

TITLE PAGE

Title: 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

Corresponding author: Dr Indrani S Bhattacharya, FRCR, Clinical Research Fellow, The Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU), 15 Cotswold Road, London, SM2 5NG

Email: indrani.bhattacharya@icr.ac.uk

Author List: Indrani S Bhattacharya* Joanne S Haviland*, Penelope Hopwood, Charlotte E Coles, John R Yarnold, Judith M Bliss** and Anna M Kirby** on behalf of the IMPORT Trialists. **joint first authors, **joint senior authors*

Indrani S Bhattacharya (FRCR) – The Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU) (indrani.bhattacharya@icr.ac.uk)

Joanne S Haviland (MSc) – The Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU) (jo.haviland@icr.ac.uk)

Penelope Hopwood (MD) - The Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU) (penny.hopwood@icr.ac.uk)

Charlotte E Coles (PhD) - Cambridge University Department of Oncology (colesc@doctors.org.uk)

John R Yarnold (FRCR) - The Institute of Cancer Research, Radiotherapy and Imaging (john.yarnold@icr.ac.uk)

Judith M Bliss (MSc) - The Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU) (judith.bliss@icr.ac.uk)

Anna M Kirby (MD) - Royal Marsden NHS Foundation Trust and Institute of Cancer Research, Radiotherapy and Imaging (anna.kirby@rmh.nhs.uk)

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 outcome measures (PROMs), clinician-reported outcomes (CRO) and photographs. This analysis investigates whether PROMs can be used as primary NTE endpoints in breast radiotherapy trials.

Methods

Analyses were conducted within IMPORT LOW (ISRCTN12852634) at 2 and 5 years. NTE were recorded by CRO, photographs and PROMs. 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 PROMS and photographs, PROMs 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 (eg. 16% patients and 4% clinicians reported breast shrinkage at year-5). Patients reported more NTE than assessed from CRO or photographs ($p < 0.001$ for most NTE). Concordance between assessments was poor on an individual patient level; eg. for year-5 breast shrinkage, % agreement=48% and weighted kappa=0.17. Risk ratios comparing radiotherapy schedules were consistent between PROMs and CRO or photographs.

Conclusions

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

250 words

Manuscript

Introduction

In the current era of low local relapse rates after adjuvant breast radiotherapy (1, 2), 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 (3) and partial-breast radiotherapy (1), 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 outcome measures (PROMs) (2, 3). 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 PROMs are able to differentiate between dose/volume regimens (1) and between small dose differences in hypofractionated regimens (2). PROMs also provide the patients' perceptions of the impact of their cancer and the consequences of treatment (4) within the framework of the question asked. This analysis investigates within the context of the IMPORT LOW partial-breast radiotherapy trial, i) the degree of concordance on an individual patient level between PROMs and CRO or photographs, ii) whether results for the randomised comparisons obtained from PROMs are consistent with those using CRO or photographs and iii) the influence of baseline characteristics on concordance, with the overall aim of assessing whether PROMs could be used as primary NTE endpoints in future breast radiotherapy trials.

Methods

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 (1).

IMPORT LOW included a comprehensive and systematic investigation of NTE including CRO in all participants, and PROMs and photographs in a subset of patients for which full details of patients and procedures have been published (1). All centres were invited to participate in the PROMs 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.

Procedures

Patients who consented to the PROMs sub-study completed the EORTC QLQ-C30 core questionnaire and QLQ-BR23 breast-specific module (5, 6) and 10-item Body Image Scale (BIS) (7), all of which asked patients to consider their symptoms during the past week. The Hospital Anxiety and Depression Scale [HADS] (8) 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 (2) 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 PROMs 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 (9).

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 (10). Clinicians were not blinded to treatment group.

Statistical analysis

PROMs were paired with the relevant CRO or photograph at 2 and 5 years for the analyses (table 1).

The “quite a bit” and “very much” categories were combined for PROMs 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 (11). 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 (12). 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 PROMs 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 PROMs 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).

Results

2018 patients were recruited to IMPORT LOW from 71 centres. 2 patients requested exclusion from analysis. In the 41 centres participating in the PROMs sub-study, 1265/1333 (95%) patients consented to PROMs, and 1318/1466 (90%) patients consented to the photographic sub-study from 37 participating centres. 1095 patients consented to both sub-studies (figure 1a).

In 1095 patients who consented to both, PROMs were available at 2 and/ or 5 years for 976 patients of whom 909 had CRO and 844 had photographs. PROMs, CRO and photographs were available for 651 and 518 patients at year-2 (figure 1b) and year-5 respectively (figure 1c). Separate analyses were conducted in patients with PROMs and CRO, and PROMs and photographs, at year-2 and year-5. Data regarding baseline characteristics (1) and PROMs questionnaire return rates (13) have been published.

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). 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, 16% patients and 4% clinicians reported moderate/ marked breast shrinkage. With respect to breast induration, 6% patients and 4% clinicians reported moderate/ marked changes. For breast appearance change, 13% patients reported moderate/ marked changes and photographic assessment reported marked changes in 1%.

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 and 3). Patients and clinicians reported similar prevalences of breast oedema, with very few events at 2 and 5 years. Concordance between PROMs and CRO or photographs of corresponding NTE on an individual patient basis was generally poor (table 2).

For breast shrinkage at year-5, patients reported more effects than clinicians (figure 3); percentage agreement was 48% and concordance was poor as evidenced by the low weighted kappa (0.17, table 2). Bowker's test of symmetry was also highly significant ($p < 0.001$) indicating discordance, with patients reporting more effects than clinicians (table 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]. Agreement was poor (35%), as was concordance (weighted kappa 0.09) [table 2].

In contrast, for breast induration at year-5, PROMs and CRO appeared better aligned with similar levels of effects reported by both (figure 3) and a higher % agreement (61%, table 2), but concordance remained poor (weighted kappa 0.12, table 2). In addition, Bowker's test for symmetry was no longer significant ($p = 0.025$), implying similar effects reported by PROMs and CRO (table 2).

Comparison of radiotherapy schedules using PROMs, 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 4). There was some evidence of differing effect sizes between the assessment methods for breast induration, but the confidence intervals overlapped (figure 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 (appendix table 1 & 2). Some baseline factors were significantly associated with concordance of PROMs and either CRO or photographs for certain NTE in logistic regression models, but predominantly on univariate analysis only and not across both time-points (appendix table 3). 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)] (appendix table 3).

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 more NTE compared with clinicians and photographs. Concordance was poor between PROMs and either CRO or photographs on an individual patient level. However, results obtained for randomised comparisons between treatment groups were consistent for PROMs 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 (1, 2, 14). It is therefore increasingly important, in an era of improving radiotherapy techniques to monitor NTE using sufficiently sensitive methods. Within IMPORT LOW, patients reported more NTE compared with clinicians or photographs; this has been previously documented in the literature (15-22). This suggests NTE may be underestimated if only clinician-reported or photographic outcomes are used. In contrast, the Cambridge IMRT trial (23) 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 also 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 (15-17, 23-26). 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 (27).

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. These findings are consistent with those from the START trials (15). Within IMPORT LOW there also appeared to be a higher sensitivity of PROMs to treatment volume, although the effect sizes obtained from PROMs remained consistent with CRO and photographs. It should be noted that the PROMs investigated in this analysis and the START trials were the protocol-specific items, which

were specifically developed to capture late radiotherapy effects (15), rather than generic PROMs related to general quality of life (5).

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 (15) and Cambridge IMRT (23) 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 (28).

Obtaining photographs is also becoming increasingly challenging. Firstly, despite consenting to participate in a photographic sub-study, patients may not attend for photographs. 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 (1-3, 23) 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 (9). However, in IMPORT LOW, there was a large discrepancy in rating overall change in breast appearance between photographs and PROMs (% agreement = 26% and 35% at year-2 and 5 respectively). Patients reported significantly more NTE at both time-points, suggesting photographs may not capture the changes which are important for patients.

PROMs provide an opportunity to understand the patients' own perception of NTE within the framework of questions asked. We know that patients report more NTE than clinicians (15-22) or photographs and therefore, without the use of PROMs, the prevalence of NTE may be underestimated. Furthermore, PROMs are able to distinguish between treatment groups (1, 2). Within the START trials, all three data ascertainment methods were able to differentiate between randomised treatment groups (29, 30) whereas in IMPORT LOW it was found that only PROMs were able to distinguish between randomised comparisons (1). 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), PROMs may have better capability in differentiating between treatment groups.

However, there are a number of issues related to PROMs. Firstly, certain patient groups may not wish to participate in a PROMs study resulting in a trial population unrepresentative of the general population. Secondly, obtaining complete datasets can be challenging (4) 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. In IMPORT LOW, women who declined participation in the PROMs sub-study were slightly older than those who did consent (13). 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 (13). Also, patients who reported more adverse effects at year-2 were more likely to return questionnaires at year-5 (13). 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, including collecting data electronically, such as via smart phone/ email. Reducing numbers of questions in PROM 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 PROMs, improvements in the standardisation of analysis, interpretation and reporting of PROMs data in clinical trials are also required to enable cross-comparison of data between trials (31).

We have discussed whether PROMs could potentially replace either CRO or photographs 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. We acknowledge CRO are still widely supported and an alternative viewpoint is that both PROMs and CROs may be necessary as they measure differing aspects of disease experience and are complementary (32).

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 PROMs and either CRO or photographs.

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

Figure Captions

Figure 1: Summary of whole trial population consenting to PROMs and photographs, and data available at 2 and 5 years

Figure 2: Comparison of year-2 patient reported outcome measures, clinician and photographic assessments of specific late NTE in IMPORT LOW

Figure 3: Comparison of year-5 patient reported outcome measures, clinician and photographic assessments of specific late NTE in IMPORT LOW

Figure 4: Comparison of the estimates of effect sizes for the randomised radiotherapy groups between PROMs and CRO/photographs at 5 years

Conflicts of Interest

JMB discloses Research Funding: AstraZeneca, Merck Sharp & Dohme, Medivation, Puma Biotechnology, Clovis Oncology, Pfizer, Janssen-Cilag, Novartis, Roche. All other authors have no conflicts of interest.

Acknowledgements

We thank the patients who participated in this study and all investigators and research support staff, past and present, at participating centres. We also thank all trials unit staff at The Institute of Cancer Research Clinical Trials and Statistics Unit (ICR-CTSU) who contributed to the central coordination of the study including Jenny Titley, Ronald Kaggwa and Marie Emson.

We acknowledge support from Cancer Research UK (grant number C1491/A6035), the National Institute for Health Research (NIHR) Cancer Research Network (CRN), National Health Service Research Scotland, Health and Care Research Wales and the National Institute of Health Research Royal Marsden/ Institute of Cancer Research Biomedical Research Centre (Dr Anna Kirby). Dr Indrani Bhattacharya is supported by a Cancer Research UK clinical trial fellowship (C1491/A8895). In addition Dr Charlotte Coles is supported by the Cambridge National Institute of Health Research Biomedical Research Centre.

The IMPORT Trialists' Group consists of the Trial Management Group, Trial Steering Committee, Independent Data Monitoring Committee and the principal and main co-investigators at the participating centres.

IMPORT LOW is sponsored by The Institute of Cancer Research, London.

References

1. Coles CE, Griffin CL, Kirby AM, Titley J, Agrawal RK, Alhasso A, 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* (London, England). 2017;390(10099):1048-60.
2. Haviland JS, Owen JR, Dewar JA, Agrawal RK, Barrett J, Barrett-Lee PJ, 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. *The Lancet Oncology*. 2013;14(11):1086-94.
3. Mukesh MB, Barnett GC, Wilkinson JS, Moody AM, Wilson C, Dorling L, et al. Randomized controlled trial of intensity-modulated radiotherapy for early breast cancer: 5-year results confirm superior overall cosmesis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2013;31(36):4488-95.
4. Faithfull S, Lemanska A, Chen T. Patient-reported Outcome Measures in Radiotherapy: Clinical Advances and Research Opportunities in Measurement for Survivorship. *Clinical oncology (Royal College of Radiologists (Great Britain))*. 2015;27(11):679-85.
5. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, 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. *Journal of the National Cancer Institute*. 1993;85(5):365-76.
6. Sprangers MA, Groenvold M, Arraras JI, Franklin J, te Velde A, Muller M, 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. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 1996;14(10):2756-68.
7. Hopwood P, Fletcher I, Lee A, Al Ghazal S. A body image scale for use with cancer patients. *European journal of cancer (Oxford, England : 1990)*. 2001;37(2):189-97.
8. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta psychiatrica Scandinavica*. 1983;67(6):361-70.
9. Haviland JS, Ashton A, Broad B, Gothard L, Owen JR, Tait D, et al. Evaluation of a method for grading late photographic change in breast appearance after radiotherapy for early breast cancer. *Clinical oncology (Royal College of Radiologists (Great Britain))*. 2008;20(7):497-501.
10. Yarnold J, Ashton A, Bliss J, Homewood J, Harper C, Hanson 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.

Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2005;75(1):9-17.

11. Bowker AH. A test for symmetry in contingency tables. *Journal of the American Statistical Association*. 1948;43(244):572-4.
12. Vitug AF, Newman LA. Complications in breast surgery. *The Surgical clinics of North America*. 2007;87(2):431-51, x.
13. Bhattacharya IS, Haviland JS, Kirby AM, Kirwan CC, Hopwood P, Yarnold JR, 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. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2018;Jco1800982.
14. Bhattacharya IS HJ, Kirby AM, Kirwan CC, Hopwood P, Yarnold JR, Bliss JM, Coles CE. Patient reported outcomes over 5 years following whole or partial breast radiotherapy: longitudinal analysis of the IMPORT LOW phase III randomised controlled trial. *Journal of Clinical Oncology*. Accepted September 2018. In Press. .
15. Haviland JS, Hopwood P, Mills J, Sydenham M, Bliss JM, Yarnold JR. 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. *Clinical oncology (Royal College of Radiologists (Great Britain))*. 2016;28(6):345-53.
16. Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2004;22(17):3485-90.
17. Velikova G, Wright P, Smith AB, Stark D, Perren T, Brown J, et al. Self-reported quality of life of individual cancer patients: concordance of results with disease course and medical records. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2001;19(7):2064-73.
18. Basch E, Jia X, Heller G, Barz A, Sit L, Fruscione M, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *Journal of the National Cancer Institute*. 2009;101(23):1624-32.
19. Bruner DW, Movsas B, Basch E. Capturing the patient perspective: patient-reported outcomes as clinical trial endpoints. *American Society of Clinical Oncology educational book American Society of Clinical Oncology Meeting*. 2012:139-44.
20. Groenvold M, Fayers PM, Petersen MA, Sprangers MA, Aaronson NK, Mouridsen HT. Breast cancer patients on adjuvant chemotherapy report a wide range of problems not identified by health-care staff. *Breast cancer research and treatment*. 2007;103(2):185-95.
21. Quinten C, Maringwa J, Gotay CC, Martinelli F, Coens C, Reeve BB, et al. Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *Journal of the National Cancer Institute*. 2011;103(24):1851-8.
22. Stephens RJ, Hopwood P, Girling DJ, Machin D. Randomized trials with quality of life endpoints: are doctors' ratings of patients' physical symptoms interchangeable with patients' self-ratings? *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 1997;6(3):225-36.
23. Mukesh MB, Qian W, Wah Hak CC, Wilkinson JS, Barnett GC, Moody AM, et al. The Cambridge Breast Intensity-modulated Radiotherapy Trial: Comparison of

- Clinician- versus Patient-reported Outcomes. *Clinical oncology (Royal College of Radiologists (Great Britain))*. 2016;28(6):354-64.
24. Davidson SE, Trotti A, Ataman OU, Seong J, Lau FN, da Motta NW, et al. Improving the capture of adverse event data in clinical trials: the role of the International Atomic Energy Agency. *International journal of radiation oncology, biology, physics*. 2007;69(4):1218-21.
25. Vistad I, Cvancarova M, Fossa SD, Kristensen GB. Postradiotherapy morbidity in long-term survivors after locally advanced cervical cancer: how well do physicians' assessments agree with those of their patients? *International journal of radiation oncology, biology, physics*. 2008;71(5):1335-42.
26. Brouwers P, van Loon J, Houben RMA, Paulissen J, Engelen SME, Heuts M, 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. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*. 2018;126(1):163-9.
27. Kirchheiner K, Nout R, Lindegaard J, Petric P, Limbergen EV, Jurgeniemi-Schulz IM, et al. Do clinicians and patients agree regarding symptoms? A comparison after definitive radiochemotherapy in 223 uterine cervical cancer patients. *Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft [et al]*. 2012;188(10):933-9.
28. Bhattacharya IS, Morden JP, Griffin C, Snowdon C, Brannan R, Bliss JM, et al. The Application and Feasibility of Using Routine Data Sources for Long-term Cancer Clinical Trial Follow-up. *Clinical oncology (Royal College of Radiologists (Great Britain))*. 2017;29(12):796-8.
29. Bentzen SM, Agrawal RK, Aird EG, Barrett JM, Barrett-Lee PJ, Bliss JM, et al. The UK Standardisation of Breast Radiotherapy (START) Trial A of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. *The Lancet Oncology*. 2008;9(4):331-41.
30. Bentzen SM, Agrawal RK, Aird EG, Barrett JM, Barrett-Lee PJ, Bentzen SM, et al. The UK Standardisation of Breast Radiotherapy (START) Trial B of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. *Lancet (London, England)*. 2008;371(9618):1098-107.
31. Pe M, Dorme L, Coens C, Basch E, Calvert M, Campbell A, et al. Statistical analysis of patient-reported outcome data in randomised controlled trials of locally advanced and metastatic breast cancer: a systematic review. *The Lancet Oncology*. 2018;19(9):e459-e69.
32. Atherton PJ, Watkins-Bruner DW, Gotay C, Moinpour CM, Satele DV, Winter KA, et al. The Complementary Nature of Patient-Reported Outcomes and Adverse Event Reporting in Cooperative Group Oncology Clinical Trials: A Pooled Analysis (NCCTG N0591). *Journal of pain and symptom management*. 2015;50(4):470-9.e9.