PRO dataset. The main objective was to determine whether breast cancer treatment-related adverse effects (AE) improve, persist or worsen over time, in order to inform future patients. In addition, it was hypothesised that baseline patient, tumour and treatment-specific factors could be identified, which influence patterns of patient AE reporting during the 5 year period.

3.3 Methods

3.3.1 Patients

As described previously, the PRO sub-study was conducted in a subset of patients in the IMPORT LOW trial (IMPORT LOW Appendix), for which full details of patients and procedures have been published [15]. All IMPORT LOW centres were invited to participate in the PRO sub-study (until sufficient accrual to the sub-study was achieved) and 41/71 centres participated. The majority of centres gave no reason for declining to participate, but a few stated lack of local research resources. There was no suggestion of a systematic difference between those centres who did and did not participate. All patients at these 41 centres were invited to participate in the sub-study until the designated sample size had been obtained. Centres that opened after sufficient accrual to the PRO sub-study was achieved were not invited to participate. Separate consent was given for the PRO sub-study. This was approved by the Oxfordshire Research Ethics Committee B (06/Q1605/128) and conducted in accordance with the principles of Good Clinical Practice.

3.3.2 Procedures

Full details of trial procedures have been previously published [15]. Women were randomly assigned (in a 1:1:1 ratio) to receive whole-breast radiotherapy or one of the experimental schedules (reduced-dose or partial-breast radiotherapy). Patients who consented to participate in the PRO sub-study completed a baseline questionnaire booklet before randomisation (post-surgery and pre-radiotherapy). Subsequent questionnaires were posted by the ICR-CTSU for completion at the patients' home at 6 months, 1, 2, and 5-years post-randomisation (after checking current health status with local hospital team or family doctor), unless they had died or withdrawn from active trial follow-up. Patients were prompted by telephone or letter if questionnaires were not returned within 3 weeks.

Patients completed the EORTC (European Organisation for Research and Treatment of Cancer) general cancer scale QLQ-C30 and QLQ-BR23 breast-

cancer specific module [25,26], Hospital Anxiety and Depression Scale (HADS): scores of 8-10 indicating borderline anxiety or depression, and scores of 11-21 indicating case levels of anxiety or depression [27], 10-item Body Image Scale [28] and protocol-specific questionnaire items asking patients to score 'change in breast appearance', 'breast hardness/firmness', 'reduction in size of breast', 'change in skin appearance', 'is the position of the nipple of your affected breast different from the other side', 'problem getting a bra to fit' and 'shoulder stiffness' which may have resulted from any prior breast cancer treatments [13]. The protocol-specific items were designed by the START trialists and added to by patient advocate members of the IMPORT Trial Management Group to assess specific post-radiotherapy adverse effects not included in the EORTC questionnaires [13]. The 2 protocol-specific items 'is the position of the nipple of your affected breast different from the other side' and 'problem getting a bra to fit' were introduced following a protocol amendment. All items were scored on a four-point scale (none, a little, quite a bit, very much). The protocol-specific items were not included initially at baseline, but were added following a protocol amendment. Individual items from the QLQ-BR23 and Body Image Scales rather than overall summary scores were used in this analysis, in order to identify which questionnaire items were most discriminating of breast cancer treatment adverse effects. This generated 24 potential AE at each time point, which focussed on AE most likely to be radiotherapy-related and all 10 items from the Body Image Scale.

Patients who consented to the PRO sub-study were also invited to participate in the photographic sub-study, which involved assessments at baseline (post-surgery but pre-radiotherapy), then at 2 and 5 years after randomisation. Breast size and surgical deficit were scored on a 3-point scale (small, medium, large) from baseline photographs by a single team of three observers blinded to patient identity and treatment allocation. Further details of the photographic sub-study in IMPORT LOW have been published [15].

3.3.3 Statistical Methods

The individual items from the QLQ-BR23, Body Image Scale and protocol-specific items were dichotomised for analysis as none/a little (none/mild) versus quite a bit/very much (moderate/marked). This follows the methodology used in previous breast radiotherapy trials [2]. The observation that very few effects were scored as marked, supported the decision to combine moderate and marked categories. All analyses conducted in the whole cohort were adjusted for treatment group.

The prevalence of moderate/marked AE at each time point was determined. If ≥10% moderate/marked AE were reported for any questionnaire item at any time point, the changes in AE reporting over time were tested using the chi-squared test for trend for those items (table 3.1). An arbitrary cut off of ≥10% prevalence was chosen as this was deemed to be clinically significant by the IMPORT Trialists. A generalised estimating equation (GEE) model [117] using a 2-way interaction between treatment group and time was built to investigate whether prevalence of moderate/marked AE over time differed between the treatment groups. A Poisson model adjusted for time and treatment group was fitted to identify whether average number of moderate/marked AE per person changed over time, and whether this varied according to treatment group by including an interaction term in the model. Potential baseline predictors of moderate/marked AE over 5 years included patient factors (age, breast size from baseline photographs, HADS anxiety and depression subscale scores, education level and whether the patient lived alone), and tumour and treatment-related factors (tumour size, surgical deficit from baseline photographs, tumour grade, lymph node status, lymphovascular invasion), oestrogen receptor (ER) status and adjuvant therapy). With respect to assessing education level as a predictor, patients with any formal education including school certificate, O-level, GCSE, NVQ, A-level, professional qualification, degree or postgraduate degree were compared with patients with no formal education. Separate GEE models for each AE were fitted including terms for time and treatment group, to investigate whether baseline factors predicted reporting of moderate/marked AE over 5 years. Each baseline factor was initially tested individually in GEE models, then factors which were statistically significant ($p \leq 0.005$) were included together in a

multivariable analysis to find those which remained statistically significantly associated with reporting of moderate/marked AE over 5 years. A significance level of ≤0.005 was chosen to allow for multiple testing in all analyses. There was no imputation of missing data.

All analyses were carried out using STATA version 14 based on a database snapshot taken on June 15th 2016 (as per the primary endpoint analysis). Analysis was on an intention to treat basis. The IMPORT LOW trial is registered in the ISRCTN registry (ISRCTN12852634) and ClinicalTrials.gov (NCT00814567).

Table 3.1: Summary of questionnaire items investigated to identify proportion of moderate and marked events

BR23 (shoulder, breast and arm) (during the past week)
Did you have any pain in your arm or shoulder?
Did you have a swollen arm or hand?
Was it always difficult to raise your arm or to move it sideways?
Have you had pain in the area of your affected breast?
Was the area of your affected breast swollen?
Was the area of your affected breast oversensitive?
Have you had skin problems on or in the area of your affected breast (e.g. itchy, dry, flaky)?
Body image scale (during the past week)
Have you been self-conscious about your appearance?
Have you felt less physically attractive as a result of your disease or treatment?*
Have you been dissatisfied with your appearance when dressed?
Have you been feeling less feminine as a result of your disease or treatment?*
Did you find it difficult to look at yourself naked?*
Have you been feeling less sexually attractive as a result of your disease or treatment?
Did you avoid people because of the way you felt about your appearance?
Have you been feeling the disease or treatment has left your body less whole?
Have you been dissatisfied with your body?*
Have you been dissatisfied with the appearance of your scar?
**refer to the past week*
Protocol specific items (any changes to your breast that may have resulted from any of your breast cancer treatments)
Has the appearance of the skin in the area of your affected breast changed?
Has the overall appearance of your breast changed, compared with the other side?
Has your affected breast become smaller?
Has your affected breast become harder/firmer to the touch?
Is the position of the nipple of your affected breast different from the other side?
Have you had a problem getting a bra to fit?
Did you have any stiffness in your shoulder?

3.4 Results

Between May 2007 and October 2010, 2018 women were enrolled into IMPORT LOW from 71 UK participating centres. 1333 patients from 41 centres were offered participation in the PRO sub-study and 1265 (95%) patients consented. The majority of women had small, approximately 1cm, grade 1 or 2 oestrogen-receptor (ER) positive and human epidermal growth factor 2 negative (HER2-ve) node negative tumours and received adjuvant hormonal therapy (table 3.2). Baseline prevalence of anxiety and depression scores in the borderline and case range as assessed by the HADS subscales were 23% and 8% respectively. Patients who declined participation in the PRO study were slightly older, median age 70 (range 63-76) versus 63 (57-67), however, other baseline characteristics were similar between those who did and did not consent to participate (table 3.2). There were no significant differences between patients who did and did not return questionnaires at year 5 in terms of baseline characteristics including age, tumour characteristics, surgery details, ER status, adjuvant therapy and body image (table 3.3). There was some evidence that those patients who did not return their 5 year questionnaire had higher HADS anxiety and depression subscale scores at baseline than those who did (table 3.3). Excluding patients who had died or withdrawn, there was a higher return rate of 5 year questionnaires in patients who reported at least 1 AE at 2 years [326/425 (85%)] compared with those who reported no AE at 2 years [601/764 (79%), p=0.008]. There was a high proportion of completed questionnaires (based on the number returned versus expected, excluding patients who had died or withdrawn from the study) at all time-points (figure 3.1).

In the whole cohort, 557/962 (58%) patients reported no AE at 5 years. Overall change in breast appearance was the most prevalent AE reported at each time point and persisted over time (19% at 1 year and 21% at 5 years) (table 3.4). Other moderate/marked AE, with a prevalence of over 10% at least once during the 5 years, were skin changes, breast hardness/firmness, breast shrinkage, nipple position affected, arm/shoulder pain, breast pain, breast swelling and breast oversensitivity (table 3.4).

Table 3.2: Baseline characteristics of patients who consented and declined participation in the IMPORT LOW PRO sub-study (in patients with available data)

	Patients who consented N=1265, n (%)	**Patients who declined N=68, n (%)**
Age (years), Median (IQR)	63 (57-67)	70 (63-76)
Side of Primary		
Left	638/1264 (50.5)	31/68 (45.6)
Right	626/1264 (49.5)	37/68 (54.4)
Pathological tumour size, Median (IQR)	1.1 (0.8-1.5)	1.25 (0.9-1.6)
Tumour Grade		
Grade 1	537/1260 (42.6)	20/68 (29.4)
Grade 2	615/1260 (48.8)	41/68 (60.3)
Grade 3	108/1260 (8.6)	7/68 (10.3)
Re-excision		
Yes	166/1264 (13.1)	5/68 (7.4)
No	1098/1264 (86.9)	63/68 (92.6)
Axillary Surgery Performed		
Yes	1263/1264 (99.9)	68/68 (100)
No	1/1264 (0.1)	-
Pathological Nodal status		
Positive	32/1264 (2.5)	3/68 (4.4)
Negative	1232/1264 (97.5)	65/68 (95.6)
Lymphovascular invasion		
Present	59/854 (6.9)	7/48 (14.6)
Absent	795/854 (93.1)	41/48 (85.4)
ER status		
Positive	1208/1262 (95.7)	60/67 (89.6)
Poor	54/1262 (4.3)	7/67 (10.4)
PR status		
Positive	695/845 (82.2)	35/49 (71.4)
Poor	150/845 (17.8)	14/49 (28.6)
HER2 status		
Positive	41/991 (4.1)	7/57 (12.3)
Negative	891/991 (89.9)	49/57 (85.9)
Inconclusive	59/991 (6.0)	1/57 (1.8)
Adjuvant systemic therapy received		
Yes	1183/1259 (94.0)	65/68 (95.6)
No	76/1259 (6.0)	3/68 (4.4)
Type of Adjuvant therapy received~:		
Chemotherapy	38/1259 (3.0)	1/68 (1.5)
Tamoxifen	756/1259 (60.0)	36/68 (52.9)
AI	475/1259 (37.7)	21/68 (30.9)
Trastuzamab	17/1259 (1.4)	2/68 (2.9)
Breast size*		
Small	437/985 (44.4)	14/35 (40.0)
Medium	344/985 (34.9)	13/35 (37.1)
Large	204/985 (20.7)	8/35 (22.9)
Surgical deficit*		
Small	639/986 (64.8)	19/35 (54.3)
Medium	258/986 (26.2)	13/35 (37.1)
Large	89/986 (9.0)	3/35 (8.6)
Hospital Anxiety and Depression Scale		
Anxiety		
Normal (0-7)	953/1243 (76.7)	
Borderline (8-10)	177/1243 (14.2)	
Case (11+)	113/1243 (9.1)	
Depression		
Normal (0-7)	1145/1242 (92.2)	
Borderline (8-10)	73/1242 (5.9)	
Case (11+)	24/1242 (1.9)	
Body image scale** (10 items): median (interquartile range)	1 (0-4)	

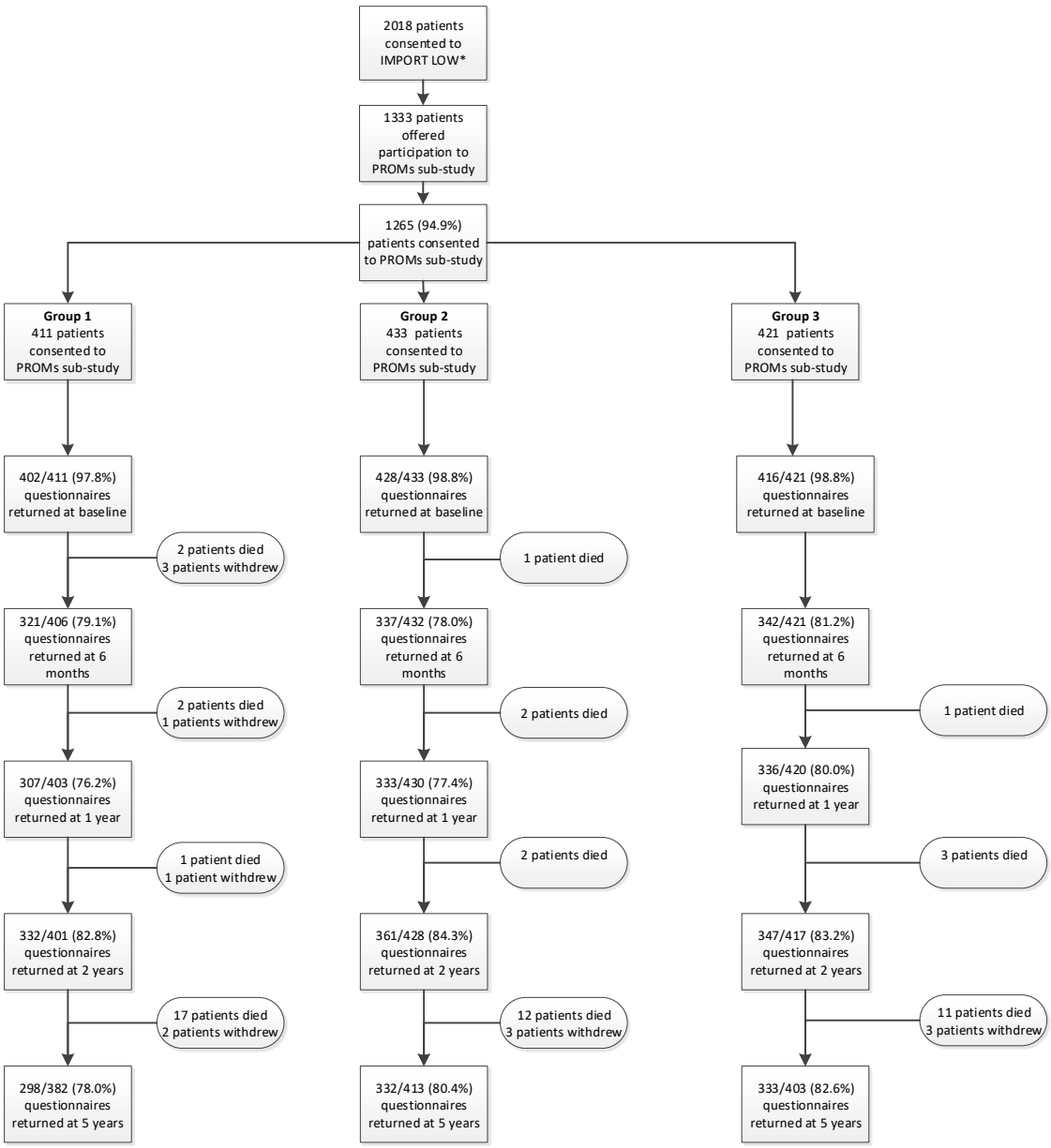
*~Not mutually exclusive, *Assessed on baseline photographs, ** Body image scale scored from 1-4 using the EORTC method (possible range 0-30). AI=Aromatase inhibitor, IQR=interquartile range.*

Table 3.3: Summary of baseline characteristics of patients who did and did not return questionnaires at year 5

	Patients who returned questionnaires at 5 years N=963, n (%)	**Patients who did not return questionnaires at 5 years N=235, n (%)**
Age, years	N=963	N=226
Mean (standard deviation)	62 (7.00)	63 (7.74)
Side of Primary	N=963	N=225
Left	488 (50.7)	111 (49.3)
Right	475 (49.3)	114 (50.6)
Pathological tumour size, cm	N=963	N=225
Mean (standard deviation)	1.19 (0.55)	1.25 (0.58)
Tumour Grade	N=960	N=224
Grade 1	418 (43.5)	87 (38.8)
Grade 2	466 (48.5)	115 (51.3)
Grade 3	76 (7.9)	22 (9.82)
Re-excision	N=963	N=225
Yes	128 (13.3)	32 (14.2)
No	835 (86.7)	193 (85.7)
Axillary Surgery Performed	N=963	N=225
Yes	962 (99.9)	225 (100)
No	1(0.1)	0
Pathological Nodal status	N=963	N=225
Positive	21 (2.2)	8 (3.5)
Negative	942 (97.8)	217 (96.4)
Lymphovascular invasion	N=642	N=162
Present	43 (6.7)	8 (4.9)
Absent	599 (93.3)	154 (95.1)
ER status	N=961	N=225
Positive	926 (96.4)	214 (95.1)
Poor	35 (3.6)	11 (4.8)
PR status	N=641	N=155
Positive	533 (83.2)	126 (81.3)
Poor	108 (16.8)	29 (18.7)
HER2 status	N=750	N=180
Positive	32 (4.3)	8 (4.4)
Negative	678 (90.4)	115 (63.9)
Inconclusive	40 (5.3)	17 (9.4)
Adjuvant systemic therapy received	N=961	N=224
Yes	905 (94.2)	210 (93.8)
No	56 (5.8)	14 (6.2)
Breast size*	N=680	N=145
Small	305 (44.9)	60 (41.4)
Medium	240 (35.3)	50 (34.5)
Large	135 (19.9)	35 (24.1)
Surgical deficit*	N=681	N=145
Small	460 (67.5)	84 (57.9)
Medium	171 (25.1)	42 (29.0)
Large	50 (7.3)	19 (13.1)
Hospital Anxiety and Depression Scale		
Anxiety	N=958	N=215
Normal (0-7)	753 (78.6)	146 (67.9)
Borderline (8-10)	130 (13.6)	37 (17.2)
Case (11+)	75 (7.8)	32 (14.9)
Depression	N=956	N=216
Normal (0-7)	891 (93.2)	189 (87.5)
Borderline (8-10)	55 (5.8)	15 (6.9)
Case (11+)	10 (1.0)	12 (5.5)
Body image scale**	N=919	N=204
Median (IQR)	1 (0-4)	1 (0-5.5)

**Assessed on baseline photographs, ** Body image scale scored from 1-4 using the EORTC method (possible range 0-30), IQR=interquartile range*

Figure 3.1: Consort diagram



**2 patients withdrew consent for any of their data to be used in the analysis*

Table 3.4: Moderate/ marked AE reported by patients in the whole cohort over 5 years

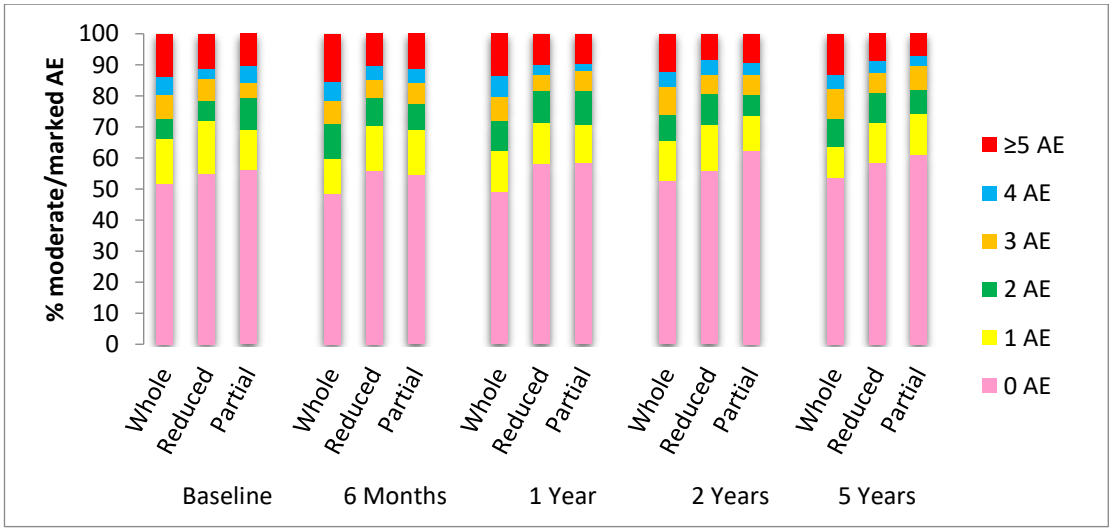
Item	Proportion of moderate/marked AE at **baseline**	Proportion of moderate/marked AE at **6 months**	Proportion of moderate/marked AE at **1 year**	Proportion of moderate/marked AE at **2 years**	Proportion of moderate/marked AE at **5 years**	Chi-square test for trend (p value)*****
BIS- Self-conscious about appearance*	58/1238 (4.7)	54/988 (5.5)	40/969 (4.1)	46/1035 (4.4)	36/955 (3.8)	-
***BIS- less physically attractive as a result of disease/treatment*	101/1220 (8.3)	74/982 (7.5)	55/958 (5.7)	57/1013 (5.6)	50/944 (5.3)	-
BIS- Dissatisfied with appearance when dressed	40/1238 (3.2)	50/991 (5.1)	35/970 (3.6)	44/1034 (4.3)	26/956 (2.7)	-
***BIS- less feminine as a result of disease/treatment*	77/1222 (6.3)	55/981 (5.6)	43/961 (4.5)	42/1015 (4.1)	36/943 (3.8)	-
***BIS-difficult looking at yourself naked*	78/1226 (6.4)	53/984 (5.4)	50/961 (5.2)	53/1014 (5.2)	55/946 (5.8)	-
BIS- Less sexually attractive	86/1224 (7.0)	54/973 (5.6)	59/955 (6.2)	65/1025 (6.3)	44/930 (4.7)	-
BIS - Avoid people because of the way you felt about appearance	15/1239 (1.2)	12/989 (1.2)	9/971 (0.9)	10/1034 (1.0)	4/956 (0.4)	-
BIS- disease/treatment has left body feeling less whole	41/1237 (3.3)	28/990 (2.8)	24/971 (2.5)	21/1036 (2.0)	24/954 (2.5)	-
***BIS- Have you been dissatisfied with your body*	106/1222 (8.7)	76/983 (7.7)	75/961 (7.8)	80/1014 (7.9)	75/943 (8.0)	-
BIS-dissatisfied with the appearance of your scar	42/1238 (3.4)	27/991 (2.7)	24/971 (2.5)	35/1036 (3.4)	28/957 (2.9)	-
*PS*** –appearance of the skin in breast changed*	73/699 (10.4)	86/985 (8.7)	59/969 (6.1)	59/1031 (5.7)	57/949 (6.0)	p<0.001
PS- overall appearance of breast changed	138/700 (19.7)	185/988 (18.7)	179/966 (18.5)	187/1034 (18.1)	195/951 (20.5)	p=0.79
PS-breast smaller	72/694 (10.4)	122/983 (12.4)	157/967 (16.2)	171/1031 (16.6)	185/951 (19.5)	p<0.001
PS-breast harder/firmer to touch	120/697 (17.2)	148/985 (15.0)	114/963 (11.8)	73/1029 (7.1)	65/947 (6.9)	p<0.001
PS-nipple position affected	70/695 (10.1)	93/776 (12.0)	104/876 (11.9)	120/1030 (11.7)	115/944 (12.2)	p=0.36
PS- problem getting bra to fit	43/698 (6.2)	63/783 (8.1)	69/884 (7.8)	83/1033 (8.0)	85/952 (8.9)	-
PS shoulder stiffness	60/1235 (4.9)	69/990 (7.0)	61/969 (6.3)	59/1035 (5.7)	47/955 (4.9)	-
*BR23****- pain in your arm or shoulder*	184/1237 (14.9)	119/987 (12.1)	115/966 (11.9)	109/1034 (10.5)	100/957 (10.5)	p=0.002
BR23-swollen arm or hand	40/1234 (3.2)	18/992 (1.8)	18/967 (1.9)	24/1033 (2.3)	22/955 (2.3)	-
BR23- difficulty raising arm or to moving sideways	71/1234 (5.8)	53/990 (5.4)	39/969 (4.0)	43/1033 (4.2)	42/956 (4.4)	-
BR23- breast pain	137/1234 (11.1)	110/983 (11.2)	67/967 (6.9)	54/1030 (5.2)	44/953 (4.6)	p<0.001
BR23- breast swollen	117/1235 (9.5)	51/988 (5.2)	19/970 (2.0)	15/1029 (1.5)	6/952 (0.6)	p<0.001
BR23-breast oversensitive	167/1237 (13.5)	96/990 (9.7)	57/966 (5.9)	60/1033 (5.8)	38/956 (4.0)	p<0.001
BR23-skin problems on breast	47/1236 (3.8)	62/990 (6.3)	47/968 (4.9)	34/1033 (3.3)	26/954 (2.7)	-

BIS=Body image scale items,Body image items originally from the body image scale however included within the BR23 questionnaire. Items were not duplicated. ***Protocol specific items, ****Items from EORTC-BR-23 subscale, *****Chi-squared test for trend performed for those items where patients reported* ≥10% moderate/ marked AEs overall.

Overall, in patients who reported at least one AE at 5 years, the median number of AE per person at this time point was 3 [interquartile range (1-4)]. Analysis by treatment group showed the median number of AE in the whole-breast group was 3 (IQR 2-5) compared with median of 2 (IQR 1-4) for both test groups at 5 years. The average number of AE reported per person at each time point was lower in the partial-breast (incidence rate ratio [IRR] 0.77, 95%CI 0.71-0.84, p<0.001) and reduced-dose (IRR 0.83, 95%CI 0.76-0.90, p<0.001) groups compared with the whole-breast group.

The number of AE reported per person reduced over time in all treatment groups (figure 3.2), and the rate of reduction was similar between treatment groups (p=0.20). Prevalence of moderate/marked breast hardness, pain, oversensitivity, oedema, skin changes (p<0.001 for each) and arm/shoulder pain (p=0.002) reduced over time (table 3.4). Breast shrinkage was the only AE where prevalence increased over time (p<0.001) [table 3.4]. There was no difference in change in prevalence of individual AE over time between treatment groups.

Figure 3.2: Number of moderate/ marked AE reported per person over time by treatment group



Certain baseline patient factors appeared to predict for some patient reported adverse effects but not for others. For example, within this largely postmenopausal population, younger age at randomisation was associated with worse AE for items in the Body Image Scale (table 3.5) over five years. In contrast, living alone was shown to be associated with reported adverse breast swelling. Education level did not predict for any AE reporting patterns (table 3.6). Baseline anxiety as measured on HADS, was associated with all adverse effects assessed except for change in overall breast appearance, breast shrinkage and problem getting a bra to fit. Baseline depression was also associated significantly with adverse effects including feeling self-conscious about appearance, less physically attractive and dissatisfaction with appearance when dressed (table 3.5). Patients with larger breast size were more likely to report AE including feeling self-conscious about appearance, dissatisfaction with appearance when dressed, and dissatisfaction with body/skin changes (table 3.5).

In relation to tumour and treatment-specific factors, larger surgical deficit predicted for reporting of change in overall appearance of the breast, breast shrinkage, nipple position affected and problem getting a bra to fit (table 3.6). In addition, higher tumour grade was associated with reporting of feeling less sexually attractive and less feminine as a result of disease or treatment, and lymph node positivity predicted for pain in arm or shoulder and swollen arm or hand (tables 3.5 and 3.6).

Table 3.5: Longitudinal analysis of baseline factors tested to predict moderate/marked AE over 5 years in Body Image Scale (BIS) items

	Baseline factors	*BIS- Self-conscious about appearance*	*BIS- less physically attractive as a result of disease/treatment*	*BIS- Dissatisfied with appearance when dressed*	*BIS- less feminine as a result of disease/treatment*	*BIS-difficult looking at yourself naked*	*BIS- Less sexually attractive*	*BIS - Avoid people because of the way you felt about appearance*	*BIS- disease/treatment has left body feeling less whole*	*BIS- Have you been dissatisfied with your body*	*BIS-dissatisfied with the appearance of your scar*
Patient Factors	Age	0.95 (0.92-0.98), p=0.002	0.94 (0.92-0.97), p<0.001	0.97 (0.94-1.00), p=0.04	0.95 (0.92-0.98), p<0.001	0.97 (0.95-1.00), p=0.04	0.95(0.92-0.97), p<0.001		0.96 (0.93-1.00), p=0.04	0.96 (0.93-0.98), p=0.001	
	Breast Size										
	Small	1		1						1	
	Medium	1.02(0.62-1.67), p=0.93		1.19 (0.70-2.03), p=0.51						1.07(0.72-1.60), p=0.74	
	Large	2.42(1.50-3.91), p<0.001		3.07 (1.84-5.12), p<0.001						2.19 (1.46-3.31), p<0.001	
	HADs anxiety										
	Normal (0-7)	1	1	1	1	1	1	1	1	1	1
	Borderline (8-10)	3.59 (2.18-5.91), p<0.001	2.16 (1.39-3.35), p=0.001	2.54 (1.48-4.36), p=0.001	2.07 (1.23-3.48), p=0.006	2.82 (1.82-4.36), p<0.001	1.90 (1.23-2.94), p=0.004	3.24 (1.28-8.18), p=0.01	2.00 (0.99-4.04), p=0.05	2.19 (1.42-3.37), p<0.001	1.61 (0.77-3.38), p=0.21
	Case (11+)	5.84 (3.41-9.99), p<0.001	3.85 (2.40-6.18), p<0.001	3.32 (1.81-6.10), p<0.001	4.23 (2.48-7.22), p<0.001	3.94 (2.39-6.52), p<0.001	3.43 (2.16-5.45), p<0.001	5.04 (1.93-13.15), p=0.001	7.35 (3.92-13.79), p<0.001	2.89 (1.76-4.75), p<0.001	5.47 (2.81-10.66), p<0.001
	HADs depression										
	Normal (0-7)	1	1	1	1	1	1	1	1	1	1
	Borderline (8-10)	2.25 (1.27-3.97), p=0.005	3.95 (2.45-6.36), p<0.001	2.76 (1.49-5.11), p=0.001	4.26 (2.49-7.29), p<0.001	2.63 (1.57-4.39), p<0.001	3.84 (2.39-6.17), p<0.001	6.88 (2.99-15.83), p<0.001	2.82 (1.48-5.38), p=0.002	3.03 (1.81-5.06), p<0.001	1.82 (0.81-4.07), p=0.15
	Case (11+)	2.16 (0.82-5.66), p=0.12	1.65 (0.66-4.12), p=0.29	3.38 (1.30-8.82), p=0.01	1.81 (0.65-5.04), p=0.26	3.51 (1.60-7.67), p=0.002	2.96 (1.29-6.83), p=0.01	8.12 (2.72-24.24), p<0.001	3.71 (1.58-8.71), p=0.003	3.71 (1.58-8.69), p=0.003	3.22 (1.20-8.61), p=0.02
	Surgical Deficit										
	Small										1
	Medium										3.65 (2.12-6.28), p<0.001
	Large										2.98 (1.28-6.93), p=0.01
Tumour and Treatment Factors	Tumour Grade										
	Grade 1				1		1				
	Grade 2				1.67 (1.08-2.58), p=0.02		1.54 (1.07-2.22), p=0.02				
	Grade 3				2.75 (1.38-5.49), p=0.004		2.77(1.57-4.90), p<0.001				
	Adjuvant Chemotherapy										
	No	1	1		1		1				
	Yes	2.30 (1.08-4.92), p=0.03	2.59(1.31-5.12), p=0.006		1.68 (0.74-3.83), p=0.22		1.58 (0.78-3.19), p=0.21				

Table 3.6: Longitudinal analysis of baseline factors tested to predict moderate/marked AE over 5 years listed in protocol-specific (PS) and QLQ-BR23 items

	PS – appearance of the skin in breast changed	*PS- overall appearance of breast changed*	*PS-breast smaller*	*PS-breast harder/firmer to touch*	*PS-nipple position affected*	*PS- problem getting bra to fit*	*PS-shoulder stiffness*	*BR23- pain in your arm or shoulder*	*BR23-swollen arm or hand*	*BR23- difficulty raising arm or to moving sideways*	*BR23- breast pain*	*BR23- breast swollen*	*BR23-breast oversensitive*	*BR23-skin problems on breast*
Age											0.98 (0.96-1.00), p=0.08		0.98 (0.96-1.00), p=0.10	
Breast Size														
Small	1													
Medium	0.67 (0.43-1.02), p=0.06													
Large	1.82 (1.22-2.72), p=0.003													
HADs anxiety														
Normal (0-7)	1	1		1	1		1	1	1	1	1	1	1	1
Borderline (8-10)	1.76 (1.12-2.76), p=0.01	1.37 (0.96-1.97), p=0.09		1.95 (1.34-2.84), p<0.001	1.55 (0.99-2.43), p=0.06		1.57 (1.04-2.39), p=0.03	1.28 (0.93-1.77), p=0.13	1.42 (0.77-2.63), p=0.27	1.20 (0.74-1.96), p=0.46	1.80 (1.27-2.56), p=0.001	2.45 (1.56-3.85), p<0.001	2.28 (1.54-3.38), p<0.001	1.80 (1.15-2.821), p=0.01
Case (11+)	2.27 (1.33-3.87), p=0.003	1.82 (1.16-2.84), p=0.009		3.11 (2.00-4.84), p<0.001	2.93 (1.83-4.71), p<0.001		3.72 (2.41-5.74), <0.001	2.90 (2.05-4.11), p<0.001	3.80 (2.12-6.79), p<0.001	4.89 (3.18-7.51), p<0.001	3.80 (2.59-5.58), p<0.001	4.23 (2.55-7.01), p<0.001	3.88 (2.49-6.05), p<0.001	2.71 (1.62-4.54), p<0.001
HADs depression														
Normal (0-7)	1	1		1			1	1	1	1	1	1	1	1
Borderline (8-10)	3.13 (1.82-5.36), p<0.001	2.10 (1.28-3.44), p=0.004		1.36 (0.80-2.33), p=0.26			1.55 (0.93-2.58), p=0.10	1.86 (1.25-2.77), p=0.002	1.41 (0.70-2.85), p=0.34	1.47 (0.86-2.50), p=0.16	1.98 (1.29-3.04), p=0.002	1.19 (0.64-2.21), p=0.59	1.57 (0.93-2.65), p=0.09	1.70 (0.96-3.02), p=0.07
Case (11+)	2.31 (0.90-5.92), p=0.08	2.34 (0.99-5.56), p=0.05		0.46 (0.15-1.46), p=0.19			1.29 (0.57-2.90), p=0.54	2.00 (1.07-3.74), p=0.03	1.24 (0.44-3.48), p=0.68	1.60 (0.75-3.39), p=0.23	1.70 (0.86-3.37), p=0.13	1.53 (0.60-3.89), p=0.37	1.98 (0.84-4.66), p=0.12	2.39 (1.06-5.38), p=0.04
Educated														
No													1	
Yes													0.73 (0.53-1.02), p=0.06	
Lives alone														
No												1		
Yes												1.92 (1.31-2.83), p=0.001		
Tumour size					1.24 (0.94-1.65), p=0.14									
Surgical Deficit														
Small	1	1	1	1	1	1								
Medium	2.06 (1.41-2.99), p<0.001	2.60 (1.96-3.45), p<0.001	2.84 (2.10-3.84), p<0.001	1.67 (1.22-2.29), p=0.001	2.87 (2.00-4.12), p<0.001	1.94 (1.31-2.85), p=0.001								
Large	2.05 (1.18-3.56), p=0.01	4.23 (2.86-6.26), p<0.001	4.55 (3.03-6.85), p<0.001	1.53 (0.94-2.50), p=0.09	3.99 (2.44-6.55), p<0.001	2.80 (1.65-4.74), p<0.001								
Tumour Grade														
Grade 1				1										
Grade 2				1.46 (1.07-1.98), p=0.02										
Grade 3				1.35 (0.77-2.36), p=0.29										
LN negative								1	1					
LN positive								2.66 (1.53-4.63), p<0.001	7.39 (3.75-14.55), p<0.001					
Adjuvant Tamoxifen														
No													1	
Yes													1.61 (1.13-2.29), p=0.008	

3.5 Discussion

IMPORT LOW provides the largest and most comprehensive report of adverse effects using PRO at serial time points from a randomised controlled trial of partial-breast radiotherapy to date. These data demonstrate that the majority of reported AE reduce over time following moderately hypofractionated external beam radiotherapy and more than half of patients report no moderate/marked AE at 5 years. In addition, the average number of AE reported per person at each time point was lower in both partial-breast and reduced-dose groups compared with the whole-breast group. Overall change in breast appearance was the most prevalent AE reported and this remained stable over time. All other AE decreased over the 5-year period with the only exception being breast shrinkage, which increased.

3.5.1 Reduction in adverse effect reporting over time

Two randomised controlled trials investigating whole-breast radiotherapy at a dose of 40Gy in 15 fractions used very similar PRO assessments, which were carried out at the same time-points as IMPORT LOW. These are the START-B hypofractionation trial [13], which compared 40Gy in 15 fractions over 3 weeks with 50Gy in 25 fractions over 5 weeks and the Cambridge IMRT trial [14], which compared 2-dimensional radiotherapy with forward planned IMRT using 40Gy in 15 fractions in both groups. Both studies allowed a tumour bed boost at clinician discretion.

These two trials had similar PRO time trends to those shown in IMPORT LOW as they also showed that the majority of reported AE reduced over a 5 year period. START-B reported reduction in breast symptoms assessed using the QLQ-BR23 subscale over 5 years following radiotherapy in both standard and hypofractionation groups [13]. The Cambridge IMRT trial [14] reported improvement in AE reporting over the same period, but also showed a slight initial worsening of toxicity at 6 months for skin changes, breast pain, breast oversensitivity and breast swelling, which then improved.

A third study, the GEC-ESTRO trial [118], also incorporated PRO using the QLQ-BR23 subscale and assessments were carried out at baseline and regularly throughout the 5-year follow up period. Patients were randomised to receive

either whole-breast radiation of 50Gy in 25 fractions with a boost of 10Gy or partial-breast radiotherapy using multicatheter brachytherapy. They found that breast symptoms assessed on the QLQ-BR23 subscale were significantly worse immediately following the last fraction of radiotherapy and at 3 months follow-up after receiving whole-breast than after partial-breast radiotherapy. There were no clinically significant differences between the 2 groups from 3 months to 5 years and therefore the initial worsening of symptoms in the whole breast group are likely to be related to acute radiotherapy toxicity. Overall, the majority of AE reported also seemed to decrease over time in both groups.

A fourth study, the Florence trial, used the QLQ-BR23 subscale, but only at baseline and 2 years following completion of treatment. Patients were randomised to partial-breast IMRT using 30Gy in 5 fractions over a week versus 50Gy in 25 fractions over 5 weeks to the whole-breast with an optional tumour bed boost [119]. The partial-breast group showed improvement in PRO at 2 years, whereas breast and arm symptoms worsened in the whole-breast radiotherapy group. This difference from other reported studies may be related to the higher biologically equivalent dose with 50Gy in 25 fractions in the whole-breast radiotherapy group compared with 40Gy in 15 fractions for other studies. It may also reflect the smaller numbers completing PRO questionnaires at both time points in the Florence study (205/520, 39%).

3.5.2 Increase in adverse effect reporting over time

As per IMPORT LOW, START B [13] also demonstrated that breast shrinkage was the only patient reported AE showing an increase in prevalence in the 5 years after completion of radiotherapy. The results of the Cambridge IMRT trial [14] suggested an increase in breast shrinkage over time as reported by patients, but this did not reach statistical significance. In the START-B trial, patients reported significantly less breast shrinkage in the hypofractionation group compared with the control group [13]. In IMPORT LOW, there was less patient-reported breast shrinkage in the reduced and partial-breast groups compared with the whole-breast group, however this did not reach statistical significance [15].

This increase in the prevalence of breast shrinkage over time is likely to be an effect of both fibrosis and atrophy, which are recognised late normal tissue

pathophysiological consequences of radiotherapy. It therefore raises the question whether breast shrinkage reported by patients may be the most appropriate endpoint for assessing dose-volume response within breast radiotherapy trials.

Although breast shrinkage has increased over time, it should be acknowledged in the current era of breast radiotherapy, toxicities which were historically seen such as brachial plexus injury are now rare. Contributions from the RAGE women have resulted in much improved quality and safety of radiotherapy delivery in the U.K [120].

3.5.3 Reporting of breast appearance over time

IMPORT LOW showed that breast appearance change was the commonest reported AE and reporting remained stable over time, with significantly lower rates in the partial-breast group. Both START B [13] and Cambridge IMRT [14] trials also reported change in breast appearance as the most prevalent PRO, which remained stable over time. The cumulative incidence across the 5 year period in these trials was similar to IMPORT LOW: approximately 20% and 18% in START B and Cambridge IMRT respectively. This stable reporting of breast appearance over the 5 year period may reflect the dynamic interaction of some surgical changes resolving whilst some radiotherapy related changes develop over time.

In contrast, different results were found in an interim analysis of the RAPID study [121] testing partial-breast radiotherapy using 3-D conformal radiotherapy versus whole-breast radiotherapy. Patients randomised to whole-breast radiotherapy showed relatively stable reporting of breast cosmesis using the EORTC cosmetic rating system [38] over the 5 year period, whereas those in the partial-breast group reported significantly worse cosmesis, which appeared to increase with time. The reasons for worse cosmesis in the partial-breast radiotherapy group are unclear, but may be related to a higher biologically equivalent dose, especially if incomplete normal tissue repair following twice-daily irradiation is considered [122].

3.5.4 Predictors of adverse effect reporting

The IMPORT LOW analysis showed that certain baseline factors were associated with some patient-reported AE. One of these factors was younger age, within the context of a cohort of peri/post-menopausal women. This observation raised the question whether association was due to biological differences, i.e. differences in breast composition, or perception of AE in the younger age group. Firstly, younger age was only associated with items in the Body Image Scale that relate to patient perception of attractiveness and sexuality as a result of their disease or treatment, whereas there was no significant association between younger age and adverse effects in the questions designed to capture breast radiotherapy effects. In contrast, the Cambridge IMRT trial showed that younger age was associated with increased rates of patient-reported skin changes and breast hardness [14]; however, this trial included women aged less than 50 years so the study population is not entirely comparable with IMPORT LOW.

Previously published results from IMPORT LOW show that there was no significant association found between age and adverse effects reported by clinicians or from photographs over 5 years [15]. Similarly, in the EORTC 'boost versus no boost' trial [54], age was not a predictor of clinician-assessed fibrosis. However, in a small subset of 348 patients, younger age predicted for poorer outcome on digitalised photographs [123]. Of note, the EORTC trial used the contralateral breast to assess photographic outcome whereas in IMPORT LOW the ipsilateral breast at 2 and 5 years was compared with the ipsilateral breast at baseline [15]. The differences in photographic assessment methods along with the EORTC boost trial eligibility criteria allowing women less than 50 years to participate, may contribute to the difference in findings. Taken together, these observations suggest that it is the perception of women at the younger end of the age spectrum within IMPORT LOW that is driving increased body image AE reporting rather than a biologic effect.

Larger breast size was a significant predictor of patient-reported adverse effects within IMPORT LOW. Patients with larger breasts were more likely to report feeling self-conscious about their appearance, dissatisfaction with appearance

when dressed, dissatisfaction with body, and skin changes. In the Cambridge IMRT study, larger breast volume was also a main risk factor influencing patient-reported breast related AE [14].

Larger surgical deficit predicted for increased change in breast appearance and breast shrinkage in IMPORT LOW. In addition, poor baseline surgical cosmesis (related to surgical deficit) predicted for increased skin changes and breast hardness within the Cambridge IMRT trial.

Positive axillary lymph nodes predicted for worse arm/shoulder AE reporting. Similarly, the GEC-ESTRO trial reported worse arm symptom scores (EORTC QLQ-BR23 subscale) in patients who underwent axillary node dissection receiving whole-breast and partial-breast radiotherapy at baseline, and in the whole-breast group at 3 and 6 months following radiotherapy [118].

In IMPORT LOW, 23% and 8% of patients were identified as at high risk (i.e. having borderline and case levels) of anxiety and for depression respectively, from baseline HADS subscale scores and these women were more likely to report AE. Anxiety predicted for almost all AE, whereas association with depression was not consistently statistically significant, possibly due to a small number of patients identified at high risk of clinical depression. Baseline prevalences of high risk anxiety and depression were higher in the START trials (32% and 12% respectively), but were not investigated as predictors of AE [124]. The lower prevalence of baseline anxiety and depression in IMPORT LOW may be partly due to the different eligibility criteria, being of lower risk with fewer patients receiving adjuvant chemotherapy, older, or other unknown factors. It has been reported that pre-treatment psychological status may affect perception of cosmetic outcome from breast conserving surgery and radiotherapy [49]. A small non-randomised study of 60 patients examined the relationship between psychological status at breast cancer diagnosis and subsequent ratings of cosmetic outcome following breast conserving surgery and adjuvant radiotherapy and found that both psychological distress and impaired quality of life were related to patients' ratings of poorer cosmesis [49].

3.5.5 Limitations

Limitations include lack of baseline data regarding patient smoking, co-morbidity, post-operative breast infections and seromas, as effects of these factors were not proven to be associated with AE during the set-up of the trial. In addition, there is a possibility of reporting bias as patients were not blinded to treatment allocation. As has already been described in the previous chapter, there is an inherent risk of informative censoring with PRO questionnaire return, for example patients with certain baseline characteristics may be more or less likely to return questionnaires. In IMPORT LOW there were no significant differences in the majority of baseline characteristics in those who did or did not return questionnaires at 5 years, with the exception of higher baseline HADS anxiety and depression subscale scores in those who did not return their year 5 questionnaire. Bias may also arise due to patients who have worse AE being more or less inclined to report or representing a different sub-population. In IMPORT LOW it was found that patients who reported at least one AE at 2 years were more likely to return questionnaires at 5 years, therefore it is possible that the prevalence of AE is overestimated in this analysis. Finally, the IMPORT LOW trial was conducted in a lower risk population and therefore may not be generalisable to all patients with early breast cancer.

3.5.6 Potential implications for practice

These results demonstrate that the majority of AE reported reduce over time. This information can provide reassurance for patients considering either whole-breast or partial-breast radiotherapy using moderately fractionated IMRT. Furthermore, baseline factors which predict adverse effects may be considered before radiotherapy and contribute to the informed discussion and shared clinician-patient decision-making process. For example, a larger-breasted woman predicted to be at very low risk of recurrence, can receive a more tailored discussion regarding her risks and benefits of radiotherapy and options such as endocrine-only treatment could be discussed where appropriate. In addition, routine psychological assessment and resources for psychosocial support could be made available to patients with suspected higher levels of anxiety and

depression at the start of treatment, as this study shows that this patient group may have long-term repercussions in terms of adverse effect reporting. However, it is unknown whether early psychological intervention would influence reporting rates of AE. Finally, this comprehensive serial analysis of PRO adds further support to the hypothesis that partial-breast radiotherapy using moderately fractionated IMRT has less toxicity compared with whole-breast irradiation.

3.6 Conclusions

These results can provide reassurance for future patients receiving either whole or partial-breast moderately-hypofractionated radiotherapy that treatment sequelae usually improve over time with more than half of patients reporting no moderate/marked AE at 5 years. Furthermore, patients receiving partial-breast radiotherapy report fewer adverse effects compared with whole-breast using this technique. In addition, baseline factors that predict adverse effects can be assessed before radiotherapy allowing tailoring of risk-benefits discussion for individuals.

Chapter 4 Is breast seroma after tumour resection associated with patient-reported breast appearance change following radiotherapy?

4.1 Abstract

Background: Seroma describes a collection of serous fluid within a cavity, occurring following surgery. Seroma is associated with normal tissue effects (NTE) following breast radiotherapy, as reported by clinicians and on photographs. This study investigates the association between seroma and the NTE breast appearance change collected using *patient*-reported outcomes (PRO) in IMPORT HIGH, as well as investigating the association between breast appearance change and patient/tumour/treatment factors.

Methods: Case-control methodology was used for seroma analysis within IMPORT HIGH. Cases were patients reporting moderate/marked breast appearance change and controls reported none/mild changes at year 3. One control was selected at random for each case. Seromas were graded as not visible/subtle or visible/highly visible on CT radiotherapy planning scans. Logistic regression tested associations, adjusting for patient/tumour/treatment factors.

Results: 1078/1149 patients consented to PRO, of whom 836 (78%) reported whether they had 3 year breast appearance change; 231 cases and 231 controls were identified. 304/462 (66%) patients received chemotherapy. Seroma prevalence was 20% (41/202) in cases and 16% (32/205) in controls, and less frequent in patients receiving adjuvant chemotherapy [10% (24/246) compared with 29% (40/138) without]. Visible seroma was not significantly associated with breast appearance change [OR 1.38 (95%CI 0.83-2.29), p=0.219]. Larger tumour size, haematoma, current smoking and body image concerns at baseline were independent risk factors for breast appearance change.

Conclusions: Seroma was not associated with patient-reported breast appearance change, but haematoma and smoking were significant risk factors. Lack of association may be related to lower prevalence of seroma compared

with previous reports, perhaps reflecting patients receiving adjuvant chemotherapy in whom seroma resolves prior to radiotherapy.

4.2 Introduction

Seroma formation describes the collection of serous fluid within a cavity and has been reported following breast surgery. Seroma prevalence of 37% and 57% was reported in the Cambridge IMRT [51] and FAST [52] trials respectively. Seroma has been associated with increased rates of post-operative infection and haematoma, and has been reported to be an independent risk factor for normal tissue effects (NTE) following radiotherapy [51].

An association between seroma and NTE has been reported in the RAPID [53] and Cambridge IMRT trials [51]. The mechanisms by which seroma may lead to NTE following radiotherapy are unknown. As well as fibrosis and retraction of the seroma cavity being possible contributing factors [50], seroma leading to larger volumes receiving radiotherapy boost doses should also be considered. In the EORTC 'boost versus no boost' trial there was an increased risk of fibrosis in those patients receiving a boost [54] and this risk was further increased in patients with a seroma. However, this was significant on univariate analysis only.

The majority of these trials used clinician assessments of NTE and/or serial photographs. Patient-reported outcomes (PRO) provide an opportunity to understand the patients' own perception of NTE and studies have found that patients report more NTE compared with clinicians and those detected on photographs [47,125]. However, the association between the presence of seroma and *patient*-reported NTE following breast radiotherapy has not been investigated to date.

This analysis from IMPORT HIGH uniquely combines comprehensive PRO data with presence/absence of seroma whilst accounting for other patient, tumour and treatment factors. The primary aim of this study was to determine whether seroma is associated with patient-reported breast appearance change following breast radiotherapy. The secondary aim was to investigate associations between other patient/tumour/treatment factors and patient-reported breast appearance change.

4.3 Methods

4.3.1 Study population of IMPORT HIGH

The study population consisted of patients recruited to IMPORT HIGH, a randomised, multi-centre phase III trial testing dose-escalated simultaneous integrated boost (SIB) against sequential boost each delivered by intensity-modulated radiotherapy for early-stage breast cancer with higher than average risk of local relapse (IMPORT HIGH Appendix). Women aged ≥18 after breast conservation surgery for pT1-3 pN0-pN3a M0 invasive carcinoma were eligible for IMPORT HIGH. Randomisation was 1:1:1 between 40Gy/15 fractions (F) to whole-breast (WB) + 16Gy/8F sequential photon boost to tumour bed (40+16Gy) [control group], 36Gy/15F to WB, 40Gy to partial-breast + 48Gy (48Gy) in 15F SIB to tumour bed [test group 1] or 36Gy/15F to WB, 40Gy to partial-breast + 53Gy (53Gy) in 15F SIB to tumour bed [test group 2] [40]. The trial was initiated with a primary endpoint of breast induration at 3 years. However, this was subsequently amended to a primary endpoint of local recurrence and patient accrual extended accordingly.

IMPORT HIGH was approved by East of England Cambridge South Research Ethics Committee (08/H0305/13) and conducted in accordance with the principles of Good Clinical Practice.

4.3.2 Study design - case-control methodology

For this exploratory analysis, patients' CT planning scans required review for the presence of seroma. Given the large numbers of patients recruited to IMPORT HIGH (>2600 patients), reviewing each patients' scan for seroma would be highly resource intensive. However, by implementing a case-control methodology [126], only a proportion of patients' scans would require review enabling a more resource-efficient trial design. In a case-control study, patients who have developed an event, for example, moderate or marked breast appearance change at 3 years (the cases) are identified and suspected aetiological factors, such as seroma on the CT radiotherapy planning scan, is

compared with patients who do not have moderate/marked breast appearance change at 3 years (the controls).

In this study, the endpoint ‘change in breast appearance’ reported by patients at 3 years was used to define cases and controls. Breast appearance change was selected as the NTE to use in this analysis. This was thought to be a clinically relevant endpoint as it is one of the most commonly patient-reported NTE in previous trials including the IMPORT LOW [15], START-B [13] and Cambridge IMRT [14] trials and persists over time. Patients scored breast appearance change using a 4-point scale of ‘none’, ‘a little’, ‘quite a bit’ and ‘very much’. Cases were defined as patients reporting ‘quite a bit’ or ‘very much’ (interpreted as moderate/marked breast appearance change) with controls reporting ‘none’ or ‘a little’ (interpreted as none/mild breast appearance change). The required number of controls (to equal the number of cases) was selected at random from all available controls. Cases and controls were not matched on known predictors of NTE such as breast size and surgical deficit, as these data were not available for all patients in our dataset, which would have reduced the number of cases and controls for analysis. Also, we wished to investigate associations between potential risk factors for patient-reported change in appearance in addition to seroma, and matching on these would have meant that we could not test them in the analyses.

4.3.3 Assessment of seroma & breast density

Radiotherapy CT planning scans for cases and controls were examined for the presence of seroma. Visualisation and Organisation of Data for Cancer Analysis (VODCA v5.4, Medical Software Solutions GmbH, Hagendorn, Switzerland) software was used to view radiotherapy planning CT scans. Seroma was identified on axial CT images and graded as not visible/subtle or visible/highly visible as per methodology used in the Cambridge IMRT study [51]. Visible seroma was contoured on axial CT slices for each case using a pre-defined protocol from the Cambridge IMRT study [51] and total seroma volume recorded. Seroma contouring was undertaken by one clinical research fellow (IB) who had received training from the Chief Investigator of the Cambridge IMRT study and was blinded to patients’ case-control status.

The clinical fellow (IB) was trained to grade seroma visibility using a series of 20 cases from the Cambridge IMRT study's training repository. The Cambridge IMRT study investigators had been trained using the same 20 cases. IB was asked to review the 20 cases and allocate a seroma visibility score as follows: 1=not visible, 2=subtle, 3=visible, 4=highly visible. This was compared to the Cambridge IMRT group score. 85% (17/20) concordance was achieved with the Cambridge IMRT cases pre-set scoring. There was agreement on the remaining 15% (3/20) cases after further review of the cases by the Cambridge IMRT Chief Investigator.

Breast density was assessed in the contralateral breast using a ranking of 1-4 (1=no or sparse distribution of fibroglandular tissue, 2=small dispersed clusters of fibroglandular tissue, 3=large cluster of fibroglandular tissue and 4=mainly fibroglandular tissue) [52].

4.3.4 Collection of dosimetric data

CT planning scan and dosimetry data were collected prospectively by the Radiotherapy Trials Quality Assurance group (RTTQA) for all IMPORT HIGH patients. Whole-breast planning target volume (PTV) dose volume histograms (DVHs) were identified in VODCA for all cases and controls. Doses were converted into equivalent dose in 2Gy (EQD_2) per fraction using the Withers formula (α/β ratio 3) [127]. An α/β ratio of 3 was used following published data from the FAST and START trials, where α/β ratios were estimated at 2.3-2.6 and 3.5-4.7 respectively [128]. The whole-breast PTV mean and maximum doses (in Gray) for each patient were calculated. The tumour bed clinical target volumes (CTV) (cm^3) were recorded on planning assessment forms (completed at the treatment centres) for all patients.

4.3.5 Collection of PRO data

Within IMPORT HIGH, NTE were assessed using PRO, photographs and annual clinician assessments. All centres were invited to participate in PRO and photographic sub-studies (until sufficient accrual was achieved). All patients at these centres were invited to participate in the PRO and photographic sub-studies until the required sample size for each sub-study was obtained.

PRO were obtained at baseline, 6 months, 1 and 3 years following radiotherapy. Baseline was pre-randomisation (post-surgery, post-chemotherapy where relevant and pre-radiotherapy). PRO included: Hospital Anxiety and Depression Scale (HADS) (scores of 8-10 indicating borderline anxiety or depression, and scores of 11-21 indicating case levels of anxiety or depression [27]); 10-item Body Image Scale where higher scores indicate worse body image [28] and protocol-specific questionnaire items including asking patients to score 'change in breast appearance' [13].

Patients consenting to the PRO sub-study were invited to participate in the photographic sub-study which involved assessments at baseline and year 3. Breast size and surgical deficit were scored on a 3 point scale (small, medium, large) from baseline photographs by a panel of observers blinded to patient identity and treatment allocation [43]. Not all patients in the PRO sub-study consented to photographs.

Information regarding smoking, co-morbidities (including diabetes mellitus, hypertension, cardiovascular disease, collagen vascular disease), antibiotics for tumour bed infection and haematoma were recorded at baseline. Details regarding timing of haematoma or whether the patient had any further surgical intervention for the haematoma were not recorded. Information regarding co-morbidities were collected (following a substantial amendment) 4-years after the trial opened to recruitment.

4.3.6 Statistical Analysis

Logistic regression was used to test associations between visible seroma and patient, tumour and treatment-related factors with moderate/marked patient-reported breast appearance change at year 3, and results summarised using odds ratios (OR, with 95% confidence intervals, CI). Each factor was initially tested in univariate analysis, and those statistically significant ($p<0.05$) were included in a multivariable analysis.

Patient-related factors tested included age, breast size and density, smoking status, comorbidity, levels of anxiety and depression measured on HADS subscales, and Body Image Scale (BIS) score. Tumour and treatment factors tested were tumour size, grade and location, use of chemotherapy, radiotherapy treatment group, tumour bed clinical target volumes (CTV), mean and maximum dose to the whole-breast PTV, axillary lymph node status, axillary surgery, post-operative infection, haematoma, surgical deficit assessed on baseline photograph, presence of visible seroma and seroma volume. For analysis of seroma volume, volume was set to zero for patients without a seroma. The factors described above were clinician-reported with the exception of the patient-reported outcome measures (HADS subscales and BIS score).

With respect to the dosimetry variables tested in the analysis, firstly, the whole-breast PTV volume receiving each dose level was tested on univariate analysis with breast appearance change at 3 years. A statistically significant association was found for each dose level ($p<0.05$ for dose levels from 1-62Gy). Treatment group was not adjusted for in these analyses as treatment group would be highly correlated with the dose level, and this would result in collinearity between dose level and treatment group. As all of the individual dose levels were significantly associated with the NTE outcome, in order to try to select which dose level to take forwards into a multivariable regression, a correlation matrix was developed including dose levels of 10Gy 20Gy 30Gy 40Gy 50Gy and 60Gy. This showed that all of the whole-breast PTV volumes receiving the specified dose levels were highly correlated with each other ($p<0.001$ for each). In IMPORT HIGH the dose levels of 36Gy, 40Gy, 48Gy, 53Gy and 56Gy were the highest prescribed doses across each of the treatment groups. As described above, a correlation matrix was developed using these dose levels and again each of the dose levels were highly correlated with each other ($p<0.001$ for each). As all dose levels were highly correlated we were unable to select a single dose level for multivariable analysis from these results. Therefore summary metrics of mean and maximum dose were used.

All analyses were carried out using STATA version 14 based on a database snapshot taken on June 15th 2018. The IMPORT HIGH trial is registered in the ISRCTN registry (ISRCTN47437448) and ClinicalTrials.gov (NCT00818051).

4.4 Results

IMPORT HIGH recruited 2621 patients from 77 centres. A total of 1078 of the 1149 patients from the 51 centres participating in the sub-study consented to PRO. Year 3 questionnaires were returned by 842/1078 (78%) patients. Of these 842 patients, 836 patients provided a response for breast appearance change at year 3 and 231/836 (28%) reported moderate or marked changes (defined as cases).

4.4.1 Seroma case-control analysis

In this study, 462 patients (231 cases and 231 controls) were identified (table 4.1). Adjuvant chemotherapy was received by 147/231 (64%) cases and 132/231 (57%) controls, and neo-adjuvant chemotherapy by 9/231 (4%) cases and 16/231 (7%) controls. In patients who received adjuvant chemotherapy, the radiotherapy planning scan would have been done approximately 16-20 weeks post-surgery (based on standard UK practice). In patients receiving neo-adjuvant chemotherapy or no chemotherapy, the radiotherapy planning scan would be approximately 4 weeks post-surgery. Radiotherapy planning CT data were available for 407 patients (missing for 29 cases and 26 controls). Reasons for missing data included the inability to retrieve dose files from centres, corrupted dose files, or deviations from trial protocol (i.e. where patients received local standard treatment, CT planning scans and dosimetric data were not collected for these patients). There were no differences in reasons for missing data between cases and controls. Seroma prevalence was 41/202 (20%) in the cases and 32/205 (16%) in the controls. In patients receiving adjuvant chemotherapy for whom seroma data were available, 10% (24/246 patients) had seroma compared with 29% (40/138) in patients not receiving chemotherapy.

Statistically significant patient factors associated with 3 year moderate/marked breast appearance change in univariate analysis included younger age, larger breast size, greater breast density, current smoking, higher baseline HADS anxiety and depression scores and body image concerns at baseline. There was

a large proportion of missing co-morbidity data and therefore these were not tested in univariate analysis, with the exception of cardiovascular disease.
Tumour and treatment factors associated with 3 year moderate/marked breast appearance change in univariate analysis were larger tumour size, post-operative infection, haematoma, larger surgical deficit on photographs, larger seroma volume, larger tumour bed CTV and mean dose (table 4.1). There was no statistically significant association found between visible/highly visible seroma and moderate/marked breast appearance change at 3 years [OR 1.38 (0.83-2.29), p=0.22]. Stratifying by adjuvant chemotherapy use, the odds ratio for the association between seroma and moderate/marked breast appearance change in patients receiving chemotherapy was 2.0 [0.82-4.86), p=0.13] compared with 1.25 [(0.60-2.61), p=0.55] in patients not receiving chemotherapy.

Factors which remained statistically significant in multivariable analysis were larger tumour size, haematoma, current smoking and body image concerns at baseline (table 4.2). The association between seroma volume and moderate/marked breast appearance change was no longer significant in multivariable analysis. As there was a large proportion of missing data for tumour bed CTV (135 patients' missing) and also for breast size and surgical deficit assessed on photographs (170 patients' data unavailable), these factors were excluded from the multivariable analysis. Whole-breast PTV recorded on CT planning scans were used in logistic regression models to represent breast size, but this was not associated with moderate/marked breast appearance change in multivariable analysis (table 4.2).

Table 4.1: Summary of univariate analyses: associations between baseline characteristics and moderate/marked change in breast appearance at 3 years in the case-control population in IMPORT HIGH

Characteristics	Cases [Patients reporting moderate/marked change in breast appearance at 3 years] N=231 (%)	Controls [Patients reporting none/mild change in breast appearance at 3 years] N=231 (%)	Univariate analyses OR (95% CI)	P value
Age years	N=231	N=231	0.98 (0.96-0.996)	0.019
Median (IQR)	49 (45-52)	49 (45-57)		
Treatment group				
Control	84/231 (36)	77/231 (33)	1	
Test group 1	62/231 (27)	74/231 (32)	0.77 (0.49-1.21)	0.258
Test group 2	85/231 (37)	80/231 (35)	0.97 (0.63-1.50)	0.905
Tumour size (cm)	N=231	N=231	1.27 (1.07-1.50)	0.005
	2.1 (1.6-2.8)	1.7 (1.3-2.5)		
Tumour grade				
Grade 1	24/231 (10)	17/231 (7)	1	
Grade 2	99/231 (43)	93/231 (40)	0.75 (0.38-1.49)	0.418
Grade 3	108/231 (47)	121/231 (52)	0.63 (0.32-1.24)	0.182
Lymph nodes				
Positive	77/231 (33)	70/231 (30)	1	
Negative	154/231 (67)	161/231 (70)	0.87 (0.59-1.29)	0.485
Tumour location				
Central	38/230 (17)	29/230 (13)	1	
Upper outer quadrant	106/230 (46)	114/230 (50)	0.71 (0.41-1.23)	0.222
Upper inner quadrant	47/230 (20)	48/230 (21)	0.75 (0.40-1.40)	0.364
Lower outer quadrant	25/230 (11)	24/230 (10)	0.79 (0.38-1.67)	0.543
Lower inner quadrant	14/230 (6)	15/230 (7)	0.71 (0.30-1.71)	0.447
CTV boost volume in cc	N=161	N=166	1.02 (1.00-1.03)	0.008
Median (IQR)	15.4 (7.5-24.6)	11.6 (6.4-18.6)		
Axillary surgery				
No	3/231 (1)	3/231 (1)	1	
Yes	228/231 (99)	228/231 (99)	1.00 (0.20-5.0)	>0.99
Post-op infection				
No	189/231 (82)	207/229 (90)	1	
Yes	42/231 (18)	22/229 (10)	2.10 (1.20-3.63)	0.009
Post-op haematoma				
No	202/231 (87)	219/229 (96)	1	
Yes	29/231 (13)	10/229 (4)	3.14 (1.49-6.61)	0.003
Smoking status				
Never smoker	123/231 (53)	141/229 (62)	1	
Current smoker	41/231 (18)	21/229 (9)	2.24 (1.25-3.99)	0.006
Previous smoker	67/231 (29)	67/229 (29)	1.15 (0.76-1.74)	0.52
Cardiovascular disease				
No	218/229 (95)	210/226 (93)	1	
Yes	11/229 (5)	16/226 (7)	0.66 (0.30-1.46)	0.307
Adjuvant chemotherapy				
No				
Yes	75/231 (32)	83/231 (36)	1	0.433
	156/231 (68)	148/231 (64)	1.17 (0.79-1.71)	
Baseline HADs anxiety				
Normal (0-7)	133/214 (62)	154/218 (71)	1	
Borderline (8-10)	38/214 (18)	46/218 (21)	0.96 (0.59-1.56)	0.858
Case (11+)	43/214 (20)	18/218 (8)	2.77 (1.52-5.03)	0.001
Baseline HADs depression				
Normal (0-7)	167/215 (78)	184/217 (85)	1	
Borderline (8-10)	30/215 (14)	26/217 (12)	1.27 (0.72-2.24)	0.405
Case (11+)	18/215 (8)	7/217 (3)	2.83 (1.15-6.95)	0.023
Body Image Scale*	N=210	N=215	1.06 (1.03-1.09)	<0.001
Median (IQR)	9 (4-15)	5 (1-11)		
Breast Size**^				
Small	57/140 (41)	69/152 (45)	1	
Medium	52/140 (37)	66/152 (43)	0.95 (0.58-1.58)	0.854
Large	31/140 (22)	17/152 (11)	2.21 (1.11-4.39)	0.024
Surgical deficit**				
Small	86/140 (61)	119/152 (78)	1	
Medium	39/140 (28)	28/152 (18)	1.93 (1.10-3.37)	0.021
Large	15/140 (11)	5/152 (3)	4.15 (1.45-11.86)	0.008
Seroma				
No	161/202 (80)	173/205 (84)	1	
Yes	41/202 (20)	32/205 (16)	1.38 (0.83-2.29)	0.219
Seroma volume (cc)	N=198	N=203	#1.21 (1.02-1.44)	0.032
Median (IQR)***	20.3 (6.8-46.1)	13.6 (7.4-19.0)		
Breast density~				
Rank 1	88/201 (44)	70/204 (34)	1	
Rank 2	51/201 (25)	57/204 (28)	0.71 (0.44-1.16)	0.175
Rank 3	50/201 (25)	51/204 (25)	0.78 (0.47-1.29)	0.33
Rank 4	12/201 (6)	26/204 (13)	0.37 (0.18-0.78)	0.009
Mean dose in Gray	N=192	N=197	1.08 (1.02-1.14)	0.009
	45.1 (43.2-49.2)	44.0 (42.4-48.6)		
Maximum dose in Gray	N=192	N=197	1.01 (0.97-1.05)	0.532
	66 (65-74)	66 (65-74)		

*IQR=interquartile range * Higher scores for body image scale indicate more problems (possible range 0-30). **Breast size and surgical deficit scored on baseline photographs (data not available for all patients as all patients in PROMs sub-study did not participate in the photographic sub-study). ***For seroma*

volume, patients without seroma included in analysis with zero volume. # Seroma volume assessed per 10cc. ~Data from 2 patients missing due to inability to assess contralateral breast and implants. Rank 1=no or sparse distribution of fibroglandular tissue, 2=small dispersed clusters of fibroglandular tissue, 3=large cluster of fibroglandular tissue and 4=mainly fibroglandular tissue. ^Breast size also assessed using whole breast PTV volume. Data for diabetes mellitus, hypertension and collagen vascular disease not shown as few patients with available data.

Table 4.2: Summary of multivariable analyses: associations between baseline characteristics and moderate/marked change in breast appearance at 3 years

Characteristics	Multivariable analyses Adjusted OR* (95% CI)	P value
Age	0.98 (0.96-1.01)	0.243
Tumour size	1.43 (1.13-1.82)	0.003
Post-op infection		
No	1	
Yes	1.45 (0.68-3.07)	0.335
Post-op haematoma		
No	1	
Yes	5.96 (2.20-16.11)	<0.001
Smoking status		
Never smoker	1	
Current smoker	2.25 (1.06-4.74)	0.034
Previous smoker	1.15 (0.67-1.97)	0.613
Baseline HADs anxiety		
Normal (0-7)	1	
Borderline (8-10)	0.70 (0.37-1.32)	0.273
Case (11+)	2.17 (0.97-4.87)	0.06
Baseline HADs depression		
Normal (0-7)	1	
Borderline (8-10)	0.90 (0.42-1.93)	0.778
Case (11+)	1.93 (0.53-6.99)	0.317
Body Image Scale	1.04 (1.00-1.09)	0.044
Whole Breast PTV volume	1.00 (0.99-1.00)	0.226
Seroma volume	1.01 (0.99-1.04)	0.209
Breast density		
Rank 1	1	
Rank 2	0.63 (0.34-1.16)	0.134
Rank 3	0.86 (0.44-1.68)	0.662
Rank 4	0.41 (0.16-1.08)	0.07
Mean dose to whole breast in Gray	1.05 (0.98-1.13)	0.19

** Odds ratios adjusted for all variables shown in the table. Rank 1=no or sparse distribution of fibroglandular tissue, 2=small dispersed clusters of fibroglandular tissue, 3=large cluster of fibroglandular tissue and 4=mainly fibroglandular tissue*

4.5 Discussion

These results show, within IMPORT HIGH, that there was no significant association between seroma and patient-reported breast appearance change at 3 years. However, haematoma, larger tumour size, current smoking and body image concerns at baseline were significant risk factors. In contrast to our findings, the Cambridge IMRT study comparing 2-dimensional radiotherapy against forward-planned IMRT using 40Gy in 15 fractions in both treatment groups, found a significant association between seroma and inferior cosmesis on photographs at 5 years [OR=1.8, (95%CI 1.0-3.4), p=0.05] [51]. Juneja et al also showed an association between seroma and breast appearance change on photographs at 2 years [OR 3.44, (95%CI 1.28-9.21), p=0.01] in the FAST-Pilot (patients received 30Gy in 5F over 15 days) and UK FAST trials (randomising to 50Gy in 25F versus 28.5 or 30Gy in 5 once weekly fractions) [52].

The lack of association between seroma and patient-reported breast appearance change may be related to the low overall prevalence of seroma within the case-control study in IMPORT HIGH: 20% in the cases and 16% in the controls. Clinically, this is lower than the 37% seroma prevalence reported in the Cambridge IMRT study [51]. It is also lower than the 57% seroma prevalence reported in a case-control study using patients from the FAST-Pilot and UK FAST trials [52].

Reasons for the lower prevalence of seroma in IMPORT HIGH may be due to a larger proportion of patients receiving chemotherapy (potentially resulting in seroma resolving prior to radiotherapy) and changes in surgical practice over time. The Cambridge IMRT and FAST trials recruited between 2003-2007, whereas IMPORT HIGH recruited from 2009-2015. In the Cambridge IMRT seroma study, 122/648 (19%) patients received chemotherapy [51] compared with 304/462 (66%) patients in our case-control study in IMPORT HIGH. In the patients receiving adjuvant chemotherapy in IMPORT HIGH (with a time lag of approximately 16-20 weeks from surgery to radiotherapy planning scan), 10% (24/246 patients) had seroma compared with 29% (40/138) in patients not

receiving chemotherapy. One study demonstrated that seroma volume decreases with a longer time interval from surgery to radiotherapy [129].

Chemotherapy was also considered a potential confounder in IMPORT HIGH. However, in our study, adjusting for adjuvant chemotherapy use made little difference to the estimate of association between seroma and breast appearance change. Nevertheless, seromas persisting after chemotherapy may be more stable during radiotherapy such that dosimetric heterogeneities within the tumour bed region incurred by fluctuating seroma volume will be minimised. In addition, seromas persisting following chemotherapy may maintain volume within the tumour bed such that any distortion associated with their resolution may be less likely.

Surgical practices have changed since the FAST and Cambridge IMRT trials were conducted, from leaving the excision cavity open (which may be associated with seroma formation) towards primary closure of the defect by either direct suturing of cavity walls together, local glandular mobilisation or therapeutic mammoplasty. In patients who develop a seroma in an open cavity, fibrosis and retraction of tissue surrounding the excision cavity (following seroma reabsorption) could result in a noticeable defect [50]. In contrast, there is also evidence to suggest that the seroma cavity may not always contract and new tissue may be laid down in concentric rings [130]. With increasing use of oncoplastic surgery to redistribute breast tissue into locations of volume loss particularly in those requiring extensive resections, rates of seroma are likely to have reduced. One study reported significantly lower rates of seroma in patients undergoing oncoplastic surgery compared with standard breast conserving surgery: 1.7% versus 4.4%, p=0.04 [131], albeit that seromas were diagnosed clinically in this study and thus rates were lower than described in the radiotherapy literature.

It is possible that our study was underpowered to detect a moderate effect of seroma; with around 200 cases and controls the study had 78% power to detect an odds ratio of 2, based on 16% seroma prevalence in our control population (alpha=0.05). Although there was no significant association between seroma

and breast appearance change, greater seroma *volume* was associated with breast appearance change on univariate analysis. For the analysis, seroma volume was set to zero for patients without seroma. Limited patient numbers with seroma may have contributed to the lack of significance on multivariable analysis, or it may be that the association between seroma and NTE is weaker than previously reported. The RAPID trial testing partial-breast radiotherapy using 3D conformal radiotherapy versus whole-breast radiotherapy reported an association between seroma volume and adverse cosmesis at 3 years [OR=1.35 (1.11-1.65), p=0.004] [53].

The choice of endpoint used in our case-control study may also explain our results (seroma not being associated with NTE) being different to those of other published studies. PRO provide the patient-perspective of side-effects and it has been shown that patients report a higher prevalence of NTE compared with clinicians or photographs [47,125]. Therefore, PRO may be a more sensitive endpoint. Furthermore, patients experiencing a large palpable seroma at baseline may be more perceptive of future NTE compared with clinician or photographic scoring (where prior seroma may not be noted). Greater volume of seroma was associated with 3 year breast appearance change in IMPORT HIGH.

With respect to other tumour and treatment factors, haematoma was significantly associated with breast appearance change within IMPORT HIGH. Similarly, haematoma predicted moderate/severe fibrosis in the EORTC 2281-10882 'boost versus no boost' trial [HR 1.80 (95%CI (1.32-2.47), p<0.0001] [54]. Post-operative haematoma leading to worse cosmetic outcome may be related to glandular necrosis. Larger tumour size was also significantly associated with breast appearance change. Tumour size may be a proxy measure for surgical deficit. Larger surgical deficit at baseline predicted patient-reported breast appearance change in IMPORT LOW [100]. Also, larger excision volumes were associated with poorer cosmetic outcome in the EORTC 'boost versus no boost' trial [123]. With regard to patient factors, current smoking was strongly associated with patient-reported breast appearance change in IMPORT HIGH. Similarly in

the RAPID trial, smoking was associated with adverse cosmesis [OR 2.42 (95%CI 1.56-3.75), p<0.001] and a deterioration in cosmesis over 3-years [OR 1.58 (95%1.01-2.46), p0.04] [53]. Smoking has been associated with impaired wound healing, post-operative complications and increased radiation toxicity [132,133]. Finally, body image concerns at baseline were also significantly associated with breast appearance change. Items in the BIS relate to patient perception of attractiveness and sexuality as a result of their disease or treatment. This association has not been previously investigated or reported in the literature.

4.5.1 Implications of findings

We were unable to show an association between seroma and patient-reported breast appearance change. However larger tumour size, haematoma, current smoking and body image concerns at baseline were independent risk factors. This suggests that measures should be taken to reduce the risk of haematoma formation. For example, by achieving adequate haemostasis with return of patient blood pressure to normal prior to wound closure and avoidance of post-operative hypertension (e.g. due to pain). Also, smoking cessation should be encouraged, although we cannot determine the time interval required from smoking cessation to start of radiotherapy to reduce the risk of patient-reported breast appearance change.

4.6 Conclusions

In conclusion, seroma was not associated with patient-reported breast appearance change, but haematoma and smoking were significant risk factors. Lack of association may be related to lower prevalence of seroma compared with previous reports, perhaps reflecting patients receiving adjuvant chemotherapy in whom seroma resolves.

Chapter 5 Can the introduction of a patient decision aid video in addition to standard written information reduce decisional conflict regarding a de-escalation of treatment clinical trial? Method development for the PRIMETIME Information Giving Study (IGS)

The PRIMETIME Information Giving Study (IGS) is split into two chapters: Chapter 5 describes the method development for the PRIMETIME IGS and Chapter 6 describes the results of the PRIMETIME IGS.

5.1 Introduction

It has been acknowledged that adjuvant breast radiotherapy following breast conserving surgery (BCS) has a number of benefits and risks. The absolute benefit is dependent on the individual's risk of relapse and can vary substantially for different prognostic risk groups of patients. Improvements in breast cancer care mean that local recurrence rates have fallen substantially over recent decades.

With respect to the risks associated with radiotherapy, commonly, patients can develop late normal tissue effects (NTE) affecting the treated breast including moderate /severe breast shrinkage, pain, tenderness and hardness [2]. Rarely, radiotherapy can also be associated with cardiac toxicity [8] and the development of second cancers [9].

For some patients with very low risk of local relapse, the risks of radiotherapy may outweigh the benefits, and for these patients de-escalation of treatment with omission of radiotherapy may be preferable. The use of modern molecular diagnostics in addition to basic clinicopathological parameters may improve the identification of patients at very low risk of local relapse, to select which patients may be suitable for de-escalation of treatment.

De-escalation of treatment has a number of consequences. Firstly, treatment de-escalation prevents overtreatment and avoids side-effects for patients.

Secondly, de-escalation enables cost savings for the National Health Service. Savings of over £14 million per year have been estimated if radiotherapy is omitted in very low risk patients in the UK [134]. However, for some patients the concept of de-escalation challenges the expectation that 'all efforts are being made to treat the cancer'. Patients may perceive that 'more is better' and clinicians may practice to be 'better safe than sorry' [55]. It has been found that patients often have quantitative misperceptions regarding adjuvant treatment, overestimating the risk of a negative outcome without treatment and overestimating the positive effect of treatment [56 57].

Also, when patients consider clinical trial entry, they will experience a degree of uncertainty. This uncertainty may be increased in de-escalation clinical trials, where a component of standard treatment may be removed. The uncertainty patients experience regarding healthcare decisions is known as decisional conflict. Strategies are required to reduce decisional conflict to ensure patients are comfortable about their decisions and optimise the decision-making process.

Understanding and communicating the risks and benefits of treatments remains a challenge for both patients and clinicians. Patient advocates now routinely have a central role in the development of information materials and this is especially important for information provided in de-escalation clinical trials. As well as presenting information in language which is understood by patients, patient advocates are able to draw on their own experiences and advise which are the essential concepts for patients to understand when considering a de-escalation clinical trial.

In general, when discussing risk with patients, presenting absolute risk has been advocated to be preferable to the relative risk, as absolute risk describes how likely an event is to a patient themselves or in a defined patient risk group [61]. In contrast, relative risk describes how much more or less likely an event is between two groups of patients and is more difficult for patients to translate to their own situation. When presenting information on absolute risks, natural frequency formats (which are numerical values expressed as event rates in groups with and without the intervention) can improve patient understanding [62]. Historically, patient information regarding clinical trials has been provided to

patients in written format. However, certain concepts such as the explanation of risk of recurrence may be more easily understood if presented in a different format, for example using a video format. Patient decision aids (PDA) are tools which help patients to understand the risks and benefits of treatment options consider the value patients place on the risk-benefit ratio and participate actively with clinicians in deciding treatment options. It can be considered that standard patient information sheets are a very simple form of a PDA although in general PDA have been developed to either build on standard patient information or provide additional information. PDA have been developed in various formats including written booklets, software programmes and videos [89]. There is evidence that suggests PDA may reduce decisional conflict [85 86].

Whether PDA can reduce decisional conflict for patients considering entry into de-escalation clinical trials can be assessed using a 'Study Within A Trial' or SWAT concept. A SWAT is where a research study is embedded within a larger clinical trial and enables us to assess different ways of designing, conducting, analysing and evaluating studies through the conduct of research within research [93]. It allows multiple questions to be answered efficiently using data from one group of patients and trialists, without the need to set-up and conduct multiple clinical trials which is a resource-intensive process. Aspects of clinical trials previously assessed using the SWAT concept include patient recruitment and patient information provision [93,94].

PRIMETIME is a biomarker directed prospective cohort study aiming to identify a group of breast cancer patients who can safely avoid adjuvant breast radiotherapy following BCS (PRIMETIME appendix). The biomarker IHC4+C (incorporating Ki-67) is used to determine the patient's recurrence risk [135]. Patients found to be at very low risk are directed to avoid radiotherapy, and patients at low, intermediate or high risk are directed to receive radiotherapy [16]. PRIMETIME provides an opportunity to embed a SWAT testing whether a PDA can reduce decisional conflict in patients considering a de-escalation clinical trial. The PRIMETIME study's Ethics approved standard information for patients consists of the PRIMETIME main study patient information sheet which includes diagrams developed in close collaboration with the patient advocates. Building on this, the explanation of risks and benefits of radiotherapy may be further

improved by delivering information using video format. Thereby, a PDA video was developed to build on the standard written information. The PRIMETIME SWAT named the PRIMETIME Information Giving Study (IGS), will test whether a PDA video in addition to standard written information can reduce decisional conflict in patients considering entry into the PRIMETIME study.

5.2 Methods

5.2.1 Method development for the PRIMETIME IGS

5.2.2 Study design: eligible population for the PRIMETIME IGS

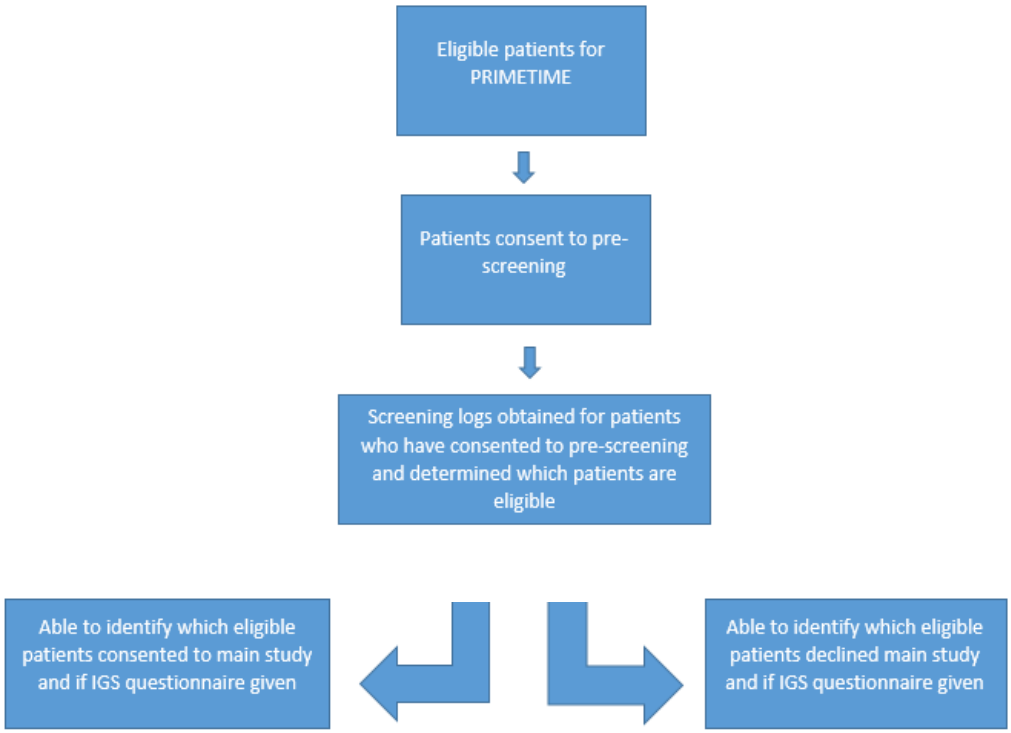
The population of patients eligible for the PRIMETIME IGS were those patients eligible for the main study. It was important that all eligible patients were offered participation in the IGS, and not only those patients who consented to the PRIMETIME main study. The reason for this was that if we were to include only patients in the PRIMETIME IGS who had consented to PRIMETIME, these patients may be characteristically different to those patients who declined PRIMETIME thereby the results obtained would not be unbiased.

Capturing patients who decline a clinical trial may be challenging as there is often not a clear pathway of follow-up for these patients. More often than not, dedicated research staff do not have further contact with patients who receive standard of care outside of a research setting. Also, the coordinating clinical trials units will only know whether a patient has declined a clinical trial via screening logs completed by sites. However, these are usually not recorded 'real-time' and are often incomplete. Furthermore, data in screening logs are anonymised such that the coordinating clinical trials unit are not able to identify the patients who decline clinical trials as the trials unit has no right to know who these patients are, whereas the details for patients who do consent are provided in the screening logs.

For the IGS, the importance of distributing decisional conflict questionnaires to *both* patients who consented *and* patients who declined PRIMETIME was emphasised to sites. The pathway of when patients were given trial information *and* questionnaires was discussed with every site to establish clear pathways to

distribute questionnaires at the correct time. Patient eligibility for the PRIMETIME main study and thereby the IGS was determined via the screening logs (figure 5.1). All patients who were eligible for the PRIMETIME main study would have consented to the pre-screening stage and allocated a pre-screening ID. For patients with a pre-screening ID it was possible from the screening logs to determine which patients were eligible for the main study, and of those patients, who consented or declined the main study. Sites were encouraged to return accurately completed screening logs in a timely manner to enable trials unit staff to determine which patients were eligible for the PRIMETIME main study, and should have been given questionnaires.

Figure 5.1: Eligibility for the PRIMETIME main study and IGS determined via screening logs which documented patient eligibility following pre-screening and which patients consented to, or declined the PRIMETIME main study



It was required that all patients who were offered the PRIMETIME IGS were able to read and write English independently to ensure that patients understood the information provided and were able to score their decisional conflict without any influence from others which may occur if a translator or family member was

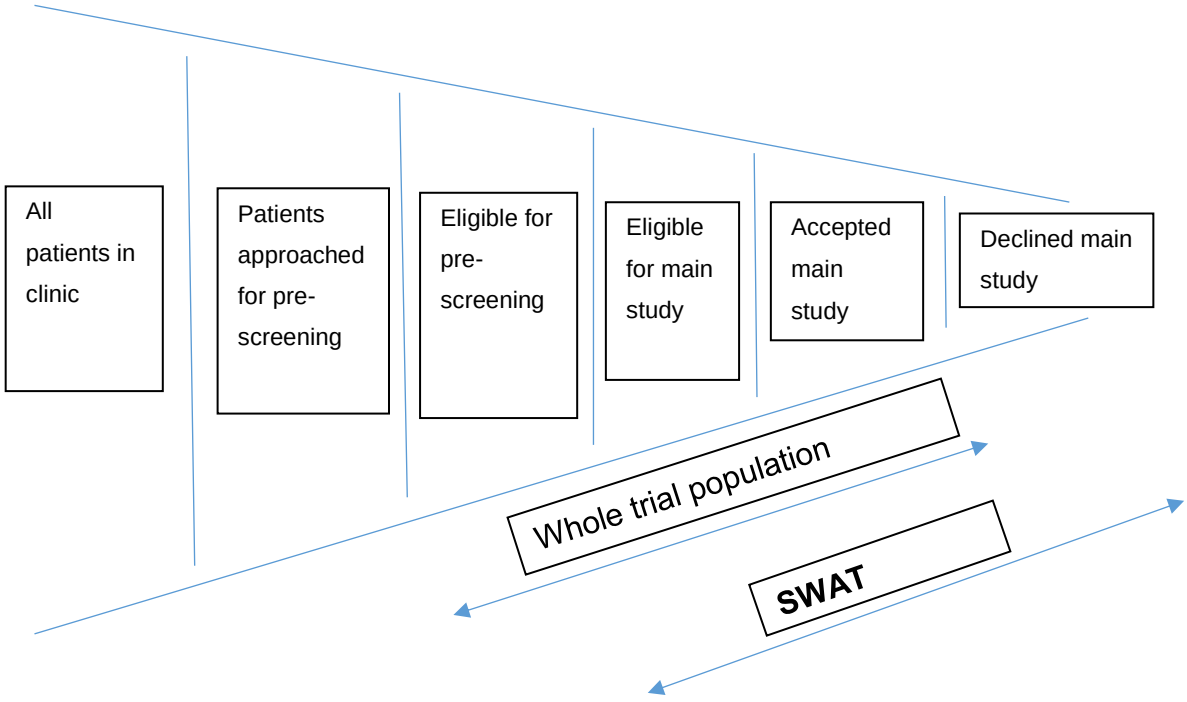
involved in the process. Due to resource constraints it was only possible to provide information in English.

5.2.3 SWAT design

Depending on the aim of a SWAT, it may encompass a broader patient group than that entered into the main trial, including those who have declined entry to the main trial (figure 5.2). Patients who decline clinical trials may have different views on the clinical trials process, and it is important that patients who decline clinical trials are approached and given the opportunity to participate in other studies albeit with a separate consent process.

For the PRIMETIME IGS, *all* sites participating in the main trial also participated in the IGS. As described above, those patients who were eligible for the PRIMETIME main study were eligible for the IGS. The denominator of potentially eligible patients for the IGS consisted of those patients who had consented to the pre-screening phase and went on to be approached and eligible for the PRIMETIME main study, including patients who consented *and* those who declined the main study (figure 5.2).

Figure 5.2: Funnel diagram demonstrating the eligible population of the population of patients in the main PRIMETIME trial was different to that of patients eligible for the PRIMETIME SWAT



5.2.4 Trial design considerations for the SWAT: Individual patient versus cluster level randomisation

When testing an intervention, the intervention can be tested either at an individual patient level or a cluster level. Individual patient level randomisation is where individual patients would be randomised to receive an intervention, whereas cluster level randomisation describes groups of individuals randomised to receive an intervention. If individual patients at a site are randomised to either receive or not receive an intervention, then patients who receive the intervention may discuss the intervention with patients receiving standard of care (unless the intervention is blinded). This could potentially compromise the interpretation of the results of the intervention. In contrast, all patients within a cluster will receive the same, either standard of care, or intervention.

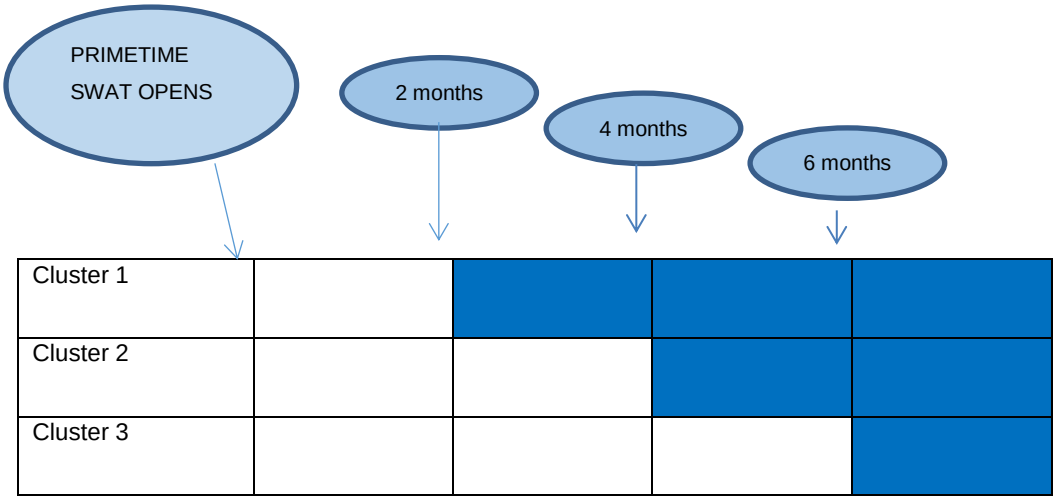
For the purposes of the IGS, when testing an intervention such as a PDA which cannot be blinded to either patients or clinicians, cluster level randomisation was preferred as this ensured all patients at a site received the same information and all healthcare professionals presented the trial in a uniform manner to all the patients at the site. For the IGS, a cluster was defined as the radiotherapy centre and any peripheral sites referring into that radiotherapy centre. The rationale for this was that generally in UK radiotherapy centres, the same clinical oncologist will see patients in referring centres and the radiotherapy centre. The effect of the introduction of a PDA on healthcare professionals should also be considered, as using the PDA may alter the manner in which a healthcare professional describes the study verbally. Therefore, it was preferred that patients within the same cluster received the same information to avoid cross contamination and enable an accurate interpretation of decisional conflict.

5.2.5 Determination of cluster design: Stepped-wedge versus parallel cluster designs

Cluster trials can be implemented using various designs including the stepped-wedge trial design. The stepped-wedge design consists of the sequential implementation of an intervention to participants grouped within clusters over a number of time periods (figure 5.3) [136]. There is an initial period where no clusters receive the intervention (control period). Then, at regular intervals described as ‘steps’, one cluster is randomised to cross over from control to

intervention. This process continues until all clusters have crossed over from control to the intervention, and by the end of the study, all clusters will receive the intervention. Data is collected throughout the study, so each cluster contributes observations from the control and intervention observation periods [137]. Outcome data is derived from single measurements from individual participants, but different participants at each step in the study. An important consideration for the stepped-wedge design is that more clusters will receive the intervention towards the end of the study compared with earlier stages which means that any effect of the intervention may be confounded by calendar time [137]. In addition, if there is potential for a 'learning curve' to affect the evaluation of an intervention in which centres' growing experience in running a trial could in itself reduce decisional conflict, the learning curve would have a greater impact on the control period, as this occurs towards the beginning of the trial, when sites are establishing trial processes. Methods to investigate the potential impact of such a learning curve on the evaluation of the PDA in the IGS are discussed in the 'method of primary outcome analysis' section below. Finally, with any cluster stepped-wedge trial design, the number of clusters, the number and lengths of steps and the number of clusters randomised at each step need to be determined.

Figure 5.3: Summary of cluster stepped-wedge trial design where each cluster begins in the control group (e.g. does not have the PDA video) and at pre-specified time points crosses over to receive the intervention (e.g. receives the PDA video)



Key: white boxes - patients receive standard patient information, blue boxes - patients receive standard patient information and video.

Another cluster level trial design which may be considered is the parallel design where clusters are randomised to receive either standard of care or the intervention up-front. As such, there is no cross-over for sites and the parallel design avoids the confounder of calendar time. However, the number and sizes of the clusters will need to be determined. An important consideration with any cluster trial is the homogeneity within a cluster (or correlations between individuals in the same cluster) [137]. This is the intra-cluster correlation (ICC). If the ICC is thought to be small, then a parallel design may deliver higher statistical power. However, if the ICC is large the stepped-wedge design will have more statistical power. Therefore, if the ICC is large, a smaller sample size of patients will be required with a stepped-wedge versus parallel design [137]. However, the ICC may not be known before a trial is conducted. Nevertheless, there are challenges in implementing any cluster trial within the context of a recruiting trial as the total number of clusters and number of patients per cluster are unknown.

A cluster stepped-wedge trial design was chosen for the PRIMETIME IGS. The reasons for this include, that firstly, it was envisaged that the PDA would reduce decisional conflict and **all** clusters would receive the PDA by the end of the trial. Secondly, the ICC was not known during the IGS set-up, but if the ICC was found to be large, then a smaller sample size of patients were required enabling the study to be conducted more efficiently given limited 3 year PhD timeframes. As mentioned above, when designing a cluster randomised stepped-wedge trial, the number of clusters, the number and lengths of steps and the number of clusters randomised at each step need to be determined. As the PRIMETIME IGS was set-up early on after PRIMETIME had opened to recruitment, the final numbers and sizes of clusters were unknown and estimates were formed based on centres' projected recruitment figures. Again, in order to enable an efficient trial design within the PhD timeframes (36 months), a limited number of steps (3 steps) with short time lengths (each 2 months long) were used.

The methodology for the IGS was submitted to 'The Northern Ireland Hub for Trials Methodology Research SWAT repository', underwent peer review and was subsequently published online [138].

5.2.6 Development of the intervention

5.2.7 Working group to develop the PDA established

A working group was set up consisting of the two PRIMETIME patient advocates (Lesley Stobart and Hillary Turner) from the Independent Cancer Patients' Voice (ICPV), clinical research fellow (Indrani Bhattacharya), trial manager (Natalie Atkins) and senior trial manager (Lisa Fox) to develop the PDA in liaison with the PRIMETIME Chief Investigator/CI (Charlotte Coles) and Scientific Lead (Judith Bliss). The first task was to determine the format of the PDA.

5.2.8 Rationale for development of PDA in video format

Prior to the development of the IGS, all PRIMETIME patients received standard written information which consisted of the PRIMETIME main study patient information sheet and diagrams. This standard written information was developed by the PRIMETIME team in collaboration with the PRIMETIME patient advocates named above. The PRIMETIME main study patient information sheet was designed for the patient to read independently whilst the diagrams were designed to be discussed with the healthcare professionals i.e. not for patients to use independently. The patient information diagrams explain the **absolute** risk of recurrence using natural frequency formats. The diagrams show the risk of recurrence in patients at very low, low and intermediate risk as determined by IHC4+C using small images of 100 women (figure 5.4). The diagrams also showed the risk of recurrence in very low risk patients in those who receive radiotherapy and those who do not (figure 5.5).

Figure 5.4: Patient information diagrams explaining the risk of distant recurrence in patients at very low, low and intermediate risk

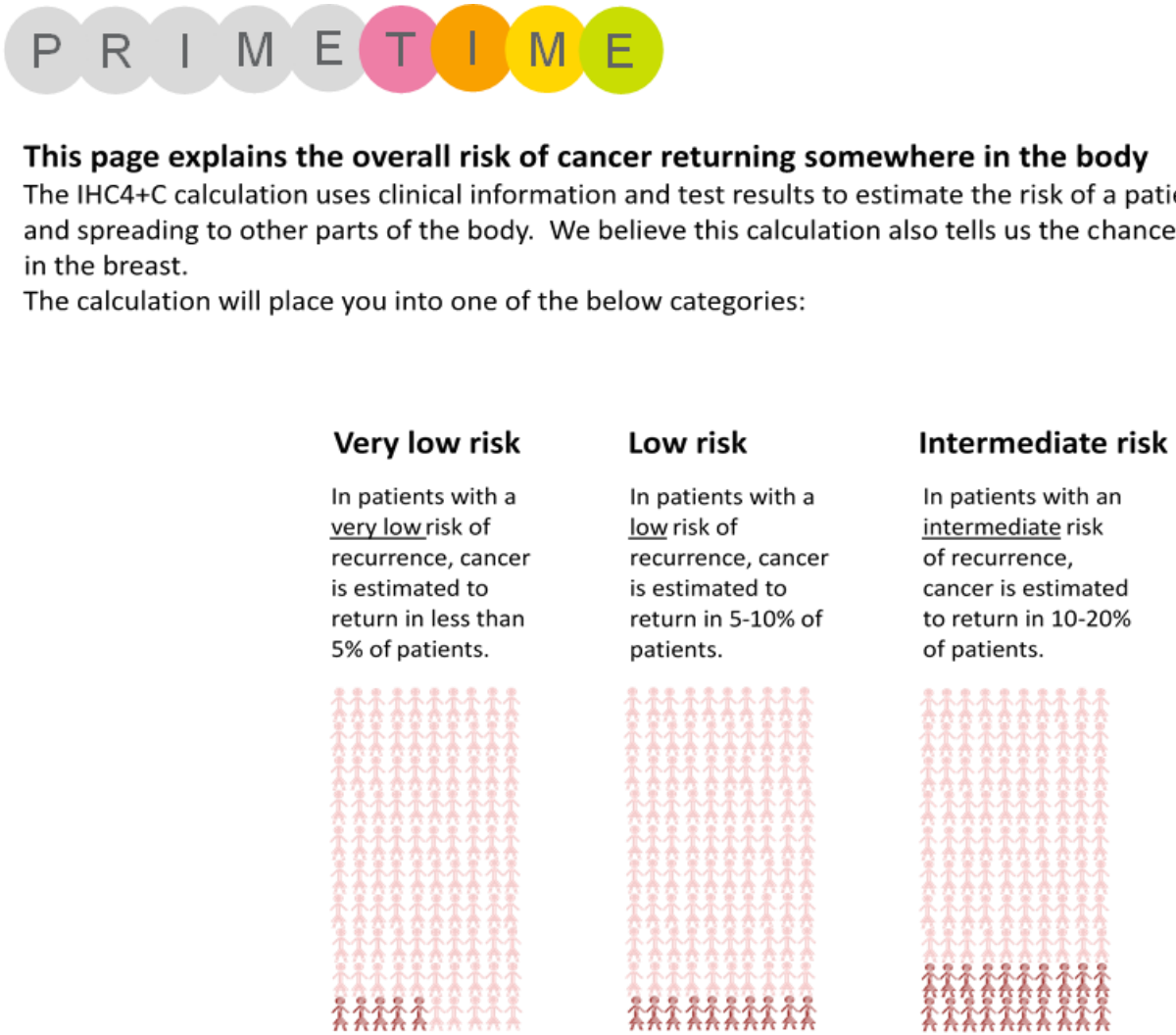


Figure 5.5: Patient information diagrams explaining the risk of local recurrence in patients who do and don't receive radiotherapy in the very low risk group

PRIMETIME

This page explains the risk of cancer returning within the breast, with and without radiotherapy

In standard practice, all patients with the same type of breast cancer as you would receive radiotherapy after surgery. Radiotherapy reduces the risk of breast cancer returning within the breast. For women who have lower risk cancers, the added benefit of radiotherapy has to be balanced against the risk of radiotherapy side-effects.

In the PRIMETIME study, patients in the 'very low' risk group can avoid having radiotherapy. The diagrams below summarise the increased risk of cancer returning in the breast, that this may result in:

'Very low' risk + <u>radiotherapy</u> to the breast

In patients in the very low risk group who have radiotherapy to the breast, cancer is estimated to return in the breast, in 1-2% of patients.

'Very low' risk + <u>no radiotherapy</u> to the breast

In patients in the very low risk group who have no radiotherapy to the breast, cancer is estimated to return in the breast, in 2-3% of patients.

The patient advocates had highlighted that the most important concept for patients to understand in the PRIMETIME study was the concept of risk of recurrence for individual patients. The PRIMETIME patient information sheet and patient information diagrams adequately explained the concept of risk and were Ethics approved. However, the explanation of risk may be more easily conveyed if a different information format is used, for example using a video. Also, the video format can be used independently by patients such that they can watch it at a time convenient for them and without increasing the duration of the clinic consultation.

5.2.9 Focus group meetings to discuss the purpose, content and presentation of the PDA video

A series of meetings with members of the working group took place in which the purpose of the PDA video was established. It was determined that the PDA video should build on the standard written information in PRIMETIME and not be a stand-alone piece of information. Also, it was thought that the PDA video should not contain any additional information i.e. the information **content** of the standard written information and PDA video should be the same. However, the PDA video enabled the concept of 'risk of recurrence' to be presented in a **different format**. It was advised that the PDA video should only be watched after the standard written information had been read.

There were a number of reasons it was thought that no additional information should be provided. Firstly, the patient advocates felt that any additional information would be overwhelming for the patient at a time when they were already receiving large amounts of information early on in their breast cancer diagnosis. Secondly, it was thought that all patients considering a clinical trial should be provided with the same level of information that had already been Ethics Committee approved. This would allow the IGS to be a purer test of only the format of information delivery. As some patients would have two information formats to go through, it was important that the PDA video was short, succinct and ideally last no longer than 10 minutes.

The PDA video concept and proposed content was presented at a patient advocate forum organised by The Institute of Cancer Research's Clinical Trials

and Statistics Unit (ICR-CTSU) at ICR, Chelsea in February 2017 which was attended by a range of patient advocates from ICPV including patients who had been previously treated for breast cancer. Patients had a number of comments and suggestions as to what would be important for them to understand in a trial such as PRIMETIME. These included highlighting the benefits and specific side-effects of radiotherapy, understanding that radiotherapy in low-risk breast cancer did not provide any survival benefit and, if patients did develop a recurrence, that this could be treated radically with surgery +/- radiotherapy. Also, it was felt important to highlight that those patients who did not receive radiotherapy would receive extra mammograms and therefore be monitored more intensively.

The most appropriate presenter of the PDA video was also discussed. Initially it was planned to have a number of presenters on the video including IB, the patient advocates and PRIMETIME CI and Scientific Lead. However, it was felt that having multiple presenters may be distracting for the patient and increase the time length of the video considerably. It was therefore decided that the video should have only one presenter. Although the PRIMETIME PDA and IGS development was led by the clinical research fellow, it was important the presenter of the PDA video was an established researcher with overall responsibility of the PRIMETIME study i.e. the PRIMETIME Chief Investigator. This would help ensure that patients would give appropriate importance to watching the video.

Table 5.1: Summary of key points regarding PDA

• PDA to be developed based on Ethics approved patient information • PDA should not be stand-alone piece of information (PIS should be read before PDA used) • No new information should be included in PDA • Recurrence risk should be explained using same concept as patient information diagrams • PDA should be presented by Chief Investigator • PDA video should be short <10min

5.2.10 Development of PDA video script

Following the focus group meetings and patient advocate forum, the clinical research fellow drafted a script which emphasised key points from the standard written information and points raised from the various meetings with patients. These included explaining the standard treatment for patients with EBC which included discussions of the risks and benefits of radiotherapy. This was followed by an explanation of what is meant by 'very low' risk and the benefit of radiotherapy in this patient group. The PRIMETIME study was then discussed including a summary of the patient journey if they were to enter PRIMETIME. Finally, a description of what happens to the patient if they do or don't have radiotherapy was included. A values-clarification exercise was also developed which enabled patients to weigh up the risks and benefits of radiotherapy.

This script then underwent multiple revisions within the working group and was also reviewed by the PRIMETIME CI, scientific lead and psychosocial oncology advisor (Penny Hopwood). The text and structure were reframed and developed in a question-based format (table 5.2). In particular, certain portions of the text were edited by the CI to wording which she felt more natural for her to say to a patient in the clinic. In the first section, 'Why are we running the PRIMETIME study?' an emphasis was placed on gaining additional information about the patient's tumour in order to individualise treatment. It was highlighted that this information is relevant whether the patient was low or very low risk. In the section detailing 'What do we need to know to work out your risk', the concept of risk recurrence was explained and components of the IHC4+C calculation were described using lay language. The fact that the IHC4+C calculation describes the risk of **distant** recurrence was stated and it was explained that it was being used as a surrogate for local recurrence. The anticipated risk-benefit ratio in patients found to be at very low, low and intermediate risk following the IHC4+C calculation was described and it was explained this was uncertain hence the reason PRIMETIME was being conducted. The standard treatment of EBC i.e. BCS, radiotherapy and hormones (in ER positive disease) was described and the benefits and side-effects of radiotherapy discussed. The fact that the benefit of radiotherapy is dependent on the patient's risk of recurrence but that side-effects were similar in all patients was also explained. The minimal benefit of

radiotherapy in patients with very-low risk of recurrence was discussed. It was advised that patients were still able to make their own decision about treatment even if it was different from the recommended trial treatment and remain within the trial. In patients not receiving radiotherapy, monitoring with extra mammograms was discussed and it was also made clear that, if a patient did have a recurrence, this could be treated with further surgery +/ radiotherapy. Furthermore, the lack of survival benefit in patients receiving radiotherapy versus not receiving radiotherapy was described.

Table 5.2: Summary of PDA video script structure

Script Structure:
1) Why are we running the PRIMETIME study?
2) What do we need to know to work out your risk?
3) What are the benefits of radiotherapy?
4) What are the side-effects of radiotherapy?
5) How do we weigh up the risks and benefits of radiotherapy?
6) Why do we think that patients who are at 'very-low' risk don't need radiotherapy?
7) What happens if I have radiotherapy and what happens if I don't?

The PDA video content was designed acknowledging the International Patient Decision Aid Standards (IPDAS) guidelines for content development which included a systematic development process, provision of information about options and probabilities, clarification of values, disclosure of conflicts of interest, a balanced presentation of options, use of plain language and information based on current evidence [88]. Of note the patient advocates felt strongly that a formal values clarification exercise would be confusing for the patient. Additionally, it was felt that most NHS departments would have insufficient resources to be able to offer patients the opportunity to discuss the exercise after it was completed. Therefore, the values clarification exercise was not included. The final script was submitted for Ethics Approval in May 2017.

5.2.11 PDA Video production

After Ethics Approval was obtained, the video was developed in collaboration with Eyewitness Productions who filmed the CI presenting the script and produced interactive graphics explaining the risk of recurrence and the risks and benefits of radiotherapy based on the standard patient information diagrams. For example, an image of 100 women shaded in grey was used to represent a group

of patients with very low risk of recurrence. The shading of 5 of these 100 women then changed to green to demonstrate the expected numbers of patients who would develop a recurrence in the very low risk group (figure 5.6). This was repeated for the other risk groups. Building on the concept of using natural frequency formats to explain risk of recurrence, the side-effects of radiotherapy were also explained in this way. For example, the risk of a change in breast appearance was shown to be 25 in 100 women who had received radiotherapy (figure 5.7)

Figure 5.6: Summary of the risk of recurrence in very low risk patients

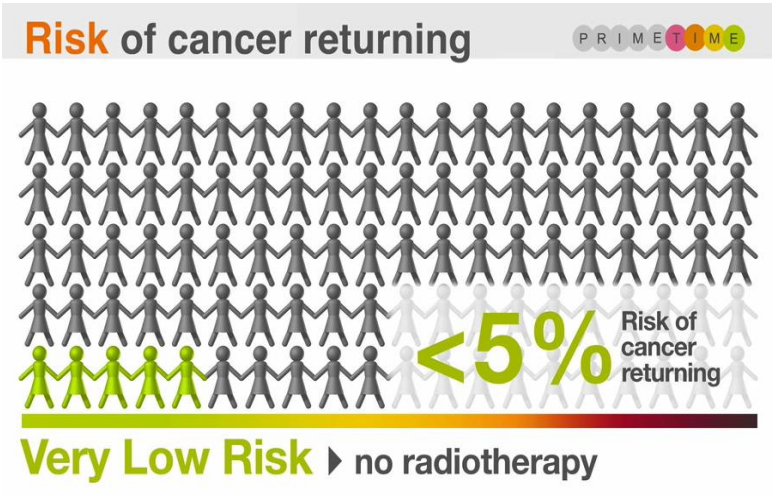
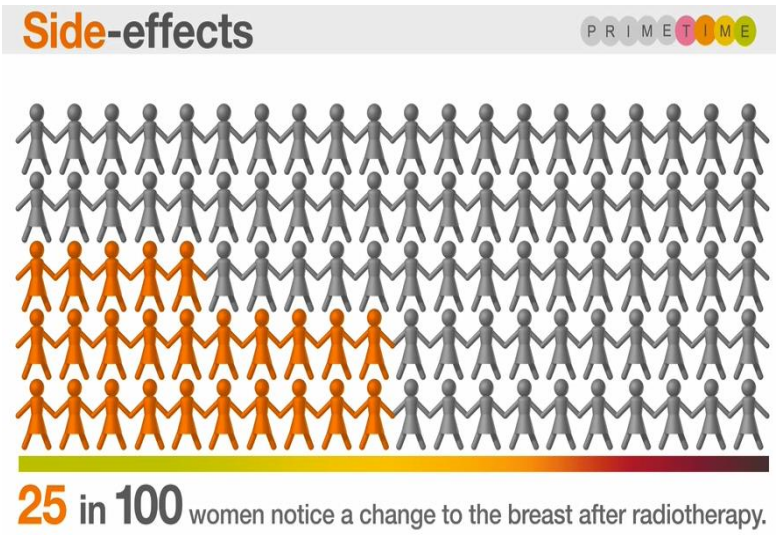


Figure 5.7: Summary of change in breast appearance in women treated with radiotherapy



As well as the text spoken from the script, a number of background images relevant to specific sections were required for the video. The purpose of this was to avoid the viewer losing interest in the video by watching the presenter continuously for the entire duration of the video. For example, when the tumour size and tumour grade were referred to, a tumour specimen was not shown as it was advised by the advocates and psychosocial oncology advisor that this may be distressing for the patient. Instead, histology slides showing tumour grade were briefly shown as a background image only. Also, when radiotherapy was mentioned, an image of a radiotherapy linear accelerator was shown. When the benefits and side-effects of radiotherapy were discussed, an image of a set of scales was shown to symbolise balancing the benefits and side-effects of radiotherapy. Images were obtained by the clinical research fellow and Eyewitness productions from a range of sources including Getty images, ICR communications department, CRUK and from pathologists within the PRIMETIME team. All externally sourced images were copyrighted.

5.2.12 Access to PDA video

Access to the PDA was restricted to the clusters until the cluster was allocated in the intervention group. PDA video access was restricted for both patients and healthcare professionals as there was a chance that healthcare professionals may adapt the way the study was explained based on having watched the video which may affect the decisional conflict results.

The final video was hosted by the ICR and accessible primarily by direct web-link https://www.icr.ac.uk/primetime. Of note, the website was not available by simply searching on any search engine in order to control access. DVDs were provided to sites for patients without internet access. Sites were given access to the PDA video one week before crossing over into the intervention group. Making provision for videos to be watched in the clinic was also discussed prior to the opening of the IGS, however several sites stated this would not be feasible due to limited clinic space. Electronic tablets could have been provided for patients to watch the video in clinic waiting rooms, but financial constraints prevented this.

5.2.13 Ethics approval process for PDA development and PRIMETIME IGS

After the PRIMETIME main study had opened and begun recruitment, the clinical research fellow (IB) liaised with the PRIMETIME Ethics Committee directly regarding obtaining ethics approval for the PDA video and PRIMETIME IGS, within the PRIMETIME main study. IB waited until the main study was successfully recruiting before approaching the Ethics Committee as it was important that the opening and recruitment of the main study was not delayed. The success of the SWAT was dependent on successful implementation and running of the main study.

Following discussions with the Ethics Committee, it was agreed to obtain approvals in a 2-stage process. Firstly, the script for the PDA video was submitted in May 2017. After ethics approval was obtained for the script in July 2017, the video was produced. Secondly, the final PDA video and study design for the IGS was submitted for ethics approval in October 2017. The rationale for this 2-stage approval process was to ensure that the script/video content was ethics approved before resources were placed in video production. The PRIMETIME IGS received ethics approval in February 2018 and opened to recruitment in April 2018.

5.2.14 Planned assessment of outcome - decisional conflict

The primary method for assessing decisional conflict in the literature uses the decisional conflict scale [84]. As this is a validated scale which has been widely used and reported on, the development of a separate questionnaire for assessing decisional conflict was not required for the PRIMETIME IGS.

The secondary endpoints were acceptance of entry into the PRIMETIME main study, and acceptance of the recommended treatment within the PRIMETIME study.

The assessment of decisional regret regarding the patient's decision of whether to participate in the PRIMETIME main study was considered. However, the patient advocates felt that this would potentially be distressing for the patient,

especially for those patients who had not received radiotherapy, and that there would be insufficient resources to manage the patient's distress in these circumstances. Assessment of decisional regret was therefore not included.

5.2.15 Analysis methods for the primary outcome

The primary endpoint of decisional conflict was assessed by an estimated difference in mean decisional conflict scores pre- and post-implementation of the PDA video, reported as a regression coefficient with 95% confidence interval, along with a test for significance (using the z-statistic), and the effect size reported. The coefficient represented a cluster level fixed effect for assigned group, obtained from a multilevel mixed effects linear regression model. To adjust for calendar time and clustering a random effect for cluster and a fixed effect for each step were included in the multilevel model. Robust standard errors were used to adjust for the clustering effect.

5.2.16 Sensitivity analysis to investigate whether a learning curve was present

When testing an intervention such as a PDA which cannot be blinded to either patients or healthcare professionals, there is a possibility that a 'learning curve', where centres' growing experience in running a trial could in itself reduce decisional conflict, rather than the reduction being due to the PDA video implementation. Furthermore, in a stepped-wedge design, the learning curve would affect the control group (i.e. those receiving standard written information) more than the intervention group as all sites will begin in the control group.

In order to investigate whether a learning curve was present, a sensitivity analysis for the primary endpoint was planned, and patients who had returned questionnaires within the first 2 months of their centre having begun recruiting to the IGS excluded.

5.3 Summary

In summary, methods to test whether a PDA video in addition to standard written information in patients considering a treatment de-escalation clinical trial, were developed using SWAT methodology. The eligible population for the SWAT was wider than that of the main trial to ensure the patients in the SWAT were representative of the general population. A cluster randomised trial design was selected to ensure all patients at a site would receive uniform information which would simplify processes for sites and avoid cross-contamination between patients. The stepped-wedge design ensured that all sites would eventually receive the intervention by the end of the trial. The intervention was designed in close collaboration with the patient advocates to explain risk of recurrence and side-effects of radiotherapy in a different format to complement the standard written information. If PDA are found to reduce decisional conflict within treatment de-escalation trials, this will provide the evidence to increase resources into the development of PDAs in future trials. Details of the implementation of the PDA video and SWAT are described in the following chapter, Chapter 6.

Chapter 6 Can the introduction of a patient decision aid video in addition to standard written information reduce decisional conflict regarding a de-escalation of treatment clinical trial? Results of the PRIMETIME Information Giving Study (IGS)

6.1 Abstract

Background: Adjuvant breast radiotherapy has a number of benefits and risks. However, with improvements in breast cancer outcomes, the risk of local relapse has fallen dramatically. For some patients at very low risk of local relapse, the risks of radiotherapy may outweigh the benefits. PRIMETIME is a prospective biomarker-directed cohort study aiming to identify a group of breast cancer patients who can safely avoid radiotherapy following breast conserving surgery. The uncertainty patients face regarding healthcare decisions, including clinical trial participation, is known as decisional conflict. Patient decision aids (PDA) are interventions which help patients to weigh up the risks and benefits of treatments. Evidence suggests PDA reduce decisional conflict. A study within a trial (SWAT) concept which enables trialists to conduct research embedded within a larger trial was used to investigate if the introduction of a PDA video reduces decisional conflict within PRIMETIME.

Methods: The PRIMETIME IGS used a cluster stepped-wedge trial design. Each cluster was defined as the radiotherapy centre and peripheral centres referring into it. All clusters began in the *standard* information group (receiving written information) and were randomised to cross over to the *enhanced* information group (receiving written information and PDA video) at either 2, 4 or 6 months. The primary endpoint was a reduction in decisional conflict following PDA implementation. Decisional conflict was assessed using a validated decisional conflict scale questionnaire (on a scale of 0-100 with greater scores indicating more decisional conflict). The target sample size was 288 patients from 24 clusters, to provide at least 84% power (alpha 0.05) with an effect size of 0.55 (assuming SD=18) across the 0-1 range of possible intra-class correlation values. For the primary endpoint an estimate of the difference in mean decisional

conflict pre- and post-implementation of the PDA video was obtained from a multilevel mixed effects linear regression model. In order to investigate the possibility that a reduction of decisional conflict could be attributed simply to a 'learning curve' in which centres' growing experience in running a trial could in itself reduce decisional conflict (beyond that due to the PDA video implementation), a sensitivity analysis for the primary endpoint was conducted excluding patients who had returned questionnaires within the first 2 months of their centre having begun recruiting to the IGS.

Results: In an interim analysis conducted on 2nd May 2019, 318 evaluable questionnaires were returned from 463 eligible patients (69% return rate) across 24 clusters; 158 questionnaires from the *standard* and 160 questionnaires from the *enhanced* information group. The majority of patients who returned IGS questionnaires had consented to the PRIMETIME study [153/158 (97%) and 155/160 (97%) patients in the *standard* and *enhanced* information groups respectively]. 60/130 (53%) patients in the *enhanced* information group (of those with available data) reportedly did not watch the PDA video.

The mean decisional conflict score in the *standard* information group [158 patients] was 10.73 (standard deviation=11.63) and 8.43 (10.83) in the *enhanced* information group [160 patients]. There was a reduction in decisional conflict in the *enhanced* group compared with the *standard* group, however the effect size was small therefore unlikely to be clinically significant [estimated difference in means = -2.50 (-4.73- -0.28), p=0.03, effect size=0.11]. In the sensitivity analysis when patients who returned questionnaires within the first 2 months of the IGS opening at their centre were excluded, the effect size decreased slightly [-3.01 (-6.72- 0.70) p=0.11, effect size=0.08].

Conclusions: The average decisional conflict scores were low in the PRIMETIME IGS, therefore there was less scope for obtaining further reductions, and the difference observed was not clinically significant. The low baseline scores may reflect the PRIMETIME study population who are at low risk of local recurrence. Almost half of patients reportedly did not watch the video; this may be due to the standard written information being sufficient to fulfil the needs of this patient population. The PRIMETIME IGS was conducted using

a SWAT concept with a cluster stepped-wedge trial design across multiple UK centres participating in the PRIMETIME study. The majority of patients who consented to the IGS also consented to the PRIMETIME main study.

6.2 Background

As previously discussed, adjuvant breast radiotherapy following breast conserving surgery (BCS) has a number of benefits and risks. The absolute benefit is dependent on the individual's risk of relapse and can vary substantially for different prognostic risk groups of patients. Local recurrence rates have fallen substantially over recent decades. This means that for some patients with a very low risk of local relapse, the risks of radiotherapy may outweigh the benefits, and for these patients, de-escalation of treatment with omission of radiotherapy may be preferable. This has led to the introduction of de-escalation of radiotherapy clinical trials.

Patients considering clinical trials will experience a degree of uncertainty. This uncertainty may be increased when a component of standard of care is removed. Uncertainty regarding healthcare decisions is known as decisional conflict. Tools which help patients to understand the risks and benefits of treatment options, consider the value they place on the risk-benefit ratio and participate actively with clinicians in deciding treatment options are patient decision aids (PDA). There is evidence that PDA can reduce decisional conflict [85] [86].

Whether the addition of a PDA video to standard written information reduces decisional conflict in patients considering entry into a de-escalation of radiotherapy trial is being tested within the PRIMETIME avoidance of radiotherapy study using a Study Within A Trial (SWAT) called the PRIMETIME Information Giving Study (IGS). The PRIMETIME IGS method development has been discussed in the previous chapter. This chapter describes the practical implementation and results of the PRIMETIME IGS across the multiple sites participating in the PRIMETIME study.

6.3 Methods

6.3.1 PRIMETIME IGS site initiation

Following ethics approval of the IGS, a PRIMETIME IGS site initiation visit (SIV) was delivered by the clinical research fellow (IB) to sites which were already recruiting to PRIMETIME. The purpose of this PRIMETIME IGS SIV was to train staff at sites for running the IGS. It was an opportunity to establish the patient pathway at each site and to determine at which hospital visit patient information materials would be distributed and decisional conflict assessed. Furthermore, the importance of including all patients who were eligible for PRIMETIME i.e. capturing patients who consent and those who decline PRIMETIME was emphasised.

Patients were included in the IGS one month post SIV. This one month lag period gave sites time to adapt to run the IGS. In those sites which opened to recruitment for PRIMETIME after ethics approval for the IGS, information regarding the IGS was given as part of the PRIMETIME main study SIV.

6.3.2 PRIMETIME IGS implementation of study design at sites

Each cluster was defined as the radiotherapy centre and any centres referring into the radiotherapy centre. All clusters began in the *standard* information group where patients were given the PRIMETIME main study patient information sheet and diagrams. At pre-specified time-points, clusters crossed over to the *enhanced* information group where patients received the PRIMETIME main study patient information sheet, diagrams and the PDA video. Of note, the only difference between the PRIMETIME main study patient information sheets provided in the standard and enhanced information groups was that in the *enhanced* group it was specified that patients will be provided with a video. The diagrams used in the standard information and enhanced information groups were the same.

6.3.3 Standard information group

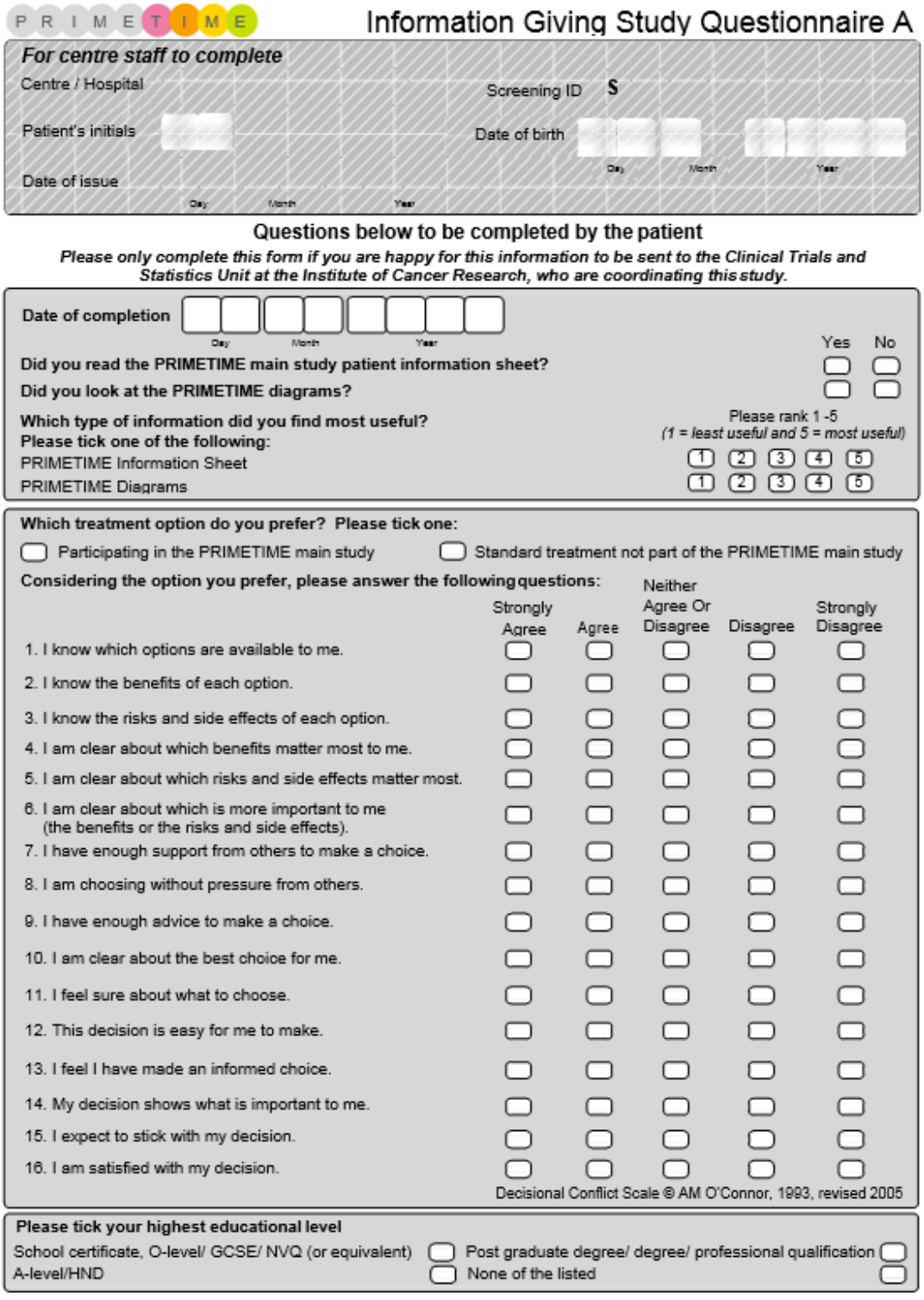
When clusters were in the *standard* information group it was requested that participants had a discussion with a healthcare professional and were provided

with the standard information. After the patient had made their decision regarding entry into PRIMETIME, they were to be presented with the standard questionnaire (figure 6.1) which included baseline demographics, questions regarding the information provided (standard information) and the decisional conflict questionnaire [84]. Return of the Questionnaire indicated consent to the PRIMETIME IGS.

6.3.4 Enhanced information group

When clusters were in the *enhanced* information group it was requested that participants had a discussion with a healthcare professional and were provided with the enhanced information. It was advised that the PDA video should be viewed after the consultation with the healthcare professional and after the patient had read the PRIMETIME main study patient information sheet but *prior* to the patient making their decision regarding entry into the PRIMETIME main study. After the patient had made their decision regarding entry into PRIMETIME, they were presented with the enhanced questionnaire (figure 6.2) which included baseline demographics, questions regarding the information provided (standard information and PDA video) and the decisional conflict questionnaire [84]. Return of the Questionnaire indicated consent to the IGS.

Figure 6.1: Questionnaire given to patients in the standard group



PRIMETIME

Information Giving Study Questionnaire A

For centre staff to complete

Centre / Hospital

Screening ID S

Patient's initials

Date of birth

Day Month Year

Date of issue

Day Month Year

Questions below to be completed by the patient

Please only complete this form if you are happy for this information to be sent to the Clinical Trials and Statistics Unit at the Institute of Cancer Research, who are coordinating this study.

Date of completion

Day Month Year

	Yes	No
Did you read the PRIMETIME main study patient information sheet?	☐	☐
Did you look at the PRIMETIME diagrams?	☐	☐

Which type of information did you find most useful?
Please tick one of the following:

Please rank 1 -5
(1 = least useful and 5 = most useful)

PRIMETIME Information Sheet	1 2 3 4 5
PRIMETIME Diagrams	1 2 3 4 5

Which treatment option do you prefer? Please tick one:

☐ Participating in the PRIMETIME main study ☐ Standard treatment not part of the PRIMETIME main study

Considering the option you prefer, please answer the following questions:

	Strongly Agree	Agree	Neither Agree Or Disagree	Disagree	Strongly Disagree
1. I know which options are available to me.	☐	☐	☐	☐	☐
2. I know the benefits of each option.	☐	☐	☐	☐	☐
3. I know the risks and side effects of each option.	☐	☐	☐	☐	☐
4. I am clear about which benefits matter most to me.	☐	☐	☐	☐	☐
5. I am clear about which risks and side effects matter most.	☐	☐	☐	☐	☐
6. I am clear about which is more important to me (the benefits or the risks and side effects).	☐	☐	☐	☐	☐
7. I have enough support from others to make a choice.	☐	☐	☐	☐	☐
8. I am choosing without pressure from others.	☐	☐	☐	☐	☐
9. I have enough advice to make a choice.	☐	☐	☐	☐	☐
10. I am clear about the best choice for me.	☐	☐	☐	☐	☐
11. I feel sure about what to choose.	☐	☐	☐	☐	☐
12. This decision is easy for me to make.	☐	☐	☐	☐	☐
13. I feel I have made an informed choice.	☐	☐	☐	☐	☐
14. My decision shows what is important to me.	☐	☐	☐	☐	☐
15. I expect to stick with my decision.	☐	☐	☐	☐	☐
16. I am satisfied with my decision.	☐	☐	☐	☐	☐

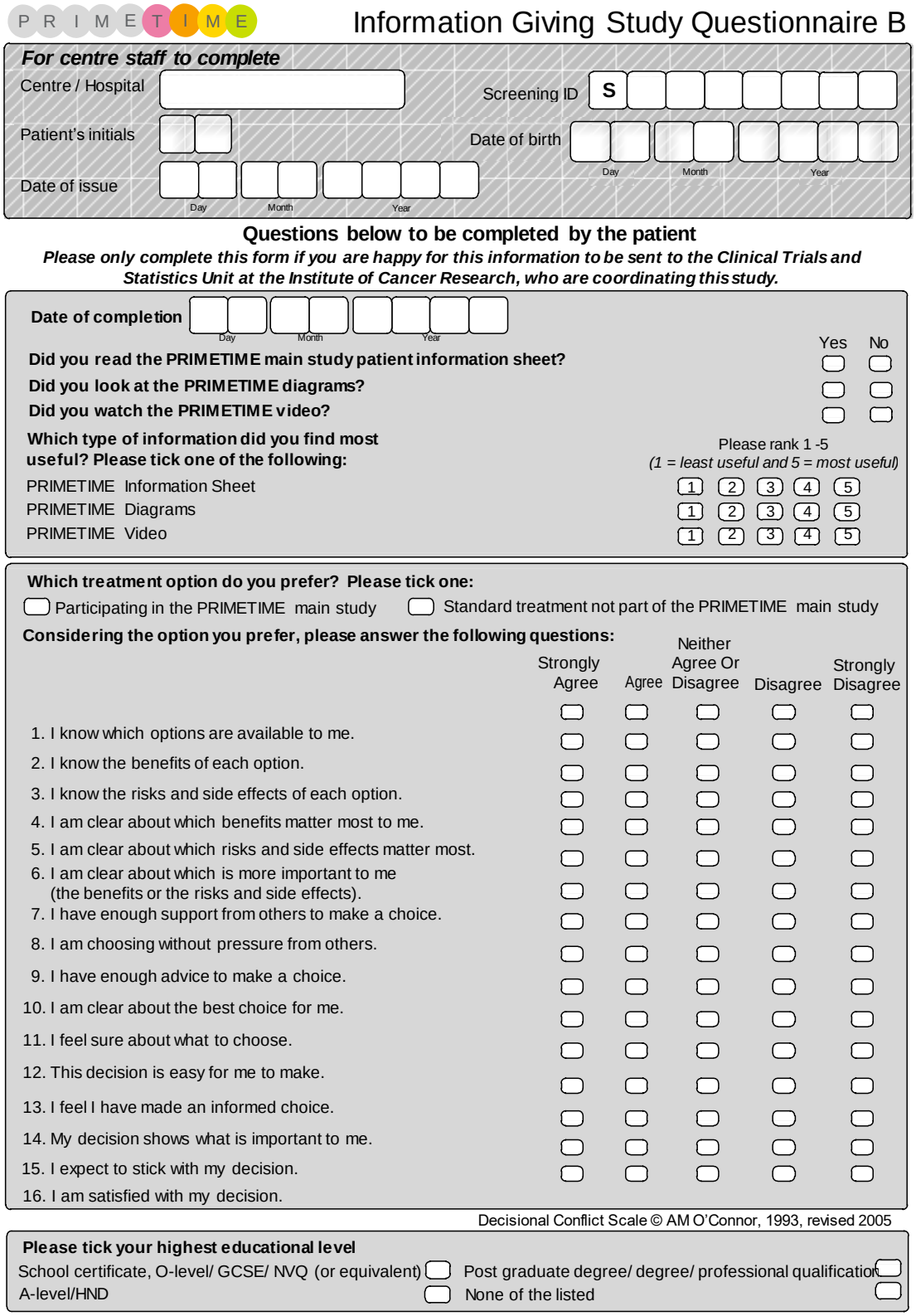
Decisional Conflict Scale © AM O'Connor, 1993, revised 2005

Please tick your highest educational level

School certificate, O-level/ GCSE/ NVQ (or equivalent) ☐ Post graduate degree/ degree/ professional qualification ☐

A-level/HND ☐ None of the listed ☐

Figure 6.2: Questionnaire given to patients in the enhanced information group



PRIMETIME

Information Giving Study Questionnaire B

For centre staff to complete

Centre / Hospital

Screening ID S

Patient's initials

Date of birth Day Month Year

Date of issue Day Month Year

Questions below to be completed by the patient

Please only complete this form if you are happy for this information to be sent to the Clinical Trials and Statistics Unit at the Institute of Cancer Research, who are coordinating this study.

Date of completion Day Month Year

	Yes	No
Did you read the PRIMETIME main study patient information sheet?	☐	☐
Did you look at the PRIMETIME diagrams?	☐	☐
Did you watch the PRIMETIME video?	☐	☐

Which type of information did you find most useful? Please tick one of the following:

Please rank 1 -5
(1 = least useful and 5 = most useful)

PRIMETIME Information Sheet	1	2	3	4	5
PRIMETIME Diagrams	1	2	3	4	5
PRIMETIME Video	1	2	3	4	5

Which treatment option do you prefer? Please tick one:

☐ Participating in the PRIMETIME main study ☐ Standard treatment not part of the PRIMETIME main study

Considering the option you prefer, please answer the following questions:

	Strongly Agree	Agree	Neither Agree Or Disagree	Disagree	Strongly Disagree
1. I know which options are available to me.	☐	☐	☐	☐	☐
2. I know the benefits of each option.	☐	☐	☐	☐	☐
3. I know the risks and side effects of each option.	☐	☐	☐	☐	☐
4. I am clear about which benefits matter most to me.	☐	☐	☐	☐	☐
5. I am clear about which risks and side effects matter most.	☐	☐	☐	☐	☐
6. I am clear about which is more important to me (the benefits or the risks and side effects).	☐	☐	☐	☐	☐
7. I have enough support from others to make a choice.	☐	☐	☐	☐	☐
8. I am choosing without pressure from others.	☐	☐	☐	☐	☐
9. I have enough advice to make a choice.	☐	☐	☐	☐	☐
10. I am clear about the best choice for me.	☐	☐	☐	☐	☐
11. I feel sure about what to choose.	☐	☐	☐	☐	☐
12. This decision is easy for me to make.	☐	☐	☐	☐	☐
13. I feel I have made an informed choice.	☐	☐	☐	☐	☐
14. My decision shows what is important to me.	☐	☐	☐	☐	☐
15. I expect to stick with my decision.	☐	☐	☐	☐	☐
16. I am satisfied with my decision.	☐	☐	☐	☐	☐

Decisional Conflict Scale © AM O'Connor, 1993, revised 2005

Please tick your highest educational level

School certificate, O-level/ GCSE/ NVQ (or equivalent) ☐ Post graduate degree/ degree/ professional qualification ☐

A-level/HND ☐ None of the listed ☐

6.3.5 Cross-over from the standard to enhanced information group

After the first patient from each centre returned a questionnaire, the site was instructed to inform the PRIMETIME team via email. This would trigger the future cross-over date to be determined. Allocation to cross-over over at either 2, 4 or 6 months post first patient entered at each centre was done using minimisation. Minimisation was performed manually using a single balancing factor which was prior recruitment to the IMPORT HIGH [40] and FAST FORWARD [33] breast radiotherapy trials, in order to give the best estimate of likely recruitment in PRIMETIME. Centres were allocated to be expected high or low recruiters based on entry into IMPORT HIGH and FAST FORWARD, and if the centre did not participate in either trial, they were assumed to be in the low recruiter group. By using recruitment status of previous breast radiotherapy trials as a balancing factor, this ensured that all the top recruiters did not receive the same information at the same time. Each centre was informed of their cross-over date via email. A reminder email one-week prior to the cross-over date was also sent which contained a web-link to the PDA video which was not password protected. DVDs were sent to the centres one-week prior to the cross-over date.

6.3.6 Study endpoints

The primary endpoint was decisional conflict as measured on the decisional conflict scale [84]. Secondary endpoints were 1) acceptance of entry into the PRIMETIME main study and 2) acceptance of the recommended treatment within the PRIMETIME study.

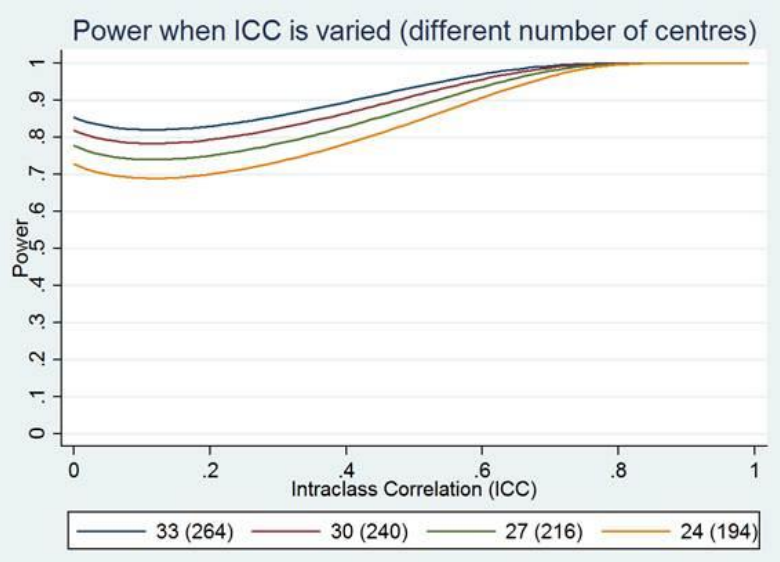
6.3.7 Statistical considerations

6.3.8 Sample size calculations

The target sample size for the PRIMETIME IGS was originally 264 patients. This sample size was based on three steps in the cluster stepped-wedge trial design (at 2, 4 and 6 months) of 33 clusters (11 per step), with 2 patients per cluster per 2 month period. There is limited literature on what is a clinically significant reduction in decisional conflict. Two small single centre studies conducted in similar populations to patients in PRIMETIME found effect sizes around 0.40,

with standard deviations for the total Decisional Conflict Scale score ranging from 11-25 [91,92]. One of the studies reported a mean decisional conflict score of [13.2 (SD=14.5)] [91]. It has been suggested than an effect size of 0.2 indicates a small effect, 0.5 a medium effect and 0.8 a large effect. As this is a cluster randomised trial, the sample size estimation needs to allow for possible clustering effects. However, there is no data available on likely values of the intraclass correlation (ICC) for the Decisional Conflict Scale, and so estimates have been calculated across the range of ICC values from 0 to 1. Assuming an alpha of 0.05, 264 patients from 33 clusters would have at least 80% power for all values of the ICC to detect a 10-point difference in total score for the Decisional Conflict Scale (effect size=0.55, assuming standard deviation=18). If this target is not achievable, then 240 patients from 30 centres would provide at least 80% power to detect an effect size of 0.55 across most of the range of ICC values (figure 6.3).

Figure 6.3: Summary of the power obtained when the intraclass correlation is varied from 0-1



An interim analysis was initially conducted from a data snapshot taken on 31/1/2019. The purpose of this interim analysis was to provide data for inclusion in this confidential PhD thesis. At the time of the first interim analysis on 31/1/2019, 215 evaluable questionnaires had been returned from 21 clusters. Based on these results, with a better than expected return rate of questionnaires but from fewer clusters, repeat sample size calculations were performed using the current estimate of the ICC (obtained in the interim analysis) for the

decisional conflict score. It was found that 288 patients from 24 clusters (8 clusters per step) would give 84% power. The target sample size for the IGS was therefore revised to 300 patients (to account for ineligible patients/ missing data) from 24 clusters.

By May 2019, 318 evaluable questionnaires had been returned from 24 clusters and therefore a second interim analysis for inclusion of updated data within this thesis was conducted on 02/05/2019. However, for the final analysis all clusters are required to have crossed over to the intervention group, as per the stepped wedge trial design. The final analysis will therefore be conducted in October 2019 after the last cluster has crossed over to the *enhanced* information group.

6.3.9 Questionnaire tools

Decisional Conflict Scale: The "statement format" of the decisional conflict scale was used which consists of 16 items and 5 response categories. The response categories are 0='strongly agree'; 1='agree', 2='neither agree nor disagree'; 3='disagree'; 4='strongly disagree'. For calculation of the total score, the 16 items [items 1-16 inclusive] are a) summed; b) divided by 16; c) multiplied by 25. The range of scores is from 0 [no decisional conflict] to 100 [extremely high decisional conflict].

Decisional Conflict Subscales: The decisional conflict subscales include the uncertainty, informed, values clarity, support and effective decision subscores.

Uncertainty subscore: 3 items [10,11,12] are a) summed; b) divided by 3; and multiplied by 25. Scores range from 0 [feels extremely certain about best choice] to 100 [feels extremely uncertain about best choice].

Informed subscore: 3 items [1, 2, 3] are a) summed; b) divided by 3; and multiplied by 25. Scores range from 0 [feels extremely informed] to 100 [feels extremely uninformed].

Values clarity subscore: 3 items [4, 5, 6] are a) summed; b) divided by 3; and multiplied by 25. Scores range from 0 [feels extremely unclear about personal

values for benefits and risks/side effects] to 100 [feels extremely clear about personal values].

Support subscore: 3 items [7, 8, 9] are a) summed; b) divided by 3; and multiplied by 25. Scores range from 0 [feels extremely supported in decision making] to 100 [feels extremely unsupported in decision making].

Effective decision subscore: 4 items [13, 14, 15, 16] are a) summed; b) divided by 4; and multiplied by 25. Scores range from 0 [good decision] to 100 [bad decision].

Information format provided: With regard to the format of information provided, patients were asked to state whether they had used the PRIMETIME main study patient information sheet, diagrams and PDA video (if applicable) with a yes/no response. Patients were also asked to rank on a scale of 1-5 with 1=least useful and 5=most useful, how useful the PRIMETIME main study patient information sheet, diagrams and PDA video were. Of note this was a non-validated scale therefore interpretations of this data are limited. Finally, patients were asked to indicate their highest level of education from the following options: Postgraduate degree/degree, A-level/HND, School certificate/O-level or not listed.

6.3.10 Statistical analysis plan

Analysis Populations: Analyses were conducted on an intention to treat (ITT) basis i.e. clusters were analysed according to the centre allocation PDA video start time regardless of when the cluster started using the PDA video and whether the patient watched the PDA video. Patients were analysed as being in the *standard* information group if the questionnaire completion date was prior to the cross-over date for that centre and as being in the *enhanced* information group if the questionnaire completion date was on or after the cross-over date for that centre. Only patients who returned questionnaires (i.e. consented to the PRIMETIME IGS) were analysed, regardless of whether they consented to participate in the PRIMETIME main study.

Review of distribution of decisional conflict data: The distribution of the decisional conflict scores was reviewed in the *standard* and *enhanced*

information groups (plotted using histograms) and the mean and standard deviations as well as the median and interquartile ranges were reported. If the data were skewed, attempts could be made to normalise the data by transforming the scores. If the transformed data remained skewed, then the decisional conflict scores could be categorised into a binary variable for analysis. Evidence suggests that decisional conflict scores of ≥25 are consistent with clinically significant decisional conflict and associated with decisional delay and decisional regret [139] [140]. The proportions of patients with clinically significant decisional conflict in the standard and enhanced information groups were reported.

Primary endpoint analysis: An estimate of the difference in mean decisional conflict scores pre- and post-implementation of the PDA video was obtained and reported as a coefficient with 95% confidence interval, along with a test for significance (using the z-statistic), and the effect size reported. The coefficient represented a cluster level fixed effect for assigned group, obtained from a multilevel mixed effects linear regression model. To adjust for calendar time and clustering a random effect for cluster and a fixed effect for each step were included in the multilevel model. Robust standard errors were used to adjust for the clustering effect. It should be noted that the output from the mixed effects multi-level regression model are parametric summary statistics, regardless of whether the data are normally distributed. Simple non-parametric tests to compare decisional conflict scores between the groups would not be appropriate in this instance, as this would not take into account the effects of clustering and the time effect in the stepped wedge trial design.

The decisional conflict scores were also dichotomised into two groups <25 (no clinically significant decisional conflict) and ≥25 (clinically significant decisional conflict) and analysed (as a binary variable) to see if there was a difference in clinically significant decisional conflict between the *standard* and *enhanced* information group. This was analysed using a multilevel mixed effects logistic regression model with a cluster level fixed effect for assigned group at that step (reported as odds ratios with 95% confidence intervals), with a random effect for cluster and a fixed effect for each step [137,141].

Sensitivity analysis: In order to investigate the possibility that a reduction in decisional conflict could be attributed simply to a 'learning curve' where centres improve in their trial delivery the longer a trial is open resulting in reduced decisional conflict (rather than the reduction being due only to the PDA video implementation), a sensitivity analysis for the primary endpoint was conducted excluding patients who had returned questionnaires within the first 2 months of their centre having begun recruiting to the IGS.

Also, the mean (SD) and median (IQR) decisional conflict values were identified in **all patients** returning questionnaires in monthly intervals from 0-6 months in the ***standard* information** group from when their centre opened to the IGS, to identify if there was any trend in the scores i.e. a reduction in decisional conflict scores over time, which may suggest a learning curve.

Secondary endpoint analyses: The proportion of patients accepting main study entry out of all patients who returned questionnaires was calculated separately for the *standard* and *enhanced* information groups. In patients who returned questionnaires, the proportion of patients in the main study who accepted recommended treatment was calculated separately for the *standard* and *enhanced* information groups. Both secondary endpoints were analysed using a multilevel mixed effects logistic regression model with a cluster level fixed effect for assigned group at that step (reported as odds ratios with 95% confidence intervals) with a random effect for cluster and a fixed effect for each step [137,141].

Exploratory analyses: Exploratory analyses were conducted to determine whether baseline factors (age and education level) were associated with decisional conflict as assessed in the primary endpoint analysis. Age was tested as a continuous variable and also categorised into 5-year intervals: aged 60-64, 65-69, 70-74 and ≥75 years. Education was tested as a categorical variable and also fitted as a continuous variable in the model to test for linear trend across the categories. Age and education were tested individually within the multilevel mixed effects model as per the primary endpoint analysis; if both factors were significant, they would then be tested together in multivariable analysis. No other clinical variables were tested as these data cannot be revealed until the primary

endpoint analysis from the main study has been completed (not expected to report before 2023).

Exploratory analyses regarding the decisional conflict subscales were conducted using the same methods as for the primary endpoint of overall decisional conflict score, and whether there was any difference in the decisional conflict subscales in the standard versus enhanced information groups.

Descriptive statistics were used to report the patient's view on the usefulness of each of the information formats according to whether the standard or enhanced questionnaire was returned. These data were not formally compared between the groups.

The ICC value for the overall DCS score was estimated from the primary endpoint analyses.

Missing data: There is no published guidance available for missing data in the decisional conflict scale. If there were 8 or more items in the decisional conflict scale missing per questionnaire, the patient was excluded from the analysis. If there were fewer than 8 questionnaire items missing per questionnaire, the missing items were imputed using a mean value from the completed responses. Similarly, for the decisional conflict subscales, if there were fewer than 2 items missing for the uncertainty, informed, values clarity and support subscales, the missing items were imputed using a mean value from the completed responses. For the effective decision subscale, if there were fewer than 3 items missing for the effective decision subscale, the missing items were imputed using a mean value from the completed responses. This is the approach used in previous studies of patient-reported outcomes, for example the EORTC questionnaires where at least half of the items need to be completed in order to calculate subscale scores [142]. If there were any other missing data on the questionnaire, these data were recorded as missing. Attempts were not made to retrieve this information from centres.

Timing of analysis: An interim analysis of the PRIMETIME IGS was planned for the first quarter of 2019 for data to be included within this confidential PhD thesis. The second interim analysis was carried out using STATA version 14 based on a data snapshot taken on 2nd May 2019 and data included in this thesis. The final analysis will take place once the revised target sample size has been accrued and all clusters have crossed over from the standard to the enhanced information groups (expected October 2019).

6.4 Results

318 evaluable questionnaires were returned from 463 eligible patients (69% return rate). Questionnaires were returned from 24 clusters (figure 6.4). All patients in the *standard* information group returned the standard questionnaire. However, a proportion of patients in the *enhanced* information group [25/160 (16%)] returned the *standard* questionnaire (figure 6.5), suggesting these patients were given the incorrect information format for that specified time. With regard to baseline characteristics, the median age was similar between the *standard* and *enhanced* information groups [71 versus 69 respectively]. Levels of education were also similar between the two groups (table 6.1). The majority of patients in the *standard* and *enhanced* information groups reported they had read the PRIMETIME main study patient information sheet and the diagrams. However, almost half of patients in the *enhanced* information group reportedly did not watch the PDA video (table 6.2).

Figure 6.4: Summary of questionnaires returned per cluster in the standard and enhanced information groups in the PRIMETIME IGS

Cluster	0 months	2 months	4 months	6months	8 months
Cluster 1	2	2	0	4	3
Cluster 2	13	12	10	10	2
Cluster 3	3	3	7	2	0
Cluster 4	6	7	6	8	2
Cluster 5	8	5	3	6	3
Cluster 6	2	0	1	0	0
Cluster 7	4	6	4	1	0
Cluster 8	3	0	0	2	0
Cluster 9	11	2	0	0	0
Cluster 10	3	1	1	0	0
Cluster 11	1	0	0	0	0
Cluster 12	5	6	6	0	4
Cluster 13	3	9	2	6	13
Cluster 14	2	0	1	1	3
Cluster 15	14	10	6	6	5
Cluster 16	2	3	0	0	3
Cluster 17	2	0	0	0	0
Cluster 18	8	1	1	1	0
Cluster 19	2	0	2	2	0
Cluster 20	2	1	1	0	0
Cluster 21	2	1	0	0	0
Cluster 22	7	1	0	0	0
Cluster 23	3	1	0	0	0
Cluster 24	1	0	0	0	0

Key: White box: Cluster receiving standard information. Pink Box: Cluster receiving enhanced information. Number in box represents the number of patients returning questionnaires in the standard and enhanced information groups per cluster

Figure 6.5: Consort diagram

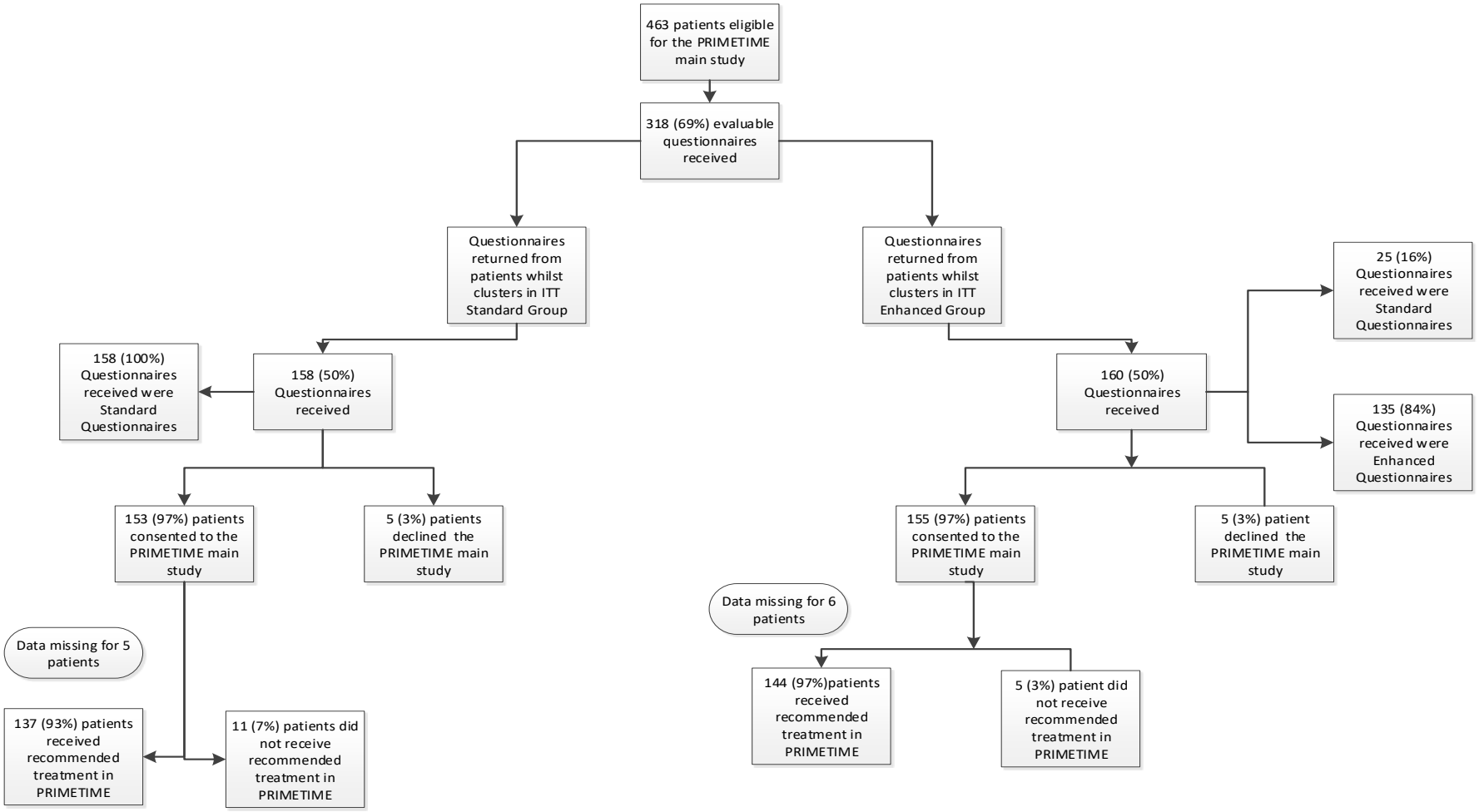


Table 6.1: Summary of baseline characteristics in patients in the standard and enhanced information groups

Baseline Characteristics	Standard Information Group n=158 (%)*	Enhanced Information Group n=160 (%)*
Age (median and IQR)	71 (67-74)	69 (65-72)
Age categories:		
60-64	28 (18)	36 (23)
65-69	45 (28)	55 (34)
70-74	52 (33)	49 (31)
≥75	33 (21)	20 (13)
Education Level*		
PG degree/degree	38 (25)	37 (25)
A-level/HND	26 (17)	23 (16)
School cert/O-level	50 (32)	53 (36)
No formal education	40 (26)	35 (24)

**Data regarding education level missing for 4 patients in standard information group and 12 patients in the enhanced information group. Percentages calculated using all available data.*

Table 6.2: Summary of information use in the standard and enhanced information groups

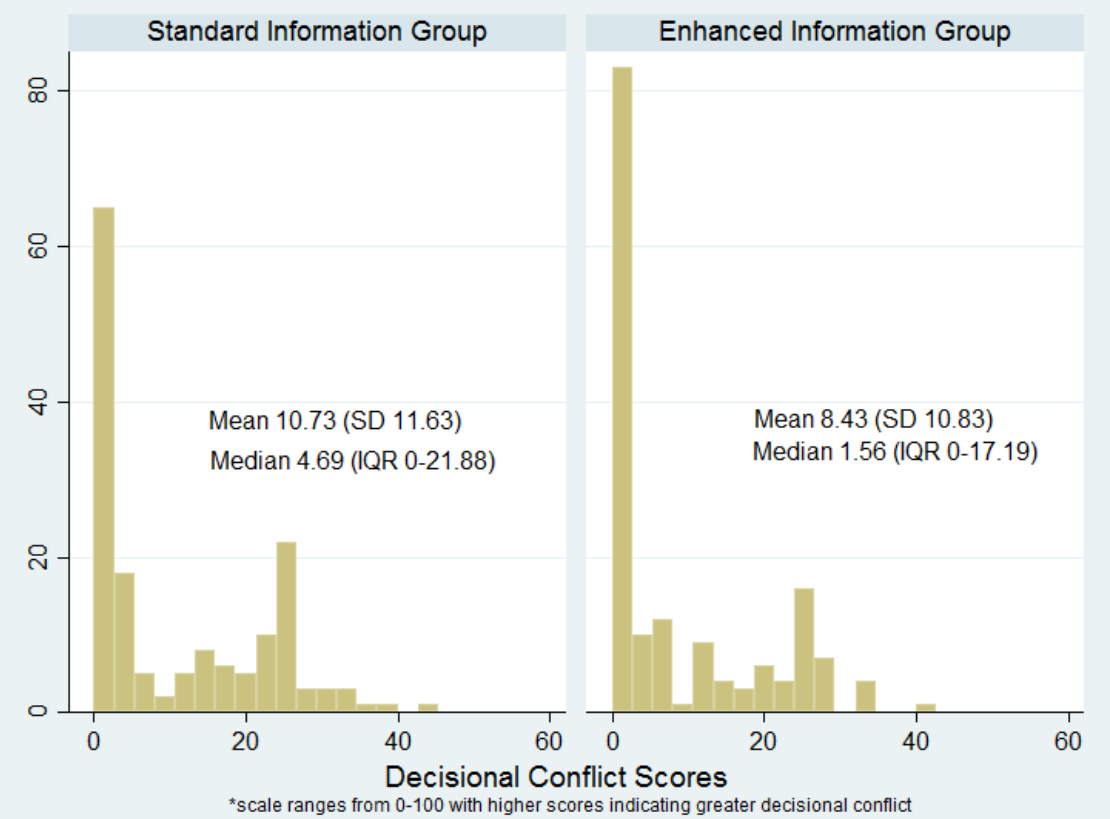
	Standard Information Group n=158 patients (%)	Enhanced Information Group n=160 patients (%)
Proportion of patients reading PIS*	N=152 patients' data available	N=156' patients data available
Yes	152 (100)	156 (100)
No	0	0
Proportion of patients looking at diagrams	N=149 patients	N=153 patients
Yes	117 (79)	132 (86)
No	32 (21)	21 (14)
Proportion of patients watching video	N/A	N=130' patients data available
Yes		61 (47)
No		69 (53)

**PIS=patient information sheet*

6.4.1 Primary endpoint

The distribution of the decisional conflict scales were skewed to the right in both the *standard* and *enhanced* information groups (figure 6.6). An attempt was made to transform the raw decisional conflict scales, however the data remained highly skewed. The mean decisional conflict score in the *standard* information group [158 patients] was 10.73 (SD=11.63) and 8.43 (10.83) in the *enhanced* information group [160 patients]. The median value (and interquartile range) was 4.69 (0-21.88) and 1.56 (0-17.19) in the *standard* and *enhanced* information groups respectively.

Figure 6.6: Histograms of decisional conflict scores in the standard and enhanced information groups



As described in the methods, the mixed effects multi-level regression model uses parametric statistics regardless of whether the data are normally distributed. There was a reduction in decisional conflict in the *enhanced* group compared with the *standard* group, however the effect size was small and

unlikely to be clinically significant [estimated difference in means = -2.50 (-4.73- -0.28), p=0.03, effect size=0.11].

Given the skewed nature of the decisional conflict scores, the data was categorised into a binary variable and analysed using the pre-specified cut-off of ≥25 ('clinically significant' decisional conflict). It was found 34/158 (22%) and 28/160 (17%) had clinically significant decisional conflict in the *standard* and *enhanced* information groups respectively. The odds of having 'clinically significant' decisional conflict in the enhanced information group were reduced compared with the standard information group, [OR 0.6 (0.35-1.03), p=0.06].

6.4.2 Sensitivity analysis

In the sensitivity analysis, when patients who returned questionnaires within the first 2 months of the IGS opening at their centre were excluded, it was found that the mean score in the *standard* information group [83 patients] was 10.9 (12.0) and 8.4 (10.9) in the *enhanced* information group [160 patients]. When the primary endpoint analysis was repeated having excluded these patients, the effect size decreased slightly [-3.01 (-6.72- 0.70) p=0.11, effect size=0.08].

Also, when the mean and median decisional conflict scores in **all patients** in the ***standard* information group**, in one monthly intervals from 0-6 months (from when the IGS had opened at the patient's respective centres) were reviewed, there was no clear trend in reduction in decisional conflict which suggests there was no obvious impact of a learning curve on patients' decisional conflict (figure 6.7). The numbers of patients in the latter time intervals were minimal and so results should be interpreted with caution.

6.4.3 Secondary endpoints

The majority of patients who returned IGS questionnaires had consented to the PRIMETIME study [153/158 (97%) and 155/160 (97%) patients in the *standard* and *enhanced* information groups respectively]. There was no statistically significant difference in patients consenting to the PRIMETIME main study

according to whether they were in the *standard* or *enhanced* information groups [OR=1.67(0.49- 5.68), p=0.42].

In those patients who had returned IGS questionnaires and consented to the main study described above, in the *standard* information group; 137 (93%) patients received their allocated treatment and 11 (7%) patients did not. In the *enhanced* information group, 144 (97%) patients received their allocated treatment and 5 (3%) did not (figure 6.5). There was no statistically significant difference in patients accepting the recommended treatment in PRIMETIME whether they were in the *standard* or *enhanced* information groups [OR=1.35 (0.46-3.95), p=0.58].

6.4.4 Exploratory endpoints

With regard to whether age was associated with decisional conflict, the median value of decisional conflict was higher in patients aged 75 or over (table 6.3). However, there was no significant association between age and decisional conflict score when allowing for the effects of the *standard* and *enhanced* information groups and centre allocation. This was the case when age was tested as a continuous variable and in categories (table 6.3). With regard to education level, there was a trend towards higher decisional conflict scores as education levels reduced (table 6.3).

With respect to the decisional conflict subscales, there were small reductions in the subscale scores in the *enhanced* information compared with the *standard* information groups. However, given the minimal reductions, these are unlikely to be of clinical significance (table 6.4).

The estimate of the ICC for the decisional conflict scale i.e. the homogeneity between the decisional conflict results within each cluster was low at 0.07 (values can range from 0-1).

Figure 6.7: Summary of decisional conflict scores in the standard information groups at varying time-points

Cluster	0-1 months	1-2 months	2-3 months	3-4months	4-5months	5-6months
N	82	27	25	14	7	3
Mean (SD)	10.45 (11.00)	10.72 (12.38)	11.88 (12.30)	13.39 (14.45)	7.81 (10.86)	3.13 (4.13)
Median (IQR)	4.84 (0-20.3)	4.69 (0-23.44)	4.69 (0-25)	8.59 (0-25)	3.13 (0-21.88)	1.56 (0-7.81)

Table 6.3: Summary of association of age and education level with decisional conflict

Baseline characteristics (n= patients with available data)	Mean decisional conflict (SD)	Median decisional conflict (IQR)	*Coefficient value (95%confidence interval), p value
Age (years)			
60-64 (n=64)	9.18 (11.98)	2.34 (0-17.19)	-
65-69 (n=100)	8.91 (10.96)	1.56 (0-20.31)	0.22 (-3.96- 4.41), p=0.916
70-74 (n=101)	8.90 (10.55)	3.13 (0-17.19)	-0.09 (-4.76-4.57), p=0.969
≥75 (n=53)	12.55 (12.55)	14.06 (0-25)	3.81 (-1.55-9.17), p=0.163
Age tested as a continuous variable	-	-	0.18 (-0.05-0.42), p=0.13
Education Level			
PG degree/degree (n=75)	6.3 (9.46)	1.56 (0-10.94)	-
A-level/HND (49)	9.99 (10.72)	4.69 (0-17.19)	3.36 (-0.72-7.44), p=0.107
School cert/O-level (103)	9.40 (11.48)	3.13 (0-20.31)	3.45 (-0.35-7.25), p=0.075
Not listed (75)	11.23 (12.06)	6.25 (0-25)	5.24 (0.29-10.19), p=0.038
Education tested as a continuous variable			1.62 (-0.05-3.29), p=0.057

**Coefficient value represents the difference between the means of the decisional conflict scales in the standard and enhanced information groups*

Table 6.4: Summary of decisional conflict subscales data

Decisional Conflict Subscales	**Mean score in Standard Information Group (standard deviation)**	**Mean score in Enhanced Information Group (standard deviation)**	***Coefficient value (95% confidence interval)**	**P value**
Uncertainty	13.45 (16.01)	11.30 (15.34)	-1.83 (-5.64-1.98)	0.348
Informed	9.72 (11.72)	7.34 (10.59)	-3.12 (-5.29- -0.93)	0.005
Values Clarity	11.67 (13.99)	9.06 (13.00)	-2.68 (-5.45-0.09)	0.058
Support	8.74 (11.61)	6.51 (10.39)	-2.51 (-4.49- -0.52)	0.013
Effective	10.22 (12.84)	8.04 (11.68)	-2.39 (-4.73- -0.05)	0.045

**Coefficient value represents the difference between the means of the decisional conflict scales in the standard and enhanced information groups*

Usefulness of information provided in the standard and enhanced information groups

With regard to the information formats provided in the two groups, the vast majority of patients scored the patient information sheet as 5 (most useful) on a scale of 1-5 [88/131 (67%) and 112/149 (75%) in the *standard* and *enhanced* information groups respectively] (figure 6.8). Similarly, in those patients who confirmed they had used the diagrams, most scored them as most useful in both groups [48/89 (54%) in the *standard* and 73/120 (61%) in the *enhanced* group] (figure 6.9). Finally, with regard to those patients who watched the video, most scored it as between 4-5 [46/54 85%] (figure 6.10).

Figure 6.8: Summary of 'how useful' patients found the PIS in the standard and enhanced information groups

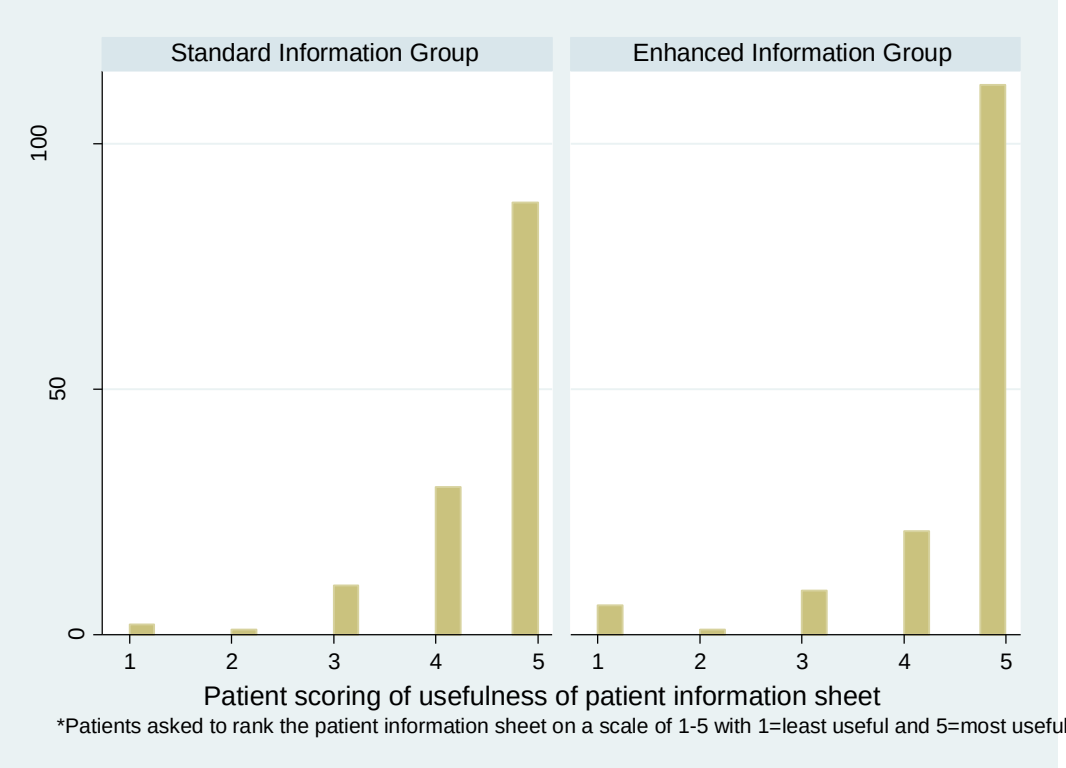


Figure 6.9: Summary of 'how useful' patients found the diagrams in the standard and enhanced information groups

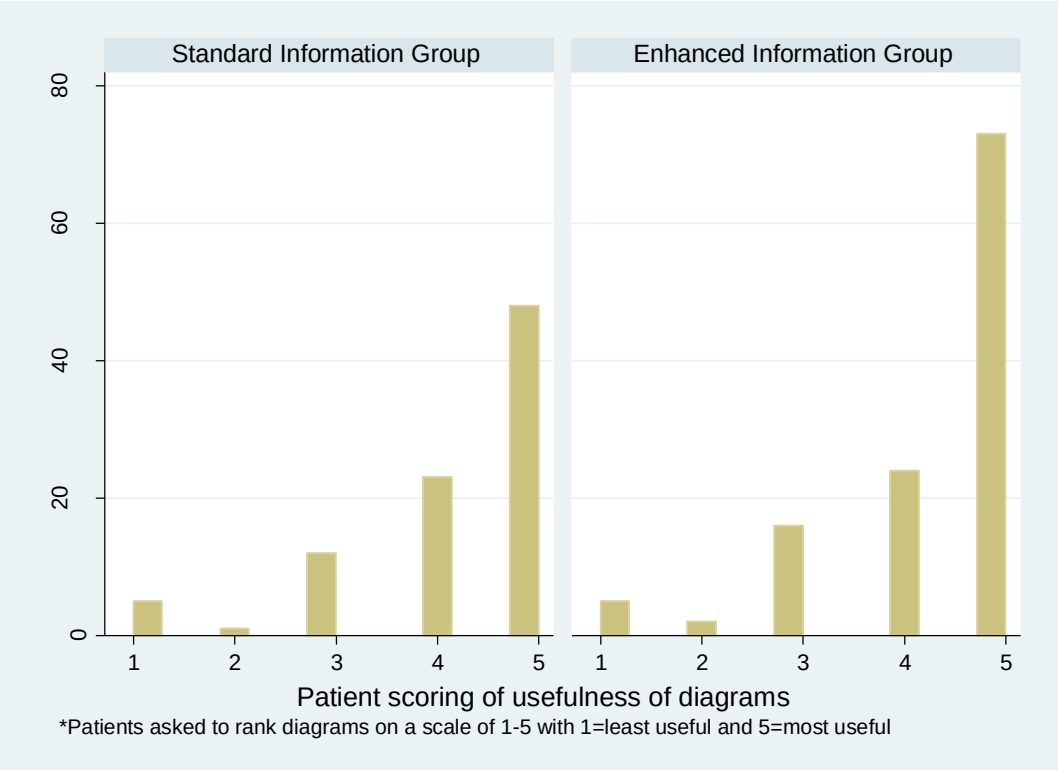
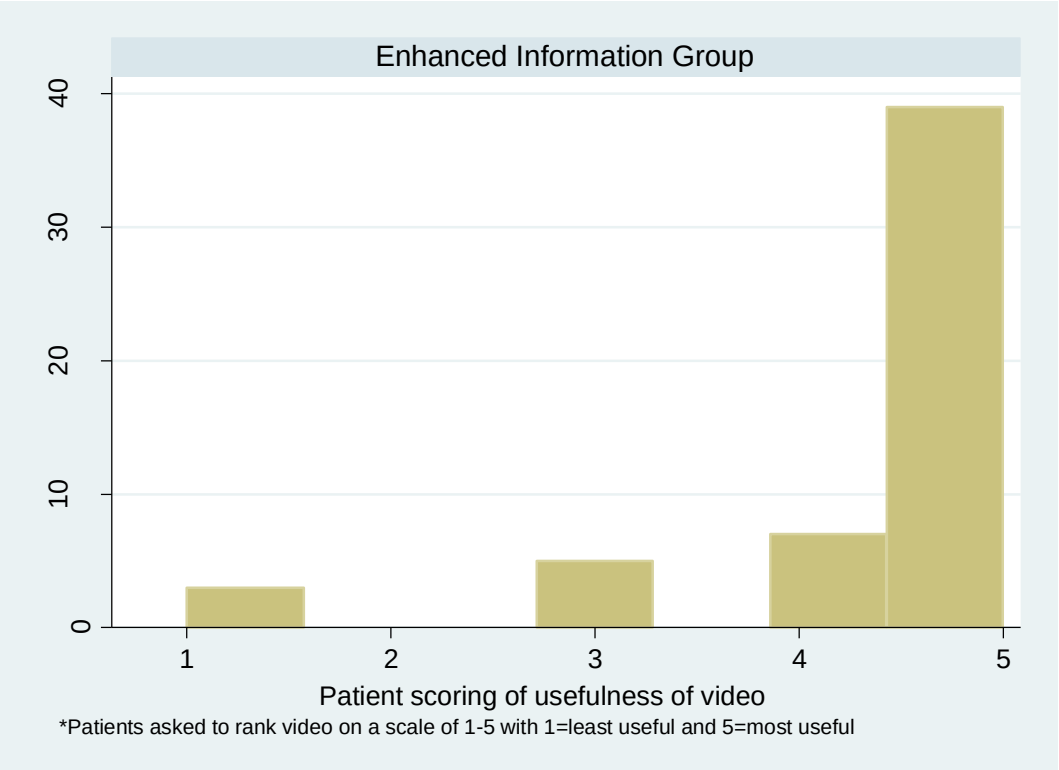


Figure 6.10: Summary of 'how useful' patients found the video in the enhanced information group



Compliance with watching the PDA video

There was a wide variation in compliance with the PDA video according to each cluster in the *enhanced* information group which ranged from 0-100% [median=48%] (table 6.5). There was no increase in the odds of a patient complying with watching the video for every month that the centre was in the *enhanced* information group [OR 1.03 (0.86-1.22), p=0.78].

Table 6.5: Summary of video compliance in centres allocated to the enhanced information group

Cluster	Watched Video N=number (%)	Did Not Watch Video	Data missing	Total
1	4 (57)	3	0	7
2	1 (5)	17	4	22
3	5 (42)	5	2	12
4	12 (52)	8	3	23
5	6 (46)	4	3	13
6	1 (100)	0	0	1
7	0 (0)	0	1	1
8	1 (50)	1	0	2
10	2 (100)	0	0	2
12	4 (100)	0	0	4
13	8 (27)	15	7	30
14	0 (0)	5	0	5
15	10 (37)	9	8	27
16	3 (100)	0	0	3
18	2 (100)	0	0	2
19	2 (50)	2	0	4
20	0 (0)	0	1	1
23	0 (0)	0	1	1

Baseline characteristics did not appear to affect compliance with watching the PDA video. The median age of patients having watched the video was 70 (IQR=66-72) and was 70 (66-73) in those who did not watch the video. Compliance with watching the video did not appear to vary with education level (table 6.6).

Table 6.6: Compliance of video stratified per patient education level in the enhanced information group in those with available data

Baseline Characteristics	**Video compliance in patients in the enhanced information group n=130 (%)***
Education Level*	
PG degree/degree	16/28 (57%)
A-level/HND	6/19 (32%)
School cert/O-level	19/41 (46%)
Not listed	18/33 (55%)

**In those with available data*

6.5 Discussion

A SWAT using a cluster stepped-wedge trial design was conducted within the PRIMETIME avoidance of radiotherapy study in order to test whether the use of a PDA video in addition to standard written information could reduce decisional conflict in patients. It was found that the absolute levels of decisional conflict scores in both the standard and enhanced information groups were low on average. Although there was a small reduction in decisional conflict in the enhanced information group, this is unlikely to be clinically significant as the effect size was small. Furthermore, the effect size decreased slightly further when the sensitivity analysis was conducted to investigate whether a learning curve was present. Almost half the patents in the enhanced information group did not reportedly watch the video. There was no statistically significant difference in the acceptance of PRIMETIME entry or recommended treatment in patients entering PRIMETIME between the two groups.

The SWAT methodology was implemented in multiple centres across the UK participating within PRIMETIME using a cluster-stepped wedge trial design. However, the vast majority of patients who participated in the IGS also participated in PRIMETIME main study, with few patients returning

questionnaires who declined the main study. Also, a proportion of patients were given the incorrect questionnaires for the pre-specified information allocation.

6.5.1 Primary endpoint - reduction in decisional conflict

This study demonstrated a reduction in decisional conflict in the enhanced compared with the standard group, however this reduction was not clinically significant. Firstly, the absolute decisional conflict scores were low on average (mean values of 11 and 8 in the standard and enhanced information groups respectively on a scale from 0-100) and the distributions of data were highly skewed with a large proportion of patients scoring 0. Various definitions of 'high' or 'clinically significant' decisional conflict have been used with score cut-offs of ≥25 or ≥37.5 [140]. Decisional conflict scores above these levels have been associated with decisional delay and regret. The PRIMETIME IGS study was conducted in a population of patients with a low clinical risk of recurrence, and it may be that this patient group have low levels of decisional conflict. Given the low average absolute scores, this suggests that it may not be possible to obtain a substantial reduction following an intervention in this study.

Secondly, the estimated difference in means between the enhanced and standard information groups was -2.50 (-4.73- -0.28) and the effect size was only 0.11. The PRIMETIME IGS had in fact been powered to detect a reduction of 10-points (effect size 0.55), although this was not based on published literature as there is no defined clinically significant reduction in decisional conflict. In a study conducted by Wong et al, investigating the use of a decision aid in booklet format for older women aged 70 or over with stage 1 breast cancer considering radiotherapy after lumpectomy [92] the baseline scores were 25.4 (17.8) and there was a reduction in decisional conflict with patients receiving the decision aid with an adjusted mean difference of -7.18 (95%CI=-13.49 to -0.85) [92]. This was in a slightly older, but similar risk-group as patients in our study. It should also be noted this study had small patient numbers (<40) and the same patients were assessed before and after receiving the decision aid. In contrast, the IBIS II trial which investigated the use of a decision aid in a randomised control trial of an aromatase inhibitor in patients at high risk of breast cancer (prevention group)

and patients with DCIS (treatment group) reported similar baseline decisional conflict scores to the PRIMETIME IGS [13.2 (SD=14.5)] [91]. Of note, patients in the IBIS II trial were randomised on an individual patient basis whereas patients in the PRIMETIME IGS were randomised in clusters. In order to better estimate a clinically significant reduction in decisional conflict in the PRIMETIME IGS, one option would have been to conduct a pilot study in a smaller, but similar population of patients. This would inform conduct of a larger study however this would not have been feasible within the limited timeframes of the PhD.

Although a non-clinically significant reduction in decisional conflict in the enhanced compared with the standard information group was found (given the small effect size), the possibility of any reduction being due to a learning curve needed to be investigated. The results of the sensitivity analysis found that when patients returning questionnaires within the first 2 months of the centre opening to the IGS were excluded, the effect size decreased slightly but remained small. Furthermore, there was also no obvious trend over time in decisional conflict scores in the standard information group, which does not support the theory of an impact of a learning curve on patients' decisional conflict, although numbers of questionnaires in the different time periods analysed were small. It is acknowledged that when using an alpha of 0.05, although the results of the primary endpoint analysis were not clinically significant, they were statistically significant and this significance was lost when the sensitivity analysis was performed. However, this was most likely due to the reduction in sample size when the sensitivity analysis was conducted and these findings are not clinically significant.

6.5.2 Secondary endpoints

With respect to the secondary endpoints, receiving enhanced group information did not make any difference to whether patients accepted entry into PRIMETIME or whether they accepted the recommended treatment. An important consideration for these results is that the majority of patients returning questionnaires consented to the PRIMETIME main study (95%), therefore there are limitations to this analysis as very few patients returning questionnaires declined the PRIMETIME main study. Also, the purpose of the PDA video was

to reduce patient uncertainty and not to improve patient recruitment. For example, the general benefits of participating in a clinical trial were not discussed which may be why the PDA video did not affect recruitment or acceptance of recommended treatment. In a study investigating the use of a decision aid (in a decision board format) in women considering radiotherapy post-lumpectomy, the decision aid also did not affect the patients' choice to select breast radiotherapy [143].

6.5.3 Exploratory endpoints

There was a suggestion of a trend towards higher decisional conflict scores in patients with lower educational levels, but the average absolute scores were low and this trend was not thought to be clinically significant. Also, although there were reductions in the decisional conflict subscales in the enhanced compared with the standard information group, the absolute scores were low with minimal reductions and not deemed to be clinically significant. Of note, in the study by Wong et al there was a significant reduction in the values clarity subscale only [92] which may be related to the values clarification exercise within their decision aid.

6.5.4 Lack of compliance with the patient information provided

Approximately half of the patients in the enhanced information group did not reportedly watch the video. Personal correspondence with research staff at sites revealed that several patients commented the written information was sufficiently informative and they did not wish to watch yet another piece of information. Others also commented that having several formats of information was too overwhelming for the patient. In the enhanced information group, the median age of patients watching and not watching the video was the same (70 years). It has been suggested that the information needs of older women may differ from those of younger women. Older women may place higher value on face-to-face or telephone conversations with cancer volunteers or friends treated for breast cancer compared with video-based resources [144]. Of note, education level did not appear to affect video compliance as compliance was 57% and 55% in the highest and lowest educational levels respectively. Also, our results found that video compliance was not affected by an increased length of time the centre

had been in the enhanced information group. It may have been expected that an increased length of time with access to the PDA video would have increased compliance as sites gained experience at going through the information provided in the enhanced information group, and also encouraging patients to watch the PDA video.

It was also apparent that although the patient information diagrams were part of the ethics approved information alongside the PRIMETIME main study information sheets in both the standard and enhanced information groups, a proportion of patients indicated that the diagrams were not used. The diagrams were designed so that healthcare professionals could talk through the diagrams with the patient, suggesting in some circumstances healthcare professionals chose not to use the diagrams. The diagrams would inevitably increase clinic time in busy NHS hospitals. In contrast, the PDA video can be watched in the patient's own time although it is not a stand-alone piece of information.

Whether other information formats are preferable for patients can be considered. In two of the studies described earlier investigating PDA in patients considering radiotherapy following lumpectomy, the format of decision aid included a decision board [143] and a booklet [92]. Decision boards are designed for use within the consultation and introduce information sequentially rather than overwhelm patients with all the information at once [90,143]. One of the disadvantages of decision boards is the requirement to restrict discussions to the consultation time limits [90]. Similarly, booklets would require discussion (in part) with a healthcare professional. As previously mentioned, the values clarification exercise was omitted from our PDA due to concerns regarding limited clinic time. However, it is important that patient concerns and questions are answered regardless of the busy NHS clinic setting to ensure patients are able to engage in the decision-making process. If PDA are to be used they need to be embraced by the clinical community as well as by the patients and established as part of the clinical workflow [145].

Few studies have directly compared different PDA formats, or asked patients to review a range of formats and express a preference. Therefore, the optimal PDA format is unclear. It is possible that a variety of formats are effective or that

certain formats are effective for certain populations [90]. In the future, acknowledging individual patient preferences and needs as well as the practicalities of information delivery within busy NHS clinic settings will need to be considered when developing PDA.

6.5.5 The IGS SWAT population

In our study, questionnaires were returned from 69% of the eligible population which included all patients who were eligible for the PRIMETIME main study. Furthermore, of those patients who returned questionnaires, 97% consented to the PRIMETIME main study. It is possible that patients who declined the PRIMETIME main study also did not want to complete questionnaires. However, verbal correspondence with sites revealed reasons why sites did not offer questionnaires to patients who declined the main study. Firstly, some research staff felt uncomfortable asking patients to complete a questionnaire after they had declined a study as they had concerns whether this was ethically appropriate. As the IGS was an ethics approved sub-protocol focussing on attitudes and behaviours as distinct from the therapeutic interventional study which they had declined, it could be considered unethical **not** to give the patient the opportunity to complete the questionnaire and this was discussed with centre staff. Secondly, in some hospital pathways, once a patient has declined a clinical trial they are not seen again by a member of the research team. In this situation the co-operation of other members of the team responsible for the patient's care is needed. Communicating with research staff and clinical teams at sites regarding the importance of these methodology based studies which may focus on attitudes and behaviours rather than 'hard' cancer outcome specific endpoints is required. Finally, some research staff voiced that patients were overwhelmed early on in the breast cancer diagnosis and they felt giving a questionnaire was not appropriate. However, questionnaire distribution was felt appropriate by our PRIMETIME patient advocates and the PRIMETIME ethics committee. Given the vast majority of patients who returned questionnaires consented to the main study, there is a risk of reporting bias as the population of patients who returned questionnaires in the PRIMETIME IGS may be

characteristically different from those who were eligible but did not return a questionnaire. Nevertheless, it should also be noted that in general there was a high level of acceptance to the PRIMETIME main study in those patients who were eligible for the main study. Following verbal feedback from sites, IB re-discussed the specific patient pathways of information provision and questionnaire distribution at individual sites to attempt to streamline the patient flow through the IGS and improve questionnaire return rates from all eligible patients. A questionnaire was sent to sites in June 2019 to evaluate their experience of running the IGS.

6.5.6 Implementation of the cluster stepped-wedge design

With respect to the implementation of the cluster stepped-wedge trial design, it was found that a proportion of patients returned questionnaires designated for the standard group whilst the cluster was allocated to the enhanced information group. This suggests that some clusters may have had a delayed cross-over despite reminders of the cross-over date. Whilst the stepped-wedge cluster design is pragmatic in that all clusters will eventually receive the intervention i.e. the PDA video, it requires cooperation and commitment from the clusters. Clusters should be ready to cross-over to the intervention as and when the randomisation order dictates [137]. This is of course not always possible in a busy NHS hospital setting where there are limited resources and increasing pressures on research staff. High staff turnover means that information regarding when the cluster was due to cross-over may not have been relayed to new staff. Also, sites may need time to get used to providing an additional format of information as part of a clinical trial. One option would be to have a 'lead in' period of time to initiate the intervention during which time the centres can gain experience with the intervention, but no data would be collected. Although this was considered for the PRIMETIME IGS, it was not feasible given the time constraints of conducting the IGS within limited PhD timeframes.

One of the options to avoid sites having to cross-over would have been to use a parallel design where clusters would be randomised to receive either standard or enhanced information up-front. The IBIS II trial utilised this trial design which enabled uniform information provision at each site throughout the study [91]. The

confounding issue of calendar time would also be avoided with the parallel design. Furthermore, our results found that the ICC was small (0.07 out of a possible range from 0-1) which means that a parallel design would have provided sufficient statistical power. Although, in a stepped-wedge design, each cluster acts as its own control as it contributes observations from both the control and intervention periods. Nevertheless, the ICC value was unknown during the set-up of the IGS and is usually only identifiable in a specific group of patients after a cluster based study has been completed. In any case, it is acknowledged that a parallel design with all sites going on to receive the video after the study had completed accrual would have been reasonable to implement. One of the major challenges in implementing a cluster trial within a recruiting trial regardless of whether a stepped-wedge or parallel design was used, is that the number of clusters and number of patients within each cluster would not be defined which was the case for the PRIMETIME IGS. This is in contrast to most cluster stepped-wedge trials often implemented in public health settings, where the number of clusters and participants per cluster are already known up-front [137].

Regardless of the stepped-wedge design, it should be acknowledged that timing of information delivery for PRIMETIME given the two-stage process of the trial may also have affected whether patients received the designated information (and returned the correct questionnaire) for the cluster at the correct time. In some circumstances, sites issued information for the PRIMETIME main study at the same time as the pre-screening information. If the patient was given the standard group information then they must be given the standard group questionnaire. It may have been possible that patients were given the standard group information but made their decision regarding PRIMETIME participation **AFTER** the cluster had crossed over, so were given the enhanced questionnaire to complete. In this circumstance, there would be a lack of compliance for the cross-over time. Given the various patient pathways at sites and some of which were co-ordinating over several sites, it would not have been reasonable to mandate specific time-points of information delivery as this may have compromised the PRIMETIME main study.

Another consideration with the stepped-wedge design is that if the intervention and outcome measure were influenced by a learning curve, this would

disproportionately affect the standard group compared with the intervention group. It is acknowledged in this study that the evaluation of a PDA could be affected by a learning curve in which sites could have improved in their delivery of information the longer they were recruiting to a study, which may have resulted in a reduction in decisional conflict irrespective of PDA implementation. An option to address the possibility of a learning curve would be to have a lead-in period in the ***standard* group**, during which time data would not be collected. Unfortunately given limited PhD timelines this was not possible for this study. In order to investigate the possibility of a learning curve, a sensitivity analysis was conducted where patients who had returned questionnaires in the first two months of the IGS being open in their centre were excluded. However, there was no evidence of a learning curve found.

6.5.7 Study limitations

The PRIMETIME IGS had a number of limitations. Firstly, there is a risk of reporting bias in the population of patients who returned questionnaires as almost all patients consented to PRIMETIME and we received minimal responses from patients who declined the main study. This was despite explaining the importance of offering questionnaires to all eligible patients to sites. It is likely this is a challenge in any clinical trial which involves capturing patients who decline trials given the current patient pathways in NHS hospitals. Furthermore, the PRIMETIME IGS was restricted to patients who were able to read and understand English independently. Therefore, the patient population in this study may not be representative of the overall patient population of PRIMETIME.

Secondly, not all patients received the correct information according to their designated information at specific time-points. This may have been related to sites not crossing over at the allocated time-point but also due to the timing of information provision to patients within the study. This highlights the challenge to conduct a stepped-wedge design in NHS hospitals with limited resources and high staff turnover.

Thirdly, the majority of patients did not comply with the intervention provided i.e. the PDA video. Reasons for this include the fact that the standard information may have been of sufficient quality to fulfil the information needs of this group of patients. Also, some patients may prefer not to receive information in video format.

6.5.8 Potential implications for practice

The absolute decisional conflict scores in this study in both groups were low on average, and it is reassuring that this population at low clinical risk of local recurrence had low levels of decisional conflict. Also, it should be acknowledged that the *standard* information consisted of the patient information sheets **and** the PDA diagrams. The standard group information itself satisfied many of the IPDAS guidelines with the exception of the values clarification exercise. It is likely therefore that the information provided in the standard group alone may have resulted in reductions in decisional conflict. This suggests that information incorporating PDA should be used as a gold standard for development of trial information. Furthermore, as most patients did not reportedly watch the PDA video, it suggests that the standard written information was sufficiently informative and that an additional PDA may not be required in this population.

A SWAT using a cluster stepped-wedge trial design was implemented in multiple centres participating in a clinical trial. However, as the majority of patients returning questionnaires were those who had also consented to participate in the PRIMETIME main study, there is a risk of reporting bias in these results.

Some of the lessons learnt during the IGS include developing information materials closely with patients to ensure that the information is relevant and presented in an easily interpretable manner i.e. developed by patients for patients. The use of PDA should be incorporated into the information provided. Furthermore, the importance of conducting studies investigating patient-reported psychological outcomes as well as those investigating hard cancer outcomes such as recurrence should be emphasised to sites. Investigating patient-reported psychological outcomes is important for the patient and sufficient

resources are required for this. Finally, using efficient and easily implementable trial designs such as the SWAT concept can enable us to maximise the information collected from individual trials and ultimately benefit the patient.

6.6 Conclusions

The average decisional conflict scores were low in the PRIMETIME IGS, therefore there was less scope for obtaining further reductions, and the difference observed was not clinically significant. The low baseline scores may reflect the PRIMETIME study population who are at low risk of local recurrence. Almost half of patients reportedly did not watch the video; this may be due to the standard written information being sufficient to fulfil the needs of this patient population. The PRIMETIME IGS was conducted using a SWAT concept with a cluster stepped-wedge trial design across multiple UK centres participating in the PRIMETIME study. The majority of patients who consented to the IGS also consented to the PRIMETIME main study.

Chapter 7 Clinical trials from the other side: Lessons learned by a clinician venturing into a clinical trials unit

7.1 Abstract

Open dialogue between clinicians recruiting to clinical trials and trials methodologists and trials managers in clinical trials units (CTU) is important to improve understanding of trials methodology and streamline the clinical trials development and implementation process. Cancer Research UK (CRUK) have developed a scheme, 'The CRUK Clinical Trials Fellowship' which embeds clinical research fellows within a CTU enabling them to gain exposure to clinical trials conduct, design and analysis thereby developing the skills required to deliver future trials successfully. IB was an early recipient of this scheme and in this chapter describes the lessons learned as a clinician based in The Institute of Cancer Research's Clinical Trials and Statistics Unit (ICR-CTSU) working within the team undertaking a portfolio of breast radiotherapy trials.

7.2 Introduction

As part of a Cancer Research UK (CRUK) 'Clinical Trials Fellowship', IB was seconded to The Institute of Cancer Research's Clinical Trials and Statistics Unit (ICR-CTSU), specifically working within the breast cancer radiotherapy trials portfolio on the PRIMETIME study [16] and IMPORT trials [15,40]. The CRUK 'Clinical Trials Fellowships' embed clinicians in clinical trials units (CTU) for a period of 1-3 years enabling the Fellow to gain experience of trials and trials methodology issues across a range of trials including those in early-stage development, in recruitment and in extended long term follow-up. In particular, the Fellowship provides an in-depth understanding of the scientific, logistical and regulatory requirements of research delivery across the complete trial lifecycle equipping the Fellow with the skills required to lead future research. These Fellowships also offer benefits for the CTU as Fellows can provide clinical expertise and develop sub-studies as part of a wider effort to ensure trials deliver maximal

outputs. These fellowships enable greater translational depth from the data collected from trials i.e. ensuring we can learn as much as possible from patients on clinical trials as per CRUK's statement of intent [146]. The CRUK 'Clinical Trials Fellowship' can facilitate effective collaboration between clinicians and CTU which can improve and streamline clinical trials.

7.3 Why do we need clinical trials and what are the practicalities?

Clinical trials are required to identify optimal treatment options for patients. However, as IB discovered, clinical trials are multi-faceted. Trial participation can fast-track implementation of new technologies or processes within a stringent quality-assured framework where appropriate support or technical advice can be provided to centres giving reassurance and safety when implementing a new technique [147]. For example, the implementation of intensity-modulated radiotherapy (IMRT) and best-practice guidance regarding surgical tumour-bed clips within IMPORT LOW [15] was a 'driver for change' within individual departments and benefitted non-trial patients long before the primary endpoint was reported. Furthermore, trial participation can facilitate cross-centre collaboration and networking where specific issues such as equipment-related problems or patient pathway issues can be discussed and resolved with centres learning from one another [147]. Thereby, trial participation can result in improvements in clinical practice across all participating sites.

Also, with respect to translational research, several immunohistochemistry and molecular profiling techniques have been developed in order to enable precision medicine. However, results of the efficacy of these techniques are highly dependent on the quality of baseline and follow-up clinical and tissue-based data collected. Therefore, complete and comprehensive data sets are needed. This requires stringent processes to be in place to safely manoeuvre biological samples between hospital sites and laboratories, and ensure the flow of accurate data back from the laboratory to the CTU, and then back to the site. Ensuring these processes occur in a timely manner is especially important in biomarker-directed studies where patient treatment is dependent on the results

of these biological samples. In order to ensure that samples could be processed in a timely manner with no delay to treatment decisions, the PRIMETIME study [16] was designed in two stages; the pre-screening stage and main study stage. The sample provision for Ki-67 was obtained at breast cancer diagnosis, whereas patient eligibility was only confirmed after the patient's breast conserving surgery at which time the IHC4+C result (incorporating Ki-67) could be determined quickly to direct treatment without delays.

Another consideration is that radiotherapy trials are almost exclusively led by academically funded CTU, utilising a resource-limited model compared with trials led by the pharmaceutical or technology industry. Similarly, resources are limited within the NHS for site participation in trials, and there is an ongoing shortage of research staff and resources in sites. As well as a lack of research practitioners who are responsible for the recruitment of patients at sites, there are also shortages in radiology and pathology time dedicated to research. Despite these limitations, optimising the quality of clinical trials is paramount.

7.4 What is the role of a CTU?

CTU are specialist multi-disciplinary academic units, usually university-based, with the specific remit to design, conduct, analyse and publish clinical trials. The CTU is an academic partner, providing specific statistical, epidemiological and other methodological, project and data management expertise to undertake clinical trials successfully. Early collaborations between clinicians and CTU are essential. As a clinician with a trial concept, it is essential to start discussions and collaborate with a CTU early who, based on their expertise and experience, will advise whether the concept is feasible and whether it will work in the 'real world'.

7.5 What is the role of the patient advocate?

The patient plays an integral role in the clinical trials process. In the contemporary era, the patient advocate is a partner in healthcare research and is involved throughout the lifespan of the trial. At trial inception, the advocate is

involved in prioritising and defining the clinical question to be asked. Advocates should be also involved in the development of the protocol as well as review of patient information materials for which they have historically been responsible for. They may also have a role in securing regulatory as well as Ethics approvals. Following trial implementation, patient advocates can have a role in improving patient recruitment. After trial data have been analysed, patient advocates may co-author publications reporting endpoints and develop lay summaries of results for patients in collaboration with the trials team.

Patient advocates were key to the concept of the PRIMETIME study [16]. They advised that the gold standard trial design of a randomised controlled trial (RCT) would not be acceptable to patients, where patients would not want to be randomised to 'endocrine therapy only'. However, a biomarker directed prospective cohort study, where patients' tumour biology (using IHC4+C) is used to direct treatment would be acceptable. Also, it was primarily the patient advocates who set the acceptable threshold of a 5-year ipsilateral disease rate of ≤4% for the selective de-escalation of radiotherapy [65].

Information delivery to patients in de-escalation studies needs careful consideration. The PRIMETIME patient advocates worked closely with the PRIMETIME trialists to develop the written information materials for PRIMETIME. As part of her Fellowship, and in close collaboration with the patient advocates, IB developed a patient decision aid (PDA) video which was tested within the PRIMETIME Information Giving Study (IGS). The PRIMETIME IGS investigated if the addition of a PDA video to standard written information reduced patient uncertainty (decisional conflict) regarding PRIMETIME entry. During the video content development process, the advocates advised that it was important for patients to understand the risk of recurrence rather than focus on details of the biomarker (IHC4+C) to determine recurrence risk. This highlighted to IB that what may seem important to a clinician, may not be so relevant to the patient. Also, International Patient Decision Aid Standards (IPDAS) guidelines state that PDA should contain a values-clarification exercise [88]. However, the advocates felt that this would be an added piece of work for patients and there would be insufficient time for patients to discuss this with healthcare professionals in busy NHS clinics. On reflection, given that the

majority of patients who had access to the PDA video in the IGS did not watch it, it is likely that the values-clarification exercise would not have been completed. As mentioned above, information should be tailored to the patient's needs and not to expectations of healthcare professionals even if mandated in guidelines.

7.6 What is the role of the clinician?

The role of the clinician is two-fold. Firstly, the clinician who is the chief investigator (CI) will provide scientific and clinical expertise, identifying the important clinical questions which need to be answered in collaboration with the patient and CTU, and have responsibility for the trial from a regulatory as well as a scientific perspective. Secondly, the site (hospital) level Principal Investigators (PI), provide oversight for trial conduct at their site, and this includes ensuring informed consent is secured for all patients, protocol adherence and that principles of Good Clinical Practice are followed.

The role of the PI is especially important in the current NHS climate given constraints in research nurse, radiology and pathology availability. Issues regarding resources and problems with trial set-up and recruitment should be communicated early to the CTU. Effective communication between sites and CTU may enable issues to be resolved in a timely manner and streamline processes. For example within the PRIMETIME study [16], pathology research availability at sites was raised as a barrier to trial set-up. In response to this, a pathology training day was arranged. Effective discussions at this meeting enabled a number of pathology related issues to be resolved. A network of pathologists was then established in which those who had successfully resolved site set-up issues partnered with those reporting ongoing set-up issues.

Another responsibility of clinicians is the acquisition of trial follow-up data and in particular, collection of normal tissue effect (NTE) data which is especially important for radiotherapy trials. These NTE may be assessed by clinicians or patients. NTE data assessed by clinicians is usually collected via the case-report form (discussed in the following section), whereas patient-reported NTE may be collected by questionnaires (either paper-based or electronic) which would be

sent to the patient directly by post or distributed in clinic by research staff. Patient-reported outcomes (PRO) provide valuable information on the patient's experience of treatment and the importance of collecting complete and timely data was apparent when IB investigated PRO in IMPORT LOW over 5-years [100]. Results showed most side-effects reduced over time providing reassurance for women considering moderately-fractionated breast radiotherapy to either the whole or partial-breast. However, one of the challenges with PRO data is in obtaining complete datasets, as whole questionnaires may not be returned or individual questionnaire items may be missing. Given the importance of PRO data, it is important for clinicians to encourage patients to complete questionnaires as fully as possible enabling high quality PRO data to be collected, benefitting future patients.

7.7 Case Report Form Design and Analysis

Case report forms (CRF) are the documents completed for each patient on each respective hospital visit in relation to the time-point of a specific trial. CRF are the single-most important data document sent from the site to the CTU. Data is uploaded onto a database and cleaned by data managers. Data cleaning describes the process of ensuring that the data is correct, consistent and useable by identifying errors such as missing, incorrect or inaccurate data. Following this, the data is analysed by trial statisticians. It is the CRF that contain the data required to analyse trial endpoints such as recurrence, survival and toxicity, thereby it is important that this form is completely fully, accurately and in a timely manner.

Previously, CRF completion may have been assigned to junior staff, who may not have fully appreciated the far-reaching benefit of clinical trials, and the importance of completing CRF 'real-time'. Retrospective CRF completion following a busy clinic may be inaccurate. CRF completion with patients in clinic allow questions to be framed appropriately to identify possible adverse effects (AE) and explore if related to the intervention. For example, is a rib fracture within the radiotherapy field? In order to determine this, information regarding laterality of the suspected rib fracture is required.

During the fellowship IB was asked to review specific AE for the IMPORT HIGH Interim Data Monitoring Committee (IDMC) reports. In particular IB was required to identify whether suspected rib fractures, lung fibrosis, ischaemic heart disease and persistent cough were due to radiotherapy. It was found that data fields regarding laterality of suspected rib fractures and lung fibrosis were frequently missing. Also, for lung fibrosis the laterality of fibrosis was also missing. Furthermore, details of whether or not these AE had been confirmed on imaging were also missing.

Despite contacting sites for missing information, it usually could not be retrieved as it was not always documented in clinical notes and there was often a high turnover of research staff in centres where staff who saw patients may have left by the time queries were raised. This missing data has implications for the patient in question, and also for future patients for whom complete trial data are crucial to the interpretation of a clinical trial and its impact on the standard of care.

As well as ensuring data fields within each CRF are correctly completed, the correct CRF must be completed. IB also reviewed events in the IMPORT LOW trial [15] including second cancers and deaths. It was found that the incorrect CRF was completed on a number of occasions. For example, a ‘second primary’ CRF was completed when a ‘disease recurrence’ form was required, which can lead to misreporting. It is also important that the CRF has no ambiguity and is easily interpretable. Guidance from the site PI aiding local members of staff in completing CRF accurately could reduce the time spent by CTU and site staff dealing with data queries.

The trial guidance documented in the trial protocol should also be understood and adhered to. There is a risk of informative censoring of data if only a sub-set of patients have CRF completed promptly as those patients may be characteristically different compared with trial patients with no CRF completed, resulting in a biased population. Furthermore, if outcomes are reported at time-points different to those specified in the trial protocol, there is a risk of inaccurate reporting and inflation or underestimation of the frequency of a given event. For example, if local recurrences are reported in real time at 15 months rather than

at the expected annual CRF return (i.e. 24 months), the denominator of patients at 15 months will be considerably less than at 24 months, such that the recurrence rate is inflated if this has not been foreseen in the analysis strategy.

7.8 Radiotherapy CT planning data

In UK radiotherapy trials, the National Cancer Research Institute (NCRI) Radiotherapy Trials Quality Assurance (RTTQA) group develop and implement radiotherapy quality assurance programmes for all National Institute for Health Research Clinical Research Network portfolio trials with a radiotherapy component [147]. As part of this quality assurance programme radiotherapy data is collected by RTTQA [148]. Participating sites use a number of treatment planning systems and therefore the nature of the data will vary. However, for the purposes of analysis, nomenclature of 'structure names' need to be uniform for entry into dosimetry software. For example, structure labels such as the whole-breast planning target volume (PTV) need to be consistently named on every scan acquired as part of the trial. Within IMPORT HIGH, IB reviewed CT planning scans to investigate whether breast seroma was associated with patient-reported breast appearance change at 3 years [149]. It was found that the whole-breast PTV was not named uniformly such that it required re-naming for almost every scan (approximately 500 scans) which was time-consuming and resource-intensive. Most radiotherapy trials will specify the nomenclature required for the trial in the radiotherapy planning pack and this guidance should be followed, thereby allowing efficient utilisation of resources.

7.9 Collaboration is the key to success...

Collaboration with colleagues from multiple disciplines is key to ensure important clinical questions are answered. As oncologists, this includes working closely with other medical disciplines e.g. surgeons, pathologists and radiologists as well as trial methodologists and patients. For example, patients potentially eligible for PRIMETIME need to be approached early in the patient's diagnostic pathway requiring close collaboration with surgeons and pathologists to ensure

that the Ki-67 and IHC4+C results are available in a timely fashion. Furthermore, the manner in which the study is explained to patients by the surgical team is crucial to whether the patient participates. If the surgeon advises the patient that 'they will be seeing the oncologist for radiotherapy' it is then challenging to discuss an avoidance of radiotherapy study with the patient. PRIMETIME is an example of a trial where multi-disciplinary collaboration was crucial to its development. PRIMETIME was closely developed with the surgical community and is an Association of Breast Surgeons (ABS) badged trial with a chief clinical coordinator who is a surgeon as well as several surgical PIs and surgical leads at participating sites.

Also, conducting research via established networks such as the NCRI Breast Clinical Studies Group (CSG) [150] can facilitate these collaborations. As a trainee representative on the NCRI Breast CSG, IB has witnessed first-hand that these networks ensure the research community is working together to answer important clinical questions, and that research is not being conducted in silos. This collaborative working also provides an opportunity for researchers to 'piggyback' onto other trials without having to set up a new trial which is a highly resource intensive process.

7.10 Earlier exposure of clinicians-in-training to CTU

As the number and complexity of clinical trials increases, additional clinician involvement and time is needed [95] to ensure appropriate trial conduct. However, there is a workforce crisis in clinical oncology due in part to a failure to recruit sufficient trainees such that it will be increasingly difficult to balance demands from service versus research on consultants' time. [96]. Furthermore, new consultants have reported that the training programme does not adequately prepare them for the research element of consultant posts [97].

Out-of-programme experience is an opportunity for trainees to gain protected time in research. Traditionally, this has been technical-radiotherapy or lab-based. However, Fellowships such as that described above allow trainees to be

embedded within CTU, obtaining day-to-day experience and conduct of clinical trials whilst also studying for a PhD.

For CTU, the Fellowship enables training of future leaders including the next generation of chief investigators, principal investigators and researchers who begin their consultant careers knowledgeable in the skills required to undertake high-quality clinical research and to understand and appreciate the multi-disciplinary team science involved.

7.11 Conclusions

Clinical trials determine optimal treatment options for patients. Clinicians and CTU share the overarching aim of improving patient care. Open dialogue and effective communication between clinicians and CTU will facilitate and streamline this process to ensure high-quality and efficient conduct of clinical trials.

Chapter 8 Final discussion and future directions

Adjuvant breast radiotherapy, like many treatments, has a number of risks and benefits. With falling local relapse rates, the absolute benefit of radiotherapy has reduced such that the risk-benefit ratio needs careful consideration. It is important that data regarding the risks of radiotherapy including normal tissue effects (NTE) related to the treated breast are accurately collected, and the collection of these data optimised. Furthermore, in some patients with a very low risk of local relapse, the risks of radiotherapy are now difficult to justify given the minimal local control benefit and, for these patients, omission of radiotherapy may be appropriate. This thesis has focussed on a number of exploratory analyses investigating how NTE data collection can be optimised using patient-reported outcome (PRO) data and what these PRO data tell us about the evolution of NTE with time. Optimisation of the information provided to patients considering omission of radiotherapy has also been investigated and is discussed below.

In Chapter 2 it was investigated whether PRO could be used as primary NTE endpoints in breast radiotherapy trials. It was found that in fact very few patients actually reported NTE. Patients reported more NTE than from clinician-reported outcomes (CRO) and photographs suggesting NTE would be underestimated if PRO were not to be used in future studies [125]. Although concordance was poor between the methods, the effect sizes from PRO were consistent with the other methods suggesting it reasonable that PRO be used as NTE endpoints in breast radiotherapy trials [125]. Furthermore PRO provide the patient's perspective of the toxicity experienced. However, one of the main challenges when using PRO data is the risk of reporting bias as inevitably not all patients will provide PRO at the specified time-points during trial follow-up. It was demonstrated that certain groups of patients were more or less likely to return questionnaires at certain time-points. For example in IMPORT LOW, there were higher baseline HADS anxiety and depression subscale scores in those who did not return their questionnaires at 5 years. Also, patients who reported at least one adverse effect (AE) at 2 years were more likely to return questionnaires at 5 years, such that it is possible that the prevalence of AE was overestimated in this analysis. It is therefore important to ensure that the PRO data collected is representative of

the whole trial population and that all trial patients are encouraged to complete questionnaires.

In order to ensure that PRO data are collected in a complete and timely manner, the required data could be limited to only the salient and discriminatory questions. However, this is challenging as questionnaires should only be used in the manner in which they have been validated. For example, if the EORTC BR23 (breast-specific) is to be used, it is advised that it is completed with the EORTC-C30 (generic cancer questionnaire) [25,26]. Yet, in patients receiving breast radiotherapy only (with no chemotherapy), the questionnaire items in the EORTC-C30 are often irrelevant. In fact, questions in the EORTC BR23 are also not breast radiotherapy specific. The START trialists therefore developed protocol-specific items which were designed to capture breast radiotherapy effects that were not assessed in the EORTC questionnaires [13]. These were added to by the IMPORT trialists. In particular the patient advocates added questions regarding bra fitting and nipple position. One option for future breast radiotherapy trials may be to include only the EORTC BR23 and protocol-specific items. It has been acknowledged that it can be a burden to expect numerous questionnaire items to be completed and the EORTC have developed a questionnaire bank where researchers can develop their own questionnaires with specific items tailored to their patients [142]. However, it should be noted that in IMPORT LOW only patient-reported breast appearance change and induration differentiated between treatment groups. Furthermore, the two protocol-specific items developed by patient advocates were not able to differentiate between treatment groups. This may of course be related to the reduction in NTE over time, but it does create a challenge in identifying which are the most appropriate PRO to use.

As well as streamlining the questions asked, the manner in which the data are collected could be improved to increase questionnaire return rates. For example, in the IMPORT LOW trial (and in many other trials) PRO data were collected via paper-based questionnaires which were completed in the clinic at baseline, and subsequently posted to the patient's home over the 5 years following radiotherapy. An alternative approach would be to provide other platforms for patients to complete questionnaires, for example via an app which is

downloadable on an electronic device or via a web-link sent to the patient's email address. It should be acknowledged however, that not all patients will have the same preferences in terms of how they would most prefer to receive a questionnaire. The options to have paper-based and electronic data completion tools should be available.

After establishing that PRO could potentially be used as primary toxicity endpoints in breast radiotherapy trials, Chapter 3 investigated the way in which PRO developed over time and identified baseline predictors of AE. Consistent with other breast radiotherapy studies it was found that most AE reduce over time with the exception of breast shrinkage which increased over time [13] [14] . Given that patient-reported breast shrinkage increased over time, breast shrinkage specifically may be a relevant toxicity endpoint as it has a reasonable event rate and also a feasible time-point for assessment of 5 years. However, patient-reported breast appearance change should also be considered as an alternative toxicity endpoint as this AE was consistently stable over time at a rate of around 20%.

Baseline predictors of patient-reported AE were identified which can be discussed during the informed consent process for patients considering adjuvant breast radiotherapy. For example, it was found that younger women were more likely to report AE pertaining to worse body image [100]. Women with larger breasts were also more likely to report AE [100]. Furthermore, women with higher levels of anxiety and depression as measured by HADS at baseline, were more likely to report worse AE. Psychosocial support could be arranged for patients with higher levels of anxiety and depression as measured on HADS albeit that it is not known whether psychosocial intervention in these patients would affect their future reporting of AE.

To further investigate factors predicting patient-reported AE, the association of breast seroma with breast appearance change was investigated using case-control methodology in Chapter 4. In contrast to other trials, an association between seroma and NTE was not found [149]. In particular, this may have been related to the prevalence of seroma being lower than described in previous reports which in turn may be due to the large proportion of patients receiving

chemotherapy in whom the seroma may have resolved prior to radiotherapy. Furthermore, changes in surgical practice with an increase in oncoplastic techniques and local glandular mobilisation may have contributed to the reduced seroma rates. Nevertheless, the study did find that larger tumour size, haematoma, current smoking and greater body image concerns were associated with breast appearance change [149]. These results suggest that measures need to be taken to reduce the risk of haematoma formation. Also, smoking cessation should be encouraged, although we cannot determine the time interval required from smoking cessation to start of radiotherapy to reduce the risk of patient-reported breast appearance change.

With respect to the study design, patient-reported breast appearance was used to determine the 'cases'. It is acknowledged that with any PRO, there is risk of reporting bias as there will not be complete data sets for all eligible patients. In this study 836/1078 (78%) patients responded to whether they had a change in breast appearance at 3 years. However, the benefit of using PRO over CRO or photographs, was that patients reported more NTE than CRO/ photographs and by not using PRO, the NTE may have been underestimated. The case-control methodology also enabled a resource-efficient study design of whether seroma was associated with patient-reported breast appearance change as it meant approximately 500 patients' scans being reviewed instead of over 2000 scans.

Similarly, a resource-efficient study design was used in Chapters 5 and 6. In Chapter 5, a patient decision aid (PDA) video was developed to be used in addition to the standard patient information with the aim of reducing decisional conflict in patients. This was tested using a Study Within A Trial (SWAT) concept within the PRIMETIME study (The PRIMETIME Information Giving Study-IGS). The IGS used a cluster stepped-wedge trial design which ensured all patients at a site received uniform information and by the end of the study all sites would have use of the PDA. The IGS was implemented across multiple UK centres participating within PRIMETIME. All patients who were eligible for the PRIMETIME main study were eligible for the IGS. This involved sites distributing questionnaires to patients who consented *as well as* patients who declined PRIMETIME. As there were not always established pathways at sites for dedicated research staff to follow-through with patients who declined clinical

trials, this was discussed with sites up-front. The timings of when information would be provided and questionnaires distributed were established with each site. It was important that data were obtained for all eligible patients as, in a study investigating decisional conflict, it would be important to capture patients who declined trial entry. If data were only obtained for patients who consented to PRIMETIME, there was a risk that the data would not be unbiased.

The results of the IGS were discussed in Chapter 6. It was apparent from the interim results of the PRIMETIME IGS that overall, the absolute levels of decisional conflict were low such that there was less scope for obtaining any substantial reductions. This suggests that this population at low risk of local recurrence do not have high levels of uncertainty regarding the management of their breast cancer. Historically there has been an attitude that 'more treatment is better' and 'we have to fight the cancer'. However, it may be that now patients are becoming increasingly aware of the potential for overtreatment in early breast cancer which is likely related to media discussions around the screening programme and contributions from patient advocates regarding clearer documentation of the prevalence and evolution of NTE.

It was also found that only 50% of patients actually watched the PDA video. It may be the case that the standard information was sufficient for this population of patients. With respect to the role of information provision in treatment de-escalation studies, having multiple formats of information which need to be used in combination may be too overwhelming for patients, especially when they are early on in their breast cancer diagnosis. The optimal information format is unknown and there is unlikely to be a consensus on this given varying individual patient preferences. In the future, stand-alone pieces of information could be developed in both written and video format so that patients can choose which information is most easily interpreted by them. The cost implications will need to be considered and accounted for in trial grant applications. Also, the information regardless of which format will need to be reviewed by the Ethics committee.

The vast majority of patients who consented to the IGS also consented to PRIMETIME with limited data collected for patients who declined the main study. The decisional conflict scores may reflect the biased population who were all

participating in the main study and perhaps their uncertainty was reduced knowing they would be monitored closely within a clinical trial. In terms of the population of patients who consented to the IGS, verbal correspondence with sites did highlight some reasons why questionnaires may not have been distributed. In order to investigate this formally, a 'service evaluation' questionnaire was developed for research practitioners at sites to complete, investigating the site's experience of the running of the IGS. This has been sent to all sites but the results are not expected until late August 2019. The results of this questionnaire may formally highlight specific issues that sites experienced whilst running the IGS and inform the running of future SWATs. Optimising the running of studies that investigate patient's attitudes and behaviours, as well as those investigating 'hard' cancer efficacy outcomes will ultimately benefit future patients.

Finally in Chapter 7, the lessons learned whilst working within the trials teams during the CRUK Clinical Trials Fellowship were described. It is known that the number and complexity of clinical trials is growing as we strive to obtain the best possible treatment options for patients. However, there is limited training in the running of clinical trials for clinical oncology trainees. For example, the importance of accurate and timely completed case report forms was highlighted during the fellowship. Also, the necessity for multidisciplinary team working to run trials successfully became apparent. This fellowship has enabled IB to obtain the relevant experience and develop the skills required in the conduct and analysis of clinical trials to lead future trials. Ongoing open dialogue and effective communication between clinicians and clinical trials units is essential to ensure the running of high-quality and efficient clinical trials.

References

1. Darby S, McGale P, Correa C, et al: Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Lancet 378:1707-16, 2011
2. Haviland JS, Owen JR, Dewar JA, et al: The UK Standardisation of Breast Radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials. Lancet Oncol 14:1086-94, 2013
3. Whelan TJ, Pignol JP, Levine MN, et al: Long-term results of hypofractionated radiation therapy for breast cancer. N Engl J Med 362:513-20, 2010
4. Smith BD, Bellon JR, Blitzblau R, et al: Radiation therapy for the whole breast: Executive summary of an American Society for Radiation Oncology (ASTRO) evidence-based guideline. Pract Radiat Oncol, 2018
5. Mannino M, Yarnold JR: Local relapse rates are falling after breast conserving surgery and systemic therapy for early breast cancer: can radiotherapy ever be safely withheld? Radiother Oncol 90:14-22, 2009
6. Davies C, Pan H, Godwin J, et al: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 381:805-16, 2013
7. Stanton AL, Krishnan L, Collins CA: Form or function? Part 1. Subjective cosmetic and functional correlates of quality of life in women treated with breast-conserving surgical procedures and radiotherapy. Cancer 91:2273-81, 2001
8. Darby SC, Ewertz M, McGale P, et al: Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med 368:987-98, 2013
9. Grantzau T, Overgaard J: Risk of second non-breast cancer after radiotherapy for breast cancer: a systematic review and meta-analysis of 762,468 patients. Radiother Oncol 114:56-65, 2015
10. Bottomley A, Pe M, Sloan J, et al: Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. Lancet Oncol 17:e510-e514, 2016
11. Lemieux J, Goodwin PJ, Bordeleau LJ, et al: Quality-of-life measurement in randomized clinical trials in breast cancer: an updated systematic review (2001-2009). J Natl Cancer Inst 103:178-231, 2011
12. Vodicka E, Kim K, Devine EB, et al: Inclusion of patient-reported outcome measures in registered clinical trials: Evidence from ClinicalTrials.gov (2007-2013). Contemp Clin Trials 43:1-9, 2015
13. Hopwood P, Haviland JS, Sumo G, et al: Comparison of patient-reported breast, arm, and shoulder symptoms and body image after radiotherapy for early breast cancer: 5-year follow-up in the randomised Standardisation of Breast Radiotherapy (START) trials. Lancet Oncol 11:231-40, 2010

14. Mukesh MB, Qian W, Wilkinson JS, et al: Patient reported outcome measures (PROMs) following forward planned field-in field IMRT: results from the Cambridge Breast IMRT trial. Radiother Oncol 111:270-5, 2014
15. Coles CE, Griffin CL, Kirby AM, et al: Partial-breast radiotherapy after breast conservation surgery for patients with early breast cancer (UK IMPORT LOW trial): 5-year results from a multicentre, randomised, controlled, phase 3, non-inferiority trial. Lancet 390:1048-1060, 2017
16. Kirwan CC, Coles CE, Bliss J: It's PRIMETIME. Postoperative Avoidance of Radiotherapy: Biomarker Selection of Women at Very Low Risk of Local Recurrence. Clin Oncol (R Coll Radiol) 28:594-6, 2016
17. Kirby AM, Bhattacharya IS, Wilcox M, et al: The IMPORT LOW Trial: Collaborative Research Accelerates Practice Change in Breast Radiotherapy. Clin Oncol (R Coll Radiol) 31:5-8, 2019
18. Mukesh MB, Barnett GC, Wilkinson JS, et al: Randomized controlled trial of intensity-modulated radiotherapy for early breast cancer: 5-year results confirm superior overall cosmesis. J Clin Oncol 31:4488-95, 2013
19. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. Silver Spring, MD, USA: Food and Drug Administration/ National Institutes of Health 2016
20. European Medicines Agency. Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man: the use of patient-reported outcome (PRO) measures in oncology studies. 2016. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2016/04/WC500205159.pdf.
21. Faithfull S, Lemanska A, Chen T: Patient-reported Outcome Measures in Radiotherapy: Clinical Advances and Research Opportunities in Measurement for Survivorship. Clin Oncol (R Coll Radiol) 27:679-85, 2015
22. Speight J, Barendse SM: FDA guidance on patient reported outcomes. Bmj 340:c2921, 2010
23. Black N, Jenkinson C: Measuring patients' experiences and outcomes. Bmj 339:b2495, 2009
24. Luckett T, King MT: Choosing patient-reported outcome measures for cancer clinical research--practical principles and an algorithm to assist non-specialist researchers. Eur J Cancer 46:3149-57, 2010
25. Aaronson NK, Ahmedzai S, Bergman B, et al: The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 85:365-76, 1993
26. Sprangers MA, Groenvold M, Arraras JI, et al: The European Organization for Research and Treatment of Cancer breast cancer-specific quality-of-life questionnaire module: first results from a three-country field study. J Clin Oncol 14:2756-68, 1996
27. Zigmond AS, Snaith RP: The hospital anxiety and depression scale. Acta Psychiatr Scand 67:361-70, 1983
28. Hopwood P, Fletcher I, Lee A, et al: A body image scale for use with cancer patients. Eur J Cancer 37:189-97, 2001
29. Cella D, Stone AA: Health-related quality of life measurement in oncology: advances and opportunities. Am Psychol 70:175-85, 2015
30. FACT-B Questionnaires http://www.facit.org/facitorg/questionnaires (accessed 6th August 2019).

31. Ong WL, Schouwenburg MG, van Bommel ACM, et al: A Standard Set of Value-Based Patient-Centered Outcomes for Breast Cancer: The International Consortium for Health Outcomes Measurement (ICHOM) Initiative. JAMA Oncol 3:677-685, 2017

32. Agrawal RK, Alhasso A, Barrett-Lee PJ, et al: First results of the randomised UK FAST Trial of radiotherapy hypofractionation for treatment of early breast cancer (CRUKE/04/015). Radiother Oncol 100:93-100, 2011

33. Brunt AM, Wheatley D, Yarnold J, et al: Acute skin toxicity associated with a 1-week schedule of whole breast radiotherapy compared with a standard 3-week regimen delivered in the UK FAST-Forward Trial. Radiother Oncol 120:114-8, 2016

34. Luckett T, King MT, Butow PN, et al: Choosing between the EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical research: issues, evidence and recommendations. Ann Oncol 22:2179-90, 2011

35. Trotti A, Colevas AD, Setser A, et al: CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol 13:176-81, 2003

36. Cox JD, Stetz J, Pajak TF: Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). Int J Radiat Oncol Biol Phys 31:1341-6, 1995

37. Harris JR, Levene MB, Svensson G, et al: Analysis of cosmetic results following primary radiation therapy for stages I and II carcinoma of the breast. Int J Radiat Oncol Biol Phys 5:257-61, 1979

38. Aaronson NK, Bartelink H, van Dongen JA, et al: Evaluation of breast conserving therapy: clinical, methodological and psychosocial perspectives. Eur J Surg Oncol 14:133-40, 1988

39. Yarnold J, Ashton A, Bliss J, et al: Fractionation sensitivity and dose response of late adverse effects in the breast after radiotherapy for early breast cancer: long-term results of a randomised trial. Radiother Oncol 75:9-17, 2005

40. Coles CE, Griffin CL, Kirby AM, et al. Dose escalated simultaneous integrated boost radiotherapy for women treated by breast conservation surgery for early breast cancer: 3-year adverse effects in the IMPORT HIGH trial (CRUK/06/003). Presented at: 2018 San Antonio Breast Cancer Symposium; December 4-8, 2018; San Antonio, TX. Abstract GS4-05.

41. Sneeuw KC, Aaronson NK, Yarnold JR, et al: Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 1. Comparison of patients' ratings, observers' ratings, and objective assessments. Radiother Oncol 25:153-9, 1992

42. Sneeuw KC, Aaronson NK, Yarnold JR, et al: Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 2. Relationship with psychosocial functioning. Radiother Oncol 25:160-6, 1992

43. Haviland JS, Ashton A, Broad B, et al: Evaluation of a method for grading late photographic change in breast appearance after radiotherapy for early breast cancer. Clin Oncol (R Coll Radiol) 20:497-501, 2008

44. Cardoso JS, Pinto da Costa JF, Cardoso MJ: Modelling ordinal relations with SVMs: An application to objective aesthetic evaluation of breast cancer conservative treatment. Neural Netw 18:808-17, 2005

45. Cardoso JS, Cardoso MJ: Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment. Artif Intell Med 40:115-26, 2007

46. Cardoso MJ, Cardoso J, Amaral N, et al: Turning subjective into objective: the BCCT.core software for evaluation of cosmetic results in breast cancer conservative treatment. Breast 16:456-61, 2007

47. Haviland JS, Hopwood P, Mills J, et al: Do Patient-reported Outcome Measures Agree with Clinical and Photographic Assessments of Normal Tissue Effects after Breast Radiotherapy? The Experience of the Standardisation of Breast Radiotherapy (START) Trials in Early Breast Cancer. Clin Oncol (R Coll Radiol) 28:345-53, 2016

48. Mukesh MB, Qian W, Wah Hak CC, et al: The Cambridge Breast Intensity-modulated Radiotherapy Trial: Comparison of Clinician- versus Patient-reported Outcomes. Clin Oncol (R Coll Radiol) 28:354-64, 2016

49. Deshields TL, Reschke A, Walker MS, et al: Psychological status at time of diagnosis and patients' ratings of cosmesis following radiation therapy for breast cancer. J Psychosoc Oncol 25:103-16, 2007

50. Clough KB, Kaufman GJ, Nos C, et al: Improving breast cancer surgery: a classification and quadrant per quadrant atlas for oncoplastic surgery. Ann Surg Oncol 17:1375-91, 2010

51. Mukesh MB, Barnett G, Cumming J, et al: Association of breast tumour bed seroma with post-operative complications and late normal tissue toxicity: results from the Cambridge Breast IMRT trial. Eur J Surg Oncol 38:918-24, 2012

52. Juneja P, Bonora M, Haviland JS, et al: Does breast composition influence late adverse effects in breast radiotherapy? Breast 26:25-30, 2016

53. Peterson D, Truong PT, Parpia S, et al: Predictors of adverse cosmetic outcome in the RAPID trial: an exploratory analysis. Int J Radiat Oncol Biol Phys 91:968-76, 2015

54. Collette S, Collette L, Budiharto T, et al: Predictors of the risk of fibrosis at 10 years after breast conserving therapy for early breast cancer: a study based on the EORTC Trial 22881-10882 'boost versus no boost'. Eur J Cancer 44:2587-99, 2008

55. Soulos PR, Yu JB, Roberts KB, et al: Assessing the impact of a cooperative group trial on breast cancer care in the medicare population. J Clin Oncol 30:1601-7, 2012

56. Ravdin PM, Siminoff IA, Harvey JA: Survey of breast cancer patients concerning their knowledge and expectations of adjuvant therapy. J Clin Oncol 16:515-21, 1998

57. Fetting JH, Siminoff LA, Piantadosi S, et al: Effect of patients' expectations of benefit with standard breast cancer adjuvant chemotherapy on participation in a randomized clinical trial: a clinical vignette study. J Clin Oncol 8:1476-82, 1990

58. Hughes KS, Schnaper LA, Bellon JR, et al: Lumpectomy plus tamoxifen with or without irradiation in women age 70 years or older with early breast cancer: long-term follow-up of CALGB 9343. J Clin Oncol 31:2382-7, 2013

59. Hughes KS, Schnaper LA, Berry D, et al: Lumpectomy plus tamoxifen with or without irradiation in women 70 years of age or older with early breast cancer. N Engl J Med 351:971-7, 2004
60. Fyles AW, McCready DR, Manchul LA, et al: Tamoxifen with or without breast irradiation in women 50 years of age or older with early breast cancer. N Engl J Med 351:963-70, 2004
61. Baglin T: Communicating benefit and risk. Br J Haematol 146:31-3, 2009
62. Naik G, Ahmed H, Edwards AG: Communicating risk to patients and the public. Br J Gen Pract 62:213-6, 2012
63. NCRI: https://www.ncri.org.uk/case-study/identifying-more-harm-than-good-the-primetime-breast-cancer-study/ (Accessed January 2020),
64. Prescott RJ, Kunkler IH, Williams LJ, et al: A randomised controlled trial of postoperative radiotherapy following breast-conserving surgery in a minimum-risk older population. The PRIME trial. Health Technol Assess 11:1-149, iii-iv, 2007
65. Bhattacharya IS, Kirby AM, Bliss JM, et al: Can Interrogation of Tumour Characteristics Lead us to Safely Omit Adjuvant Radiotherapy in Patients with Early Breast Cancer? Clin Oncol (R Coll Radiol) 30:158-165, 2018
66. LUMINA: https://clinicaltrials.gov/ct2/show/NCT01791829?term=LUMINA+breast+cancer&rank=1 (accessed 1st August 2018),
67. IDEA: https://clinicaltrials.gov/ct2/show/NCT02400190?term=IDEA+breast+cancer&rank=1 (accessed 6th August 2019),
68. PRECISION: https://clinicaltrials.gov/ct2/show/NCT02653755?term=PRECISION+breast+cancer&rank=1 (accessed 1st August 2018),
69. EXPERT: https://clinicaltrials.gov/ct2/show/NCT02889874?term=EXPERT+breast+cancer&rank=1 (accessed 6th August 2019),
70. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. www.healthaffairs.org/WebExclusives/Wennberg_Web_Excl_021302.htm
71. O'Connor AM, Légaré F, Stacey D: Risk communication in practice: the contribution of decision aids. BMJ (Clinical research ed.) 327:736-740, 2003
72. O'Conner AM, Jacobsen MJ. Decisional Conflict: Assessing and Supporting Patients Experiencing Uncertainty About Choices Affecting Their Health. June 2004
73. Keeney RL: Decision analysis: an overview. Oper Res 30:803-38, 1982
74. Sjöberg L: To Smoke or Not to Smoke: Conflict or Lack of Differentiation?, in Humphreys P, Svenson O, Vári A (eds): Advances in Psychology, North-Holland, 1983, pp 383-400
75. Janis IL, & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York, NY, US: Free Press:
76. North American Nursing Diagnosis Association. Tenth Conference for Classification of Nursing Diagnose; 1992; San Diego C:

77. O'Connor AM. Decisional conflict. In: McFarland G, McFarlane, EA, editors. Nursing diagnosis and intervention. 3rd ed. Toronto: The C.V. Mosby Company; 1997. p.486-496., 1997
78. O'Connor AM, Drake ER, Wells GA, et al: A survey of the decision-making needs of Canadians faced with complex health decisions. Health Expect 6:97-109, 2003
79. Cranney A, O'Connor AM, Jacobsen MJ, et al: Development and pilot testing of a decision aid for postmenopausal women with osteoporosis. Patient Educ Couns 47:245-55, 2002
80. Drake ER, Engler-Todd L, O'Connor AM, et al: Development and Evaluation of a Decision Aid About Prenatal Testing for Women of Advanced Maternal Age. J Genet Couns 8:217-33, 1999
81. Stacey D, O'Connor AM, DeGrasse C, et al: Development and evaluation of a breast cancer prevention decision aid for higher-risk women. Health Expect 6:3-18, 2003
82. O'Connor AM, Tugwell P, Wells GA, et al: A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. Patient Educ Couns 33:267-79, 1998
83. Grant FC, Laupacis A, O'Connor AM, et al: Evaluation of a decision aid for patients considering autologous blood donation before open-heart surgery. Cmaj 164:1139-44, 2001
84. O'Connor AM: Validation of a decisional conflict scale. Med Decis Making 15:25-30, 1995
85. O'Connor AM, Stacey D, Entwistle V, et al: Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev:Cd001431, 2003
86. Stacey D, Légaré F, Lewis K, et al: Decision aids for people facing health treatment or screening decisions. The Cochrane database of systematic reviews 4:CD001431-CD001431, 2017
87. Stacey D, Legare F, Col NF, et al: Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev:Cd001431, 2014
88. Elwyn G, O'Connor A, Stacey D, et al: Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. Bmj 333:417, 2006
89. O'Brien MA, Whelan TJ, Villasis-Keever M, et al: Are cancer-related decision aids effective? A systematic review and meta-analysis. J Clin Oncol 27:974-85, 2009
90. Nicholas Z, Butow P, Tesson S, et al: A systematic review of decision aids for patients making a decision about treatment for early breast cancer. Breast 26:31-45, 2016
91. Juraskova I, Butow P, Bonner C, et al: Improving decision making about clinical trial participation - a randomised controlled trial of a decision aid for women considering participation in the IBIS-II breast cancer prevention trial. Br J Cancer 111:1-7, 2014
92. Wong J, D'Alimonte L, Angus J, et al: Development of patients' decision aid for older women with stage I breast cancer considering radiotherapy after lumpectomy. Int J Radiat Oncol Biol Phys 84:30-8, 2012

93. Smith V, Clarke M, Begley C, et al: SWAT-1: The effectiveness of a 'site visit' intervention on recruitment rates in a multi-centre randomised trial. Trials 16:211, 2015
94. Education section - Studies Within A Trial (SWAT). J Evid Based Med 5:44-5, 2012
95. Tharmalingam H, Vinayan A, Anyamene N: UK Training in Clinical Oncology: Tasters, Coasters and the National Recruitment Crisis. Clin Oncol (R Coll Radiol) 30:599-601, 2018
96. Casswell G, Shakir R, Macnair A, et al: UK Training in Clinical Oncology: The Trainees' Viewpoint. Clin Oncol (R Coll Radiol) 30:602-604, 2018
97. Dickson J, Liu D, Bloomfield D: Training in Clinical Oncology and the Transition from Trainee to Consultant: Results of the Royal College of Radiologists' 2015 Post-Certificate of Completion of Training Survey. Clin Oncol (R Coll Radiol) 29:e64-e71, 2017
98. Bowker AH: A test for symmetry in contingency tables. J Am Stat Assoc 43:572-4, 1948
99. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
100. Bhattacharya IS, Haviland JS, Kirby AM, et al: Patient-Reported Outcomes Over 5 Years After Whole- or Partial-Breast Radiotherapy: Longitudinal Analysis of the IMPORT LOW (CRUK/06/003) Phase III Randomized Controlled Trial. J Clin Oncol:Jco1800982, 2018
101. Fromme EK, Eilers KM, Mori M, et al: How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. J Clin Oncol 22:3485-90, 2004
102. Velikova G, Wright P, Smith AB, et al: Self-reported quality of life of individual cancer patients: concordance of results with disease course and medical records. J Clin Oncol 19:2064-73, 2001
103. Basch E, Jia X, Heller G, et al: Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. J Natl Cancer Inst 101:1624-32, 2009
104. Bruner DW, Movsas B, Basch E: Capturing the patient perspective: patient-reported outcomes as clinical trial endpoints. Am Soc Clin Oncol Educ Book:139-44, 2012
105. Groenvold M, Fayers PM, Petersen MA, et al: Breast cancer patients on adjuvant chemotherapy report a wide range of problems not identified by health-care staff. Breast Cancer Res Treat 103:185-95, 2007
106. Quinten C, Maringwa J, Gotay CC, et al: Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. J Natl Cancer Inst 103:1851-8, 2011
107. Stephens RJ, Hopwood P, Girling DJ, et al: Randomized trials with quality of life endpoints: are doctors' ratings of patients' physical symptoms interchangeable with patients' self-ratings? Qual Life Res 6:225-36, 1997
108. Davidson SE, Trotti A, Ataman OU, et al: Improving the capture of adverse event data in clinical trials: the role of the International Atomic Energy Agency. Int J Radiat Oncol Biol Phys 69:1218-21, 2007
109. Vistad I, Cvancarova M, Fossa SD, et al: Postradiotherapy morbidity in long-term survivors after locally advanced cervical cancer: how well

do physicians' assessments agree with those of their patients? Int J Radiat Oncol Biol Phys 71:1335-42, 2008

110. Brouwers P, van Loon J, Houben RMA, et al: Are PROMs sufficient to record late outcome of breast cancer patients treated with radiotherapy? A comparison between patient and clinician reported outcome through an outpatient clinic after 10years of follow up. Radiother Oncol 126:163-169, 2018

111. Kirchheiner K, Nout R, Lindegaard J, et al: Do clinicians and patients agree regarding symptoms? A comparison after definitive radiochemotherapy in 223 uterine cervical cancer patients. Strahlenther Onkol 188:933-9, 2012

112. Bhattacharya IS, Morden JP, Griffin C, et al: The Application and Feasibility of Using Routine Data Sources for Long-term Cancer Clinical Trial Follow-up. Clin Oncol (R Coll Radiol) 29:796-798, 2017

113. Bentzen SM, Agrawal RK, Aird EG, et al: The UK Standardisation of Breast Radiotherapy (START) Trial A of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. Lancet Oncol 9:331-41, 2008

114. Bentzen SM, Agrawal RK, Aird EG, et al: The UK Standardisation of Breast Radiotherapy (START) Trial B of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. Lancet 371:1098-107, 2008

115. Pe M, Dorme L, Coens C, et al: Statistical analysis of patient-reported outcome data in randomised controlled trials of locally advanced and metastatic breast cancer: a systematic review. Lancet Oncol 19:e459-e469, 2018

116. Atherton PJ, Watkins-Bruner DW, Gotay C, et al: The Complementary Nature of Patient-Reported Outcomes and Adverse Event Reporting in Cooperative Group Oncology Clinical Trials: A Pooled Analysis (NCCTG N0591). J Pain Symptom Manage 50:470-9.e9, 2015

117. Hanley JA, Negassa A, Edwardes MD, et al: Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol 157:364-75, 2003

118. Schafer R, Strnad V, Polgar C, et al: Quality-of-life results for accelerated partial breast irradiation with interstitial brachytherapy versus whole-breast irradiation in early breast cancer after breast-conserving surgery (GEC-ESTRO): 5-year results of a randomised, phase 3 trial. Lancet Oncol, 2018

119. Meattini I, Saieva C, Miccinesi G, et al: Accelerated partial breast irradiation using intensity modulated radiotherapy versus whole breast irradiation: Health-related quality of life final analysis from the Florence phase 3 trial. Eur J Cancer 76:17-26, 2017

120. RAGE: https://www.macmillan.org.uk/assets/yesterdays-women-radiotherapy-action-group-exposure-rage.pdf (Accessed January 2020).

121. Olivotto IA, Whelan TJ, Parpia S, et al: Interim cosmetic and toxicity results from RAPID: a randomized trial of accelerated partial breast irradiation using three-dimensional conformal external beam radiation therapy. J Clin Oncol 31:4038-45, 2013

122. Yarnold J, Bentzen SM, Coles C, et al: Hypofractionated whole-breast radiotherapy for women with early breast cancer: myths and realities. Int J Radiat Oncol Biol Phys 79:1-9, 2011

123. Immink JM, Putter H, Bartelink H, et al: Long-term cosmetic changes after breast-conserving treatment of patients with stage I-II breast

cancer and included in the EORTC 'boost versus no boost' trial. Ann Oncol 23:2591-8, 2012

124. Hopwood P, Sumo G, Mills J, et al: The course of anxiety and depression over 5 years of follow-up and risk factors in women with early breast cancer: results from the UK Standardisation of Radiotherapy Trials (START). Breast 19:84-91, 2010

125. Bhattacharya IS, Haviland JS, Hopwood P, et al: Can patient-reported outcomes be used instead of clinician-reported outcomes and photographs as primary endpoints of late normal tissue effects in breast radiotherapy trials? Results from the IMPORT LOW trial. Radiother Oncol 134:220-230, 2019

126. https://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/8-case-control-and-cross-sectional (accessed 6th August 2019).

127. Withers HR, Thames HD, Jr., Peters LJ: A new isoeffect curve for change in dose per fraction. Radiother Oncol 1:187-91, 1983

128. Ray KJ, Sibson NR, Kiltie AE: Treatment of Breast and Prostate Cancer by Hypofractionated Radiotherapy: Potential Risks and Benefits. Clin Oncol (R Coll Radiol) 27:420-6, 2015

129. Sharma R, Spierer M, Mutyala S, et al: Change in seroma volume during whole-breast radiation therapy. Int J Radiat Oncol Biol Phys 75:89-93, 2009

130. Whipp EC, Halliwell M: Magnetic resonance imaging appearances in the postoperative breast: the clinical target volume-tumor and its relationship to the chest wall. Int J Radiat Oncol Biol Phys 72:49-57, 2008

131. Crown A, Scovel LG, Rocha FG, et al: Oncoplastic breast conserving surgery is associated with a lower rate of surgical site complications compared to standard breast conserving surgery. Am J Surg, 2018

132. Sorensen LT: Wound healing and infection in surgery. The clinical impact of smoking and smoking cessation: a systematic review and meta-analysis. Arch Surg 147:373-83, 2012

133. Zimmermann-Nielsen E, Dahl MB, Graversen HP: [Effects of tobacco smoking on the incidence of flap necrosis after mastectomy]. Ugeskr Laeger 159:4974-6, 1997

134. Han K, Yap ML, Yong JH, et al: Omission of Breast Radiotherapy in Low-risk Luminal A Breast Cancer: Impact on Health Care Costs. Clin Oncol (R Coll Radiol) 28:587-93, 2016

135. Cuzick J, Dowsett M, Pineda S, et al: Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. J Clin Oncol 29:4273-8, 2011

136. Brown CA, Lilford RJ: The stepped wedge trial design: a systematic review. BMC Med Res Methodol 6:54, 2006

137. Hemming K, Haines TP, Chilton PJ, et al: The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. Bmj 350:h391, 2015

138. SWAT56: https://www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyR

esearch/SWATSWARInformation/Repositories/SWATStore/ (accessed 6th August 2019),

139. Thompson-Leduc P, Turcotte S, Labrecque M, et al: Prevalence of clinically significant decisional conflict: an analysis of five studies on decision-making in primary care. BMJ Open 6:e011490, 2016

140. https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Decisional_Conflict.pdf (accessed 6th August 2019),

141. Hussey MA, Hughes JP: Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 28:182-91, 2007

142. EORTC: https://qol.eortc.org/manuals/ (accessed 6th August 2019),

143. Whelan TJ, Levine MN, Gafni A, et al: Breast irradiation postlumpectomy: development and evaluation of a decision instrument. J Clin Oncol 13:847-53, 1995

144. Wong JJ, D'Alimonte L, Angus J, et al: What do older patients with early breast cancer want to know while undergoing adjuvant radiotherapy? J Cancer Educ 26:254-61, 2011

145. An introduction to patient decision aids. Bmj 347:f4147, 2013

146. CRUK: https://www.cancerresearchuk.org/funding-for-researchers/clinical-research/clinical-research-statement-of-intent (Accessed January 2020),

147. Miles E, Venables K: Radiotherapy quality assurance: facilitation of radiotherapy research and implementation of technology. Clin Oncol (R Coll Radiol) 24:710-2, 2012

148. RTTQA: http://www.rttrialsqa.org.uk/rttqa/ (accessed 1st December 2018),

149. Bhattacharya IS, Haviland JS, Perotti C, et al: Is breast seroma after tumour resection associated with patient-reported breast appearance change following radiotherapy? Results from the IMPORT HIGH (CRUK/06/003) trial. Radiother Oncol 136:190-196, 2019

150. NCRI Breast Clinical Studies Group https://csg.ncri.org.uk/about-the-csgs/clinical-studies-groups/breast/ (Accessed 15th August 2019).

Abbreviations

ABS - Association of Breast Surgeons

AE - Adverse effects

BIS - Body Image Scale

BCS - Breast conserving surgery

BCTOS - Breast Cancer Treatment Outcomes Scale

CI - Chief Investigator

CRUK - Cancer Research UK

CSG - Clinical Studies Group

CTCAE - Common Terminology Criteria for Adverse Events

CTV - Clinical target volume

CTU - Clinical trials units

CRO - Clinician-reported outcomes

DCIS - Ductal Carcinoma In Situ

EBC - Early breast cancer

EBCTCG - Early Breast Cancer Trialists' Collaborative Group

ER - Oestrogen receptor

EORTC - European Organisation for Research and Treatment of Cancer

GEE - Generalised estimating equation

Gy - Gray

HADS - Hospital Anxiety and Depression subscales

HER2 - Human epidermal growth factor receptor 2

ICC - Intra-cluster correlation

IGS - Information Giving Study

IHC - Immunohistochemistry

IHC4+C - IHC4+Clinical

ICR-CTSU - Institute of Cancer Research's Clinical Trials and Statistics Unit

IPDAS - International Patient Decision Aid Standards

IMRT - Intensity-modulated radiotherapy

IRR - Incidence rate ratio

NCRI - National Cancer Research Institute

NTE - Normal tissue effect

PDA - Patient decision aids

PI - Principal Investigator

PRO - Patient-reported outcomes

PROM - Patient reported outcome measures

PTV - Planning target volume

PR - Progesterone receptor

RCT - Randomised controlled trial

RTOG - Radiation Therapy Oncology Group/ European Organisation for Research

RTTA - Radiotherapy Quality Assurance Team

SD - Standard deviation

SWAT - Study within a Trial

SIB - Simultaneous integrated boost

VODCA - Visualisation and Organisation of Data for Cancer Analysis

WB - Whole breast

Appendices

1.1. The IMPORT LOW trial

(Intensity Modulated and Partial Organ Radiotherapy)

Randomised Trial Testing Intensity Modulated and Partial Organ Radiotherapy following Breast Conservation Surgery for Early Breast Cancer

Chief Investigator: Dr Charlotte Coles (Addenbrookes Hospital NHS Trust)

Chief Clinical Co-ordinator: Dr Anna Kirby (Institute of Cancer Research/The Royal Marsden Hospital NHS Foundation Trust)

Scientific Lead: Professor Judith Bliss (Institute of Cancer Research- Clinical Trials and Statistics Unit)

Trial Statistician: Mrs Clare Griffin (Institute of Cancer Research- Clinical Trials and Statistics Unit)

Sponsor: The Institute of Cancer Research

Funder: Cancer Research UK

ISRCTN12852634

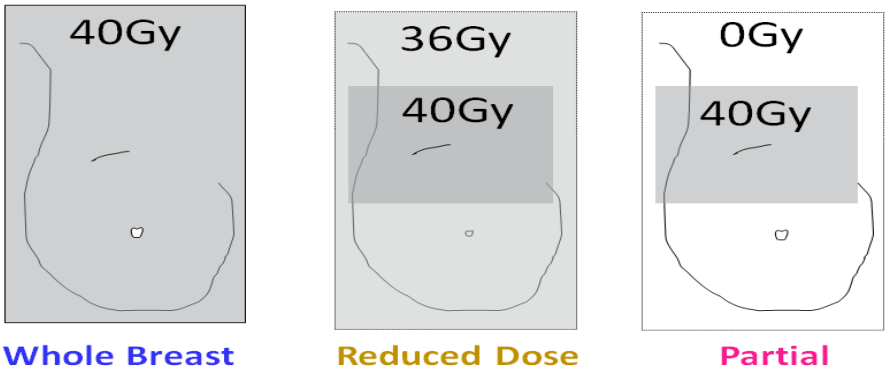
ICR-CTSU/2006/10001

MREC No: Oxford Research Committee B 06/Q1605/128

IMPORT LOW is a multicentre, randomised, controlled, phase III, non-inferiority trial involving 30 radiotherapy centres across the UK. Women aged 50 years or older who had undergone breast conserving surgery (BCS) for unifocal invasive ductal adenocarcinoma of grade 1–3, with a tumour size of 3 cm or less (pT1–2), zero to three positive axillary nodes (pN0–1), and minimum microscopic excision margins of non-cancerous tissue of 2 mm or more, were recruited. Patients were randomly assigned (1:1:1) to receive 40Gy whole-breast radiotherapy (control), 36Gy whole-breast radiotherapy and 40Gy to the partial-breast (reduced-dose group), or 40Gy to the partial-breast only (partial-breast

group) in 15 daily treatment fractions. Patients and clinicians were not blinded to treatment allocation.

Figure 1: Treatment groups in the IMPORT LOW trial



There was a comprehensive collection of normal tissue effect (NTE) data with patient-reported outcomes (PRO) and photographic sub-studies conducted in a subset of patients and clinician-reported outcomes (CRO) in all patients. All centres were invited to participate in the PRO and photographic sub-studies (until sufficient accrual was achieved). All patients at these centres were invited to participate in the sub-studies until the designated sample size for each sub-study was obtained. The protocol-specified sample size for the IMPORT LOW PRO sub-study was a minimum of 400 patients per treatment group, providing >80% power to detect differences between treatment groups of at least 15% in prevalence of adverse effects (AE), allowing for 10% attrition due to death or illness. This was powered in relation to the randomised comparisons previously reported and provided a large dataset for these exploratory analyses. Patient-reported outcome measures (PROM) including the EORTC general cancer scale QLQ-C30 and QLQ-BR23 breast-cancer specific module, Hospital Anxiety and Depression Scale (HADS), 10-item Body Image Scale and protocol-specific questionnaire items were assessed at baseline, 6 months and 1, 2 and 5 years following radiotherapy in a subset of patients. The protocol-specific questionnaire items include asking patients to score ‘change in breast appearance’, ‘breast hardness/firmness’, ‘reduction in size of breast’, ‘change in skin appearance’, ‘is the position of the nipple of your affected breast different from the other side’, ‘problem getting a bra to fit’ and ‘shoulder stiffness’. The protocol-specific questionnaire items were scored on a four-point scale: none, a

little, quite a bit, very much (interpreted as none, mild, moderate, marked). CRO including breast shrinkage, breast induration, telangiectasia and breast oedema were scored using the contralateral breast as a comparator with a four-point graded scale (none, a little, quite a bit, very much; interpreted as none, mild, moderate, marked) at 1, 2 and 5 years following radiotherapy in all patients. Photographs were taken at baseline, 2 and 5 years.

IMPORT LOW recruited 2018 women between 2007 and 2010. Results showed, at a median follow-up of 72 months, partial-breast and reduced-dose radiotherapy were non-inferior to whole-breast radiotherapy in terms of local control. Five year local relapse rates were 1.1% (95% confidence interval 0.5-2.3), 0.2% (0.02-1.2) and 0.5% (0.2-1.4) in the whole-breast, reduced-dose and partial-breast groups respectively. Assessments of NTE showed that the partial-breast radiotherapy technique was associated with significantly lower rates of patient-reported change in breast appearance (15% at 5 years compared with 27% in the whole-breast radiotherapy group) and both the partial-breast and reduced-dose techniques were associated with significantly lower rates of patient-reported breast firmness and hardness than for whole-breast radiotherapy.

Following the reporting of the primary endpoint analysis of IMPORT LOW in 2017, partial-breast radiotherapy has now become the standard of care for patients with low-risk breast cancer following BCS, and has been recommended as per the RCR consensus guidelines.

1.2. The IMPORT HIGH Trial

(Intensity Modulated Partial Organ Radiotherapy)

Randomised trial testing dose escalated intensity modulated radiotherapy for women treated by breast conservation surgery and appropriate systemic therapy for early breast cancer

Chief Investigator: Dr Charlotte Coles (Addenbrookes Hospital NHS Trust)

Chief Clinical Co-ordinator: Dr Anna Kirby (Institute of Cancer Research/The Royal Marsden Hospital NHS Foundation Trust)

Scientific Lead: Professor Judith Bliss (Institute of Cancer Research- Clinical Trials and Statistics Unit)

Trial Statistician: Mrs Clare Griffin (Institute of Cancer Research- Clinical Trials and Statistics Unit)

Sponsor: The Institute of Cancer Research

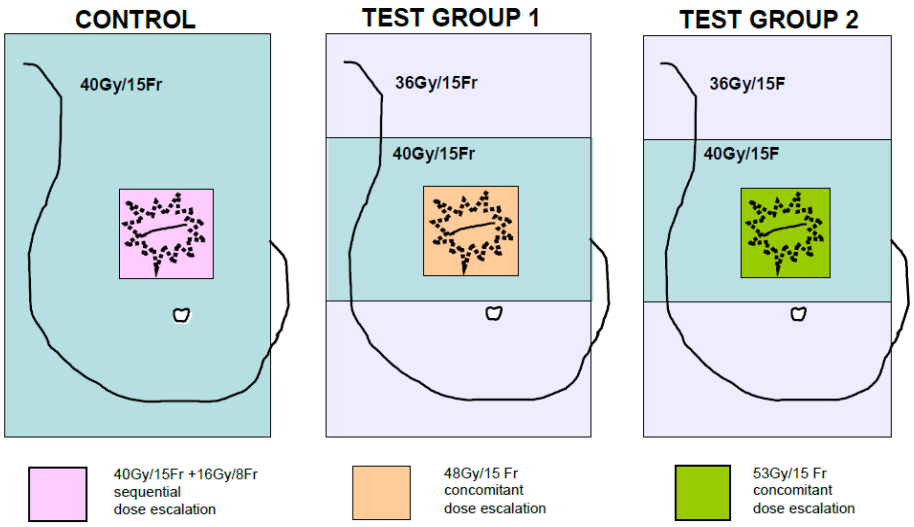
Funder: Cancer Research UK

Main REC Reference Number: 08/H0305/13

ISRCTN 47437448

IMPORT HIGH is a randomised, multi-centre, phase III trial, testing dose escalated simultaneous integrated boost (SIB) against sequential boost each delivered by IMRT for early breast cancer with higher risk of local relapse. Women aged 18 or over after BCS for pT1-3 pN0-pN3a M0 invasive carcinoma were eligible. Randomisation was 1:1:1 between 40Gy/15F to whole breast (WB) + 16Gy/8F sequential photon boost to tumour bed (40+16Gy), 36Gy/15F to WB, 40Gy to partial breast + 48Gy (48Gy) or + 53Gy (53Gy) in 15F SIB to tumour bed. CT planning scan data for all patients recruited into IMPORT HIGH were collected by the Radiotherapy Quality Assurance Team (RTTQA).

Figure 2: Treatment groups in the IMPORT HIGH trial



Normal tissue effect (NTE) data in IMPORT HIGH were collected using patient-reported outcomes (PRO), clinician-reported outcomes (CRO) and photographs. PRO were obtained at baseline, 6 months, 1, 3 and 5 years following radiotherapy. CRO were assessed annually and photographs taken at baseline, 3 and 5 years. Patient-reported outcome measures (PROM) including the EORTC general cancer scale QLQ-C30 and QLQ-BR23 breast-cancer specific module, Hospital Anxiety and Depression Scale (HADS), 10-item Body Image Scale and protocol-specific questionnaire items were assessed at baseline, 6 months and 1, 2 and 5 years following radiotherapy in a subset of patients. The protocol-specific questionnaire items include asking patients to score ‘change in breast appearance’, ‘breast hardness/firmness’, ‘reduction in size of breast’, ‘change in skin appearance’, ‘is the position of the nipple of your affected breast different from the other side’, ‘problem getting a bra to fit’ and ‘shoulder stiffness’. The protocol-specific questionnaire items were scored on a four-point scale: none, a little, quite a bit, very much (interpreted as none, mild, moderate, marked). CRO including breast shrinkage, breast induration, telangiectasia, breast oedema, breast discomfort and breast tenderness on palpation were scored using the contralateral breast as a comparator with a four-point graded scale (none, a little, quite a bit, very much; interpreted as none, mild, moderate, marked) at 1, 2 and 5 years following radiotherapy in all patients.

The trial was initiated with a primary endpoint of breast induration at 3 years. It was initially planned that the year 3 breast induration data would inform a separate phase 3 trial with local control as the primary endpoint. However, training and implementation of the advanced radiotherapy techniques required to deliver IMPORT HIGH took longer than anticipated, and recruitment was lower than planned for the first 2 years. During this time, analyses of induration in other radiotherapy trials (START and FAST trials) confirmed that a safety analysis based on clinician and patient-reported adverse effects at 1 and 2 years would provide reliable indicators of subsequent (5 year) relationships between IMPORT HIGH schedules. Closing IMPORT HIGH to recruitment until 3 year data were collected just as the trial momentum had finally reached its peak, would have risked undermining its ability to answer the question relating to local recurrence. The primary endpoint was therefore amended to local recurrence and patient accrual extended accordingly. By extending the trial and continuing recruitment with the revised primary endpoint of local recurrence (and secondary endpoint of breast induration), the trial was able to continue recruitment in centres without breaking momentum. It was still planned for the 3 year toxicity endpoint to be reported prior to the revised primary endpoint. Between March 2009 and September 2015, 2617 women consented to IMPORT HIGH from 39 radiotherapy centres. The results of the 3 year toxicity analysis found that rates of moderate/marked NTE were broadly similar between the randomised groups; with a suggestion of a slightly increased risk for breast induration in 53Gy compared with the control-group (borderline significance). The results of the local recurrence endpoint are expected in 2020.

1.3. The PRIMETIME Trial

Post-operative avoidance of radiotherapy: biomarker selection of women categorised to be in a very low risk group by IHC4+C

Chief Investigator: Dr Charlotte Coles (Addenbrookes Hospital NHS Trust)

Chief Clinical Co-ordinator: Miss Cliona Kirwan (University of Manchester)

Scientific Lead: Professor Judith Bliss (Institute of Cancer Research- Clinical Trials and Statistics Unit)

Trial Statistician: Mrs Joanne Haviland (Institute of Cancer Research- Clinical Trials and Statistics Unit)

Sponsor: The Institute of Cancer Research

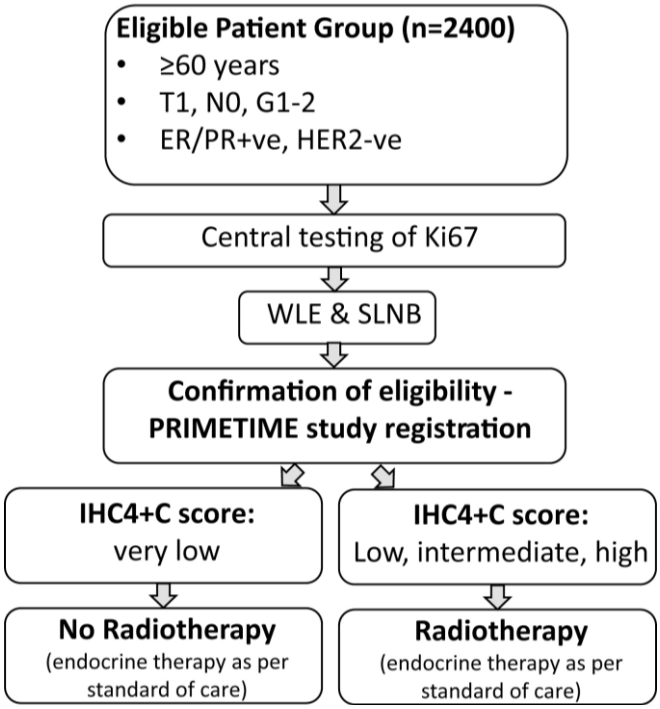
Funder: Cancer Research UK

ISRCTN 41579286

MREC No: 16/EE/0305

PRIMETIME 'Post-operative avoidance of radiotherapy: biomarker selection of women categorised to be in a very low risk group by lHC4+C' is a biomarker directed prospective cohort study aiming to identify a group of breast cancer patients who can safely avoid adjuvant breast radiotherapy following breast conserving surgery. The biomarker being used is IHC4+Clinical (IHC4+C) which combines expression of ER, progesterone receptor (PR), HER2 and Ki-67 with clinico-pathological parameters (tumour size, grade, nodal status, age and endocrine treatment) to identify breast cancer patients at very low, low, intermediate or high risk of distant disease recurrence. Eligibility criteria include women ≥60 years with T1, N0, G1-2, ER+ve and HER2-ve breast cancers who have undergone BCS for invasive disease, with complete resection of tumour tissue.

Figure 3: Summary of patient eligibility and pathway in PRIMETIME



In order to ensure sufficient time for IHC4+C calculation, there are two stages to patient recruitment: 1) pre-operative following diagnostic biopsy and 2) post-operative following definitive surgery and multidisciplinary team confirmation of eligibility. In stage 1, patients are pre-operatively assessed as potentially eligible for study entry and approached before definitive BCS. Following explanation of the PRIMETIME study, consent is sought for sample provision to a central laboratory for IHC4+C testing. In stage 2, after definitive BCS and confirmation of eligibility, patients are offered the option of participating in the main study. Patients who are determined to be 'very low' risk using IHC4+C will be directed to avoid radiotherapy, whilst patients determined to be 'low, intermediate or high' risk will be directed to receive radiotherapy. Patients can decide against the directed treatment and remain within the trial.

The primary endpoint is ipsilateral breast disease rate at 5 years. PRIMETIME requires recruitment of 2400 patients at the preoperative stage, to allow 1550 patients to actively avoid radiotherapy, based on a local relapse rate, in the absence of radiotherapy, of ≤4% at 5 years. The primary endpoint is expected to report in 2021.