Practical Radiation Oncology® (2022) 12, e317–e328

practical radiation oncology
pro
www.practicalradonc.org

Physics Contribution

Evaluation of Dose Accuracy in the Near-Surface Region for Whole Breast Irradiation Techniques in a Multi-institutional Consortium

Alexander Moncion, PhD,[a,*] Melissa Wilson, MS,[b] Ruimei Ma, PhD,[c] Robin Marsh, CMD,[a] Jay Burmeister, PhD,[d] Daniel Dryden, MS,[e] Danielle Lack, MS,[f] Margaret Grubb, MS,[a] Alan Mayville, MS,[g] Paul Jursinic, PhD,[h] Kathryn Dess, MS,[d] Justin Kamp, MS,[c] Kellen Young, CMD,[e] Joshua T. Dilworth, MD, PhD,[f] Larry Kestin, MD,[b] Reshma Jagsi, MD, DPhil,[a] Melissa Mietzel, MS,[a] Frank Vicini, MD,[b] Lori J. Pierce, MD,[a] and Jean M. Moran, PhD[a] on behalf of the Michigan Radiation Oncology Quality Consortium

[a]University of Michigan, Department of Radiation Oncology, Ann Arbor, Michigan; [b]Michigan Healthcare Professionals Radiation Oncology Institute/GenesisCare USA, Farmington Hills, Michigan; [c]Sparrow Health Systems, Department of Cancer Care and Oncology, Lansing, Michigan; [d]Wayne State University, Karmanos Cancer Center, Detroit, Michigan; [e]Covenant HealthCare, Covenant Radiation Center, Saginaw, Michigan; [f]Beaumont Health, Beaumont Cancer Center, Royal Oak, Michigan; [g]Mercy Health, Lacks Cancer Center, Grand Rapids, Michigan; and [h]West Michigan Cancer Center, Kalamazoo, Michigan

Received 3 August 2021; accepted 10 January 2022

Abstract

Purpose: To assess the accuracy of dose calculations in the near-surface region for different treatment planning systems (TPSs), treatment techniques, and energies to improve clinical decisions for patients receiving whole breast irradiation (WBI).

Methods and Materials: A portable custom breast phantom was designed for dose measurements in the near-surface regions. Treatment plans of varying complexities were created at 8 institutions using 4 different TPSs on an anonymized patient data set (50 Gy in 25 fractions) and peer reviewed by participants. The plans were recalculated on the phantom data set. The phantom was aligned with

Sources of support: This work is financially supported by Blue Cross Blue Shield of Michigan and the Blue Care Network of Michigan as part of the BCBSM Value Partnerships Program.

All data generated and analyzed during this study are included in this published article. Raw data are readily available upon requests made to the corresponding author.

Disclosure: R. Marsh, M. Grubb, M. Mietzel, Dr. Jagsi, Dr. Pierce, and Dr. Moran report salary support from Blue Cross Blue Shield of Michigan, during the conduct of the study. Dr Jagsi reports other from Equity Quotient, personal fees and nonfinancial support from Amgen, personal fees from the Greenwall Foundation, personal fees from Dressman Benziger Lavalle Law, other from JAMA Oncology editorial board, grants and personal fees from the Komen Foundation, grants from the National Institutes of Health (NIH), grants and personal fees from the Doris Duke Foundation, personal fees from Sherinian and Hasso, and personal fees from Kleinbard LLC, outside the submitted work. Dr Pierce is cofounder of PFS Genomics, and has a patent METHOD FOR THE ANALYSIS OF RADIOSENSITIVITY pending. Dr Moran reports grants from Varian Medical Systems and the NIH, outside the submitted work. The remaining authors have nothing to disclose.

* Corresponding author: Alexander Moncion, PhD; E-mail: ambaez@med.umich.edu

https://doi.org/10.1016/j.prro.2022.01.013

predetermined shifts and laser marks or cone beam computed tomography, and the irradiation was performed using a variety of linear accelerators at the participating institutions. Dose was measured with radiochromic film placed at 0.5 and 1.0 cm depth and 3 locations per depth within the phantom. The film was scanned and analyzed >24 hours postirradiation.

Results: The percentage difference between the mean of the measured and calculated dose across the participating centers was -0.2 % ± 2.9%, with 95% of measurements within 6% agreement. No significant differences were found between the mean of the calculated and measured dose for all TPSs, treatment techniques, and energies at all depths and laterality investigated. Furthermore, no significant differences were observed between the mean of measured dose and the prescription dose of 2 Gy per fraction.

Conclusion: These results demonstrate that dose calculations for clinically relevant WBI plans are accurate to within 6% of measurements in the near-surface region for various complexities, TPSs, linear accelerators, and beam energies. This work lays the necessary foundation for future studies investigating the correlation between near-surface dose and acute skin toxicities.

Introduction

Radiation-induced skin toxicities (RISTs) are a common side effect of breast radiation therapy, and the severity has been shown to correlate with higher treatment dose.[1-4] These side effects can occur during or after radiation treatment and ultimately reduce the patient's quality of life.[5] RISTs manifest in the form of erythema, moist/dry desquamation, and pain, among other toxicities. The severity of the reaction is commonly graded using 2 scales: the Radiation Therapy Oncology Group/European Organization for Research and Treatment of Cancer (RTOG/EORTC) and the National Cancer Institute Common Toxicity Criteria for Adverse Events (NCI CTCAE v5.0).[1,6,7] Risk factors that predispose patients to the development of RISTs are generally in 2 categories: patient-related risk factors and treatment-related risk factors.[1] Patient-related risk factors include breast size, smoking, body mass index (BMI), comorbidities, and genetic predispositions,[8-15] and treatment-related risk factors are more technical and include treatment dose, fractionation scheme, treatment technique used, beam quality, energy, and systemic therapy use.[16-19]

Technical treatment-related risk factors may be easily adjusted to reduce a patient's risk of RISTs. Thus, research studies have investigated the correlation between treatment-related risk factors and RISTs, and how adjusting them affects treatment outcomes and quality of life for patients. Hypofractionated treatments (eg, 40.0 Gy in 15 fractions or 42.6 Gy in 16 fractions) have become a widely used treatment option for adjuvant WBI.[20] Randomized trials and prospective studies have shown that hypofractionated treatment regimens can provide comparable disease control relative to conventional fractionation and reduce the rate of RISTs.[19,21,22] Additionally, several studies have demonstrated that intensity modulated radiation therapy (IMRT), whether forward or inverse planned, reduces the occurrence of RISTs compared with 2-dimensional (2D) or 3-dimensional (3D) conformal radiation therapy (CRT), including wedge techniques.[9,23-25] However, these studies are unable to correlate their findings to surface or near-surface dose, and instead rely on the prescription dose. Unfortunately, there is a knowledge gap in this space arising from the lack of confidence in the treatment planning system (TPS) to accurately calculate dose in the surface and near-surface region.

Surface dose measurements require specialized dosimetric methods and may not be routinely addressed in commissioning of a TPS. Clinical trials such as RTOG 1004 address the lack of confidence in surface dose calculations through the creation of an evaluation structure for dose assessment. Other investigators have reported on discrepancies between dose calculations and measurements at depths of 2 to 3 mm.[26,27] The Michigan Radiation Oncology Quality Consortium (MROQC) is a multicenter collaborative quality initiative. A preliminary retrospective analysis within MROQC showed interesting correlations between acute skin toxicities and the mean dose in the near-surface region (0.5-1.0 cm) for different plan complexities.

In this work, we present the results from breast plans created for a common patient geometry and then delivered to a custom breast phantom at 8 clinics within the state of Michigan with a variety of TPSs. The goal of this study was to assess the agreement between the calculated dose from the TPS and the measured dose from the phantom across all the participating institutions for different TPSs, treatment techniques, energy, and measurement location laterality at the near-surface depths of 0.5 cm and 1.0 cm. This work was motivated by a panel discussion among physicists and physicians at one of our consortium meetings aiming to understand the uncertainty in the accuracy of the dose at depth. We seek to better understand the accuracy of TPSs in the near-surface region before an in-depth study looking at the relationship between near-surface dose and skin toxicity.

Materials and Methods

Custom breast phantom setup and dose measurements

In this work, we sought to use the same dosimetry system for measurements while minimizing the equipment

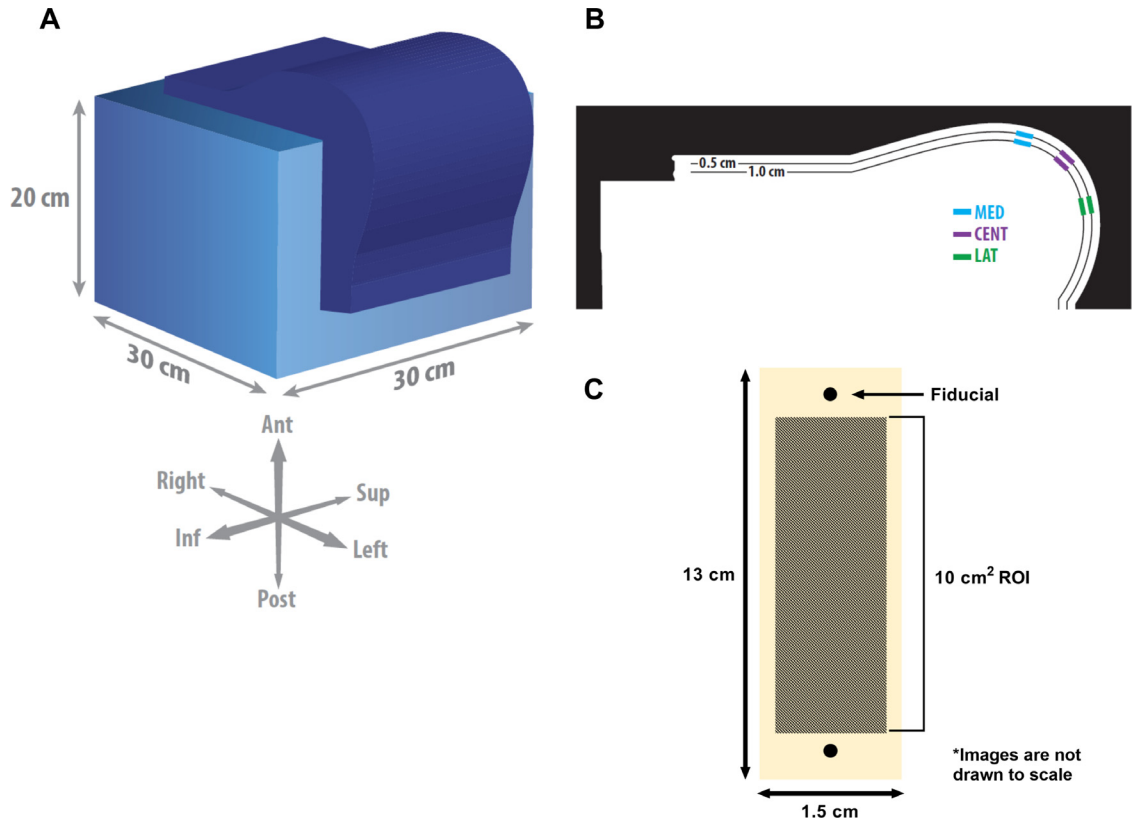

Figure 1 (A) The custom breast bolus was designed to rest on top of each institution's solid water phantom of dimensions 20 × 30 × 30 cm^3. (B) An axial slice of the phantom shows the 3 separate layers used to simulate depths of 0.5 and 1.0 cm. We were also able to simulate laterality within each depth: medial (MED), central (CENT), and lateral (LAT). The anonymized patient's separation was 20.2 cm and the phantom's was 19.1 cm (defined at the patient's origin). (C) Each film was labeled with 2 fiducials used for registration to the calculated dose. Post registration, a 10.0 × 1.0 cm^2 region of interest (ROI) was drawn on the registered dose planes (ie, the calculated and measured) to exclude artifacts from the film edges caused by cutting and preparation. *Abbreviations:* Ant = Anterior; Inf = Inferior; Pos = Posterior; Sup = Superior.

needed to be transferred to each institution. We designed an overlay for a solid water phantom, which was then created of a proprietary polyethylene wax (.decimal, Sanford, FL) to represent breast anatomy of an anonymized patient data set, which was used for initial plan creation, from the tangential beam orientation (Fig 1A). The phantom consisted of 3 layers (Fig 1B): the base layer that mounts onto each participating institution's solid water, and the middle and top layers are each 0.5 cm thick, allowing measurements at 0.5 cm and 1.0 cm depth underneath each layer, respectively. The phantom was scanned at the coordinating center using the coordinating center's solid water, and the computed tomography (CT) data were sent to the different sites for teams to create plans consistent with their practice and specific to their planning and delivery systems. Based on the CT data, the only air gaps identified were small and owing to the presence of the film in place. Measurements were made using Gafchromic EBT3 film (Ashland Advanced Materials, Bridgewater, NJ).[28] The film was cut to dimensions of 13.0 cm × 1.5 cm, and then 0.05-cm holes were punched on the short edges of the film, as shown in Figure 1C, to serve as fiducials. The dimensions of the film were chosen to fit in narrow flat regions of the phantom at depths of 0.5 cm and 1.0 cm, thus eliminating film curvature while maximizing the surface area of the measured dose plane. The film was placed in 3 different locations (medially [MED], central [CENT], and laterally [LAT]) and at 2 depths (0.5 and 1.0 cm) within the custom breast phantom (see Fig 1B). Thus, there were a total of 6 measurements made per plan. All film was prepared at the coordinating center and transported with the phantom to the different clinics for measurements. Two separate film calibration strips were included with each transport. The first calibration strip, with instructions to not irradiate, accounted for any changes due to temperature and handling of the film and also served as the unexposed film of the published single scan protocol.[29-31] The second calibration strip was exposed to a known dose (2 Gy) at the participating center and used as the calibration strip for a single scan protocol. Initial localization of the breast phantom was performed by aligning the in-room lasers to markings on

the phantom. The final alignment was then performed with either a cone beam computed tomography (CBCT) or predetermined shifts with laser marks, based on the preference of the local institution. After alignment was complete, the plans created by each participating institution were delivered to the phantom. A new set of unexposed film was used for each plan.

Treatment planning

From within the consortium, 8 centers with different planning and delivery systems participated in this study. Each system (TPS and/or delivery system) was represented by at least 2 institutions, with the exception of TomoTherapy. Treatment plans were generated with the TPSs for the techniques and energies defined in Table 1. The energy combinations depended on the institution's standard practice.

All plans were generated at the participating institution using their respective TPS and the same anonymized patient data set to mimic planning complexities introduced from the presence of tissue heterogeneities and organs-at-risk. The plans represent the range of treatment techniques at that institution. This fits into a broader assessment of the impact of treatment technique on

Table 1 Details of the TPSs and algorithms, treatment machines, plan complexities, and calculation grid size used for treatment planning and delivery at the different participating institutions

Institution	TPS (algorithm)	Treatment machine	Plan complexities	Calculation grid size (mm)
A	Eclipse 15.6 (AAA 15.6)	Varian	Open (1 plan: 6X) FiF (1 plan: 6X) FiF mixed energy (1 plan: 6X/16X mixed) IMRT (1 plan: 6X)	2.5
B	Pinnacle 9.1 (Adaptive Convolution Superposition)	Elekta	Open (1 plan: 6X) FiF (1 plan: 6X) FiF mixed energy (1 plan: 6X/16X mixed) IMRT (1 plan: 6X)	2.5
	Raystation 8.1 (Collapsed Cone v 5.0)	Elekta	Open (1 plan: 6X) FiF (1 plan: 6X) FiF mixed energy (1 plan: 6X/16X mixed) IMRT (1 plan: 6X)	2.0
C	Pinnacle 16.2 (Adaptive Convolution Superposition)	Elekta	Open (2 plans: 6X, 18X) IMRT (2 plans: 6X, 18X)	2.5
D	Eclipse 15.6 (AAA 15.6)	Varian	Open (1 plan: 6X) FiF (1 plan: 6X) FiF mixed energy (1 plan: 6X/15X mixed) IMRT (1 plan: 6X)	2.5
E	Raystation 8.0 (Collapsed Cone v 4.1)	Elekta	Open (1 plan: 6X) FiF (1 plan: 6X) FiF mixed energy (1 plan: 6X/18X mixed) IMRT (1 plan: 6X)	2.0
F	Eclipse 15.6 (AAA 15.6)	Varian	Open (1 plan: 6X) FiF (2 plans: 6X, 10X) IMRT (2 plans: 6X, 10X)	2.0
G	Tomotherapy 6.0 (Collapsed Cone v5.0)	Tomotherapy	Open (1 plan: 6X) FiF (1 plan: 6X) IMRT (2 plans: direct, helical)	1.0
H	Eclipse 15.6 (AAA 15.6)	Varian	Open (1 plan: 6X) FiF (1 plan: 6X) FiF mixed energy (1 plan: 6X/16X mixed) IMRT (1 plan: 6X)	2.5

Abbreviations: AAA = analytical anisotropic algorithm; FiF = field-in-field; IMRT = intensity modulated radiation therapy; TPS = treatment planning system.
The number of plans and energy used in the plan are shown in parentheses next to each plan complexity. All sites delivered an open field and an IMRT plan. The low energy was considered 6X for all sites, and all energies other than 6X were considered high energy.

patient toxicity.[32] The quality of the plan was verified by a dosimetrist, physician, and physicist from each center to confirm that coverage goals were similar and reasonable for the plan types. The plan goals were for 95% of the breast target volume to receive ≥95% of the prescribed dose, mean dose to the heart ≤1.3 Gy, and ≤33% of the ipsilateral lung to receive no more than 20 Gy. All of the plans had a prescription of 50 Gy in 25 fractions, and the maximum dose was kept <107% of the prescription. For each TPS, a set of plans that represented the beam energies and continuum of treatment techniques was created. The range of treatment complexity included open fields, field-in-field (FiF), FiF with mixed energies, and IMRT. All institutions were asked to create a simple plan with open tangents and no wedges for 6 MV. This plan was intended to assess a baseline performance without any beam modifiers or modulation. For all other plans, institutions were asked to use their preferred energy, which included 6 MV, mixed (eg, 6 MV and 10 or 16 MV), or high (10 MV or 16 MV) energies. For non-IMRT, flash was achieved by opening the jaws about 20 mm beyond the external body contour. The amount of flash for IMRT was left to each institution's preference and expertise, and no specifications on planning objectives were given beyond the previously mentioned target and OAR goals. These instructions were provided because each center had patient data submitted to the consortium based on its typical techniques. Centers were asked to create plans across the spectrum of techniques even if they preferentially used a particular technique for the majority of their patients. For TomoTherapy, the open and FiF plans were delivered using TomoDirect with pre-established static gantry angles. The IMRT plan was delivered with arc-based TomoHelical technique with directional blocking. Other IMRT plans were inverse-planned with a static gantry sliding window technique.

After plan creation, a summary of the plan with screenshots in the planes with maximum doses was shared with all participants in the project. After review, personnel at each institution then transferred each plan to the registered CT data set of the custom breast phantom shown in Figure 1A. The dose was then recalculated on the custom phantom using the institution's respective TPS with the monitor units from the patient treatment plan. This is consistent with how pretreatment IMRT QA measurements are done at the participating centers.

Extraction of dose calculations

The DICOM-RT data associated with all the treatment plans on the custom phantom data set were exported from each center's TPS and imported into Eclipse 15.6 (Varian) at the coordinating center. For each unique plan, dose planes in the TPS, which overlapped with the physical location of the film in the phantom, were exported as DICOM files from the TPS using a plane with identical physical dimensions to that of the film (ie, 13 cm × 1.5 cm or 394 × 45 pixels) with a resolution of 3 pixels/mm.

Film analysis

All film was analyzed at the coordinating center. The film was scanned (Epson Expression 10000XL) with 48-bit color at a 72 dpi resolution (3 pixels/mm) and without any color corrections. The film was registered to the exported TPS calculated dose plane using the fiducials in the film (see Fig 1C). A calibration curve from 0 to 5.0 Gy was generated by exposing a set of film strips from the same lot as the experimental film to a known dose corresponding to points along the calibration curve. The dose map on the film was generated using a triple channel uniformity optimization. The registration, calibration curve, and dose map were all created using FilmQA Pro 2016 (version 5.0.6161.42070, Ashland). After registration and generation of the film's dose map, a 10.0 cm × 1.0 cm region-of-interest (ROI) was drawn on the registered measured and calculated dose planes, and the dose within the ROI was exported to Matlab for batch analysis (MathWorks, Natick, MA).

Statistics

All statistical analyses were performed using GraphPad Prism (GraphPad Software, Inc, La Jolla, CA). All data are expressed as the mean ± standard deviation of the measured quantities unless stated as otherwise. All n values are listed with each corresponding figure. Statistically significant differences of all data sets were determined with a Student t test corrected for multiple comparisons using the Holm-Sidak method, with differences deemed significant for $P < .05$.

Results

Figure 2 shows the correlation between the mean of the calculated dose plane and the mean of the measured dose plane for all treatment complexities and measurement locations within the custom breast phantom. The data in all the figures is representative of all participating centers. There were a total of 60, 54, 36, and 72 individual measurements performed across all the participating centers for the open field, FiF, FiF mixed energy, and IMRT plans, respectively. The number of plans per participating center is shown in Table 1. These measurements account for all measurement locations and depths in the custom breast phantom. Overall, there is a positive correlation between the calculated and measured mean dose, with Pearson correlation coefficients of 0.8, 0.9, 0.7, and 0.8 for the

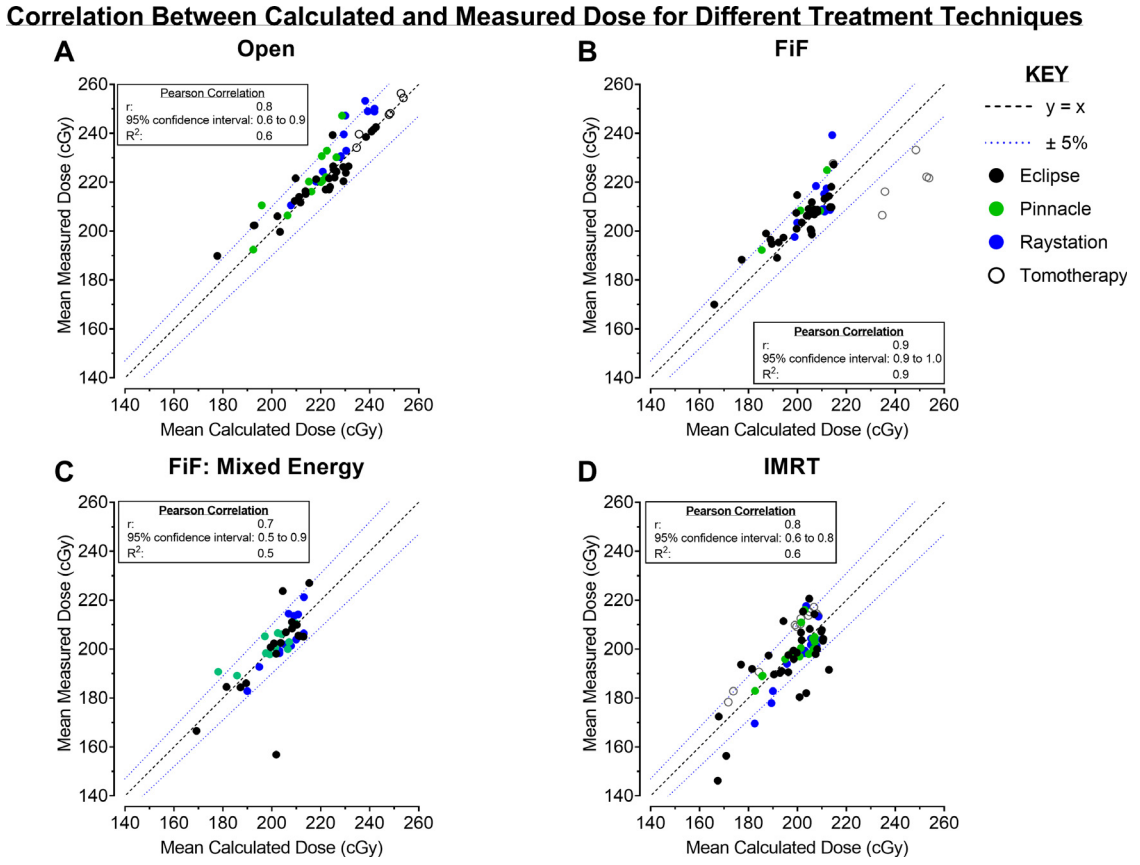

Figure 2 The correlation between the mean of the measured and calculated dose is shown for (A) open fields, (B) field-in-field (FiF), (C) FiF mixed energy, and (D) intensity modulated radiation therapy. Each planning complexity contains measurements from Eclipse, Pinnacle, Raystation, and TomoTherapy treatment planning systems, where applicable. A total of 60, 54, 36, and 72 individual measurements was performed across all the participating centers for the open field, FiF, FiF mixed energy, and intensity modulated radiation therapy plans, respectively. A $y = x$ and $\pm 5\%$ lines were overlaid with the data.

open field, FiF, FiF mixed energy, and IMRT, respectively. For all individual measurements, 86.7%, 77.8%, 91.9%, and 80.6% of the measured mean doses were within ±5% of calculated mean dose for open field, FiF, FiF mixed energy, and IMRT, respectively.

A Gaussian fit of a histogram of the percentage dose difference between the mean of the calculated and measured dose planes is shown in Figure 3 for different depths. The measurements performed at 0.5 cm depth trended toward a higher mean measured dose relative to the calculation. This resulted in the mean of the distribution being $-1.3 \pm 2.3\%$ across 111 measurements (see Fig 3A). The opposite trend was observed for the measurements performed at 1.0 cm depth with the mean of the distribution being $0.8 \pm 2.2\%$ across 111 measurements (see Fig 3B). The combined data (see Fig 3C) show that the mean percentage difference is $-0.2 \pm 2.9\%$; thus, 95% of all the measurements are within ±6% of zero dose difference regardless of measurement depth, location, treatment modality, and treatment complexity.

Figure 4 shows the data as a function of measurement location and depth on the phantom for all 4 treatment complexities. There was good agreement between the calculated and measured mean dose for all treatment complexities, measurement location, and depth. The largest difference between the measured and calculated means was 7.0 cGy at the 0.5 cm LAT measurement location for the open field. Similarly, the largest differences for the FiF, FiF mixed energy, and IMRT plans were 2.3, 6.9, and 4.6 cGy at the 0.5 cm LAT, 0.5 cm CENT, and 0.5 cm LAT measurement locations, respectively. The largest range for the mean of the calculated dose was observed for the FiF (0.5 cm LAT) treatment technique, with a minimum and maximum of 166.0 cGy and 234.6 cGy (range 68.6 cGy) for the calculated dose, and a minimum and maximum of 170.0 and 206.5 (range 36.5 cGy) for the corresponding measured dose. The largest range for the mean of the measured dose was observed for the IMRT (0.5 cm CENT) treatment technique, with a minimum and maximum of 146.2 cGy and 214.3 cGy

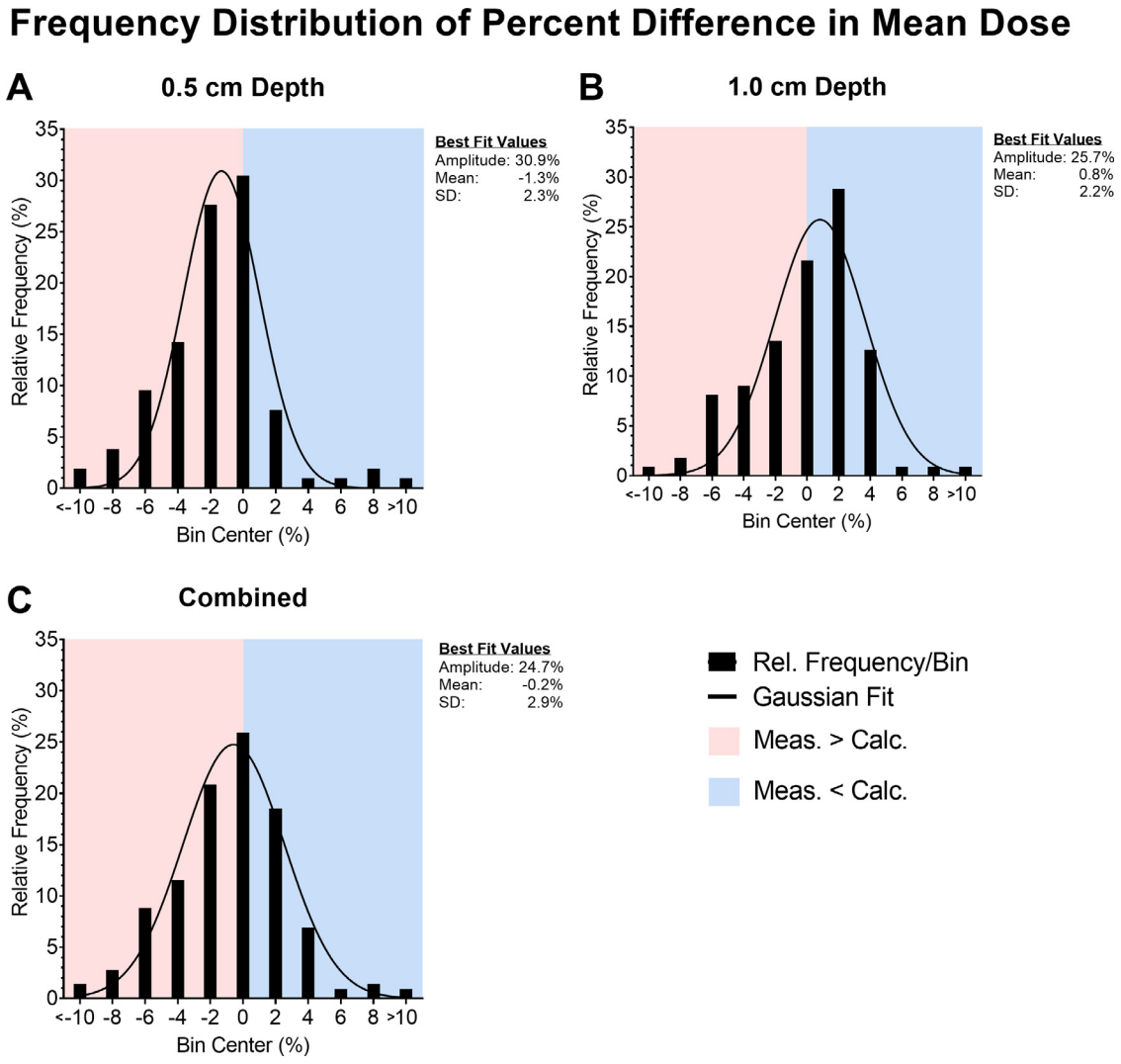

Figure 3 A Gaussian fit of a histogram of the percentage dose difference between the mean of the calculated (Calc.) and measured dose (Meas.) planes for measurements performed at (A) 0.5 cm and (B) 1.0 cm depths. The combined data are shown in (C), and all data have a bin width of 2%. Based on the Gaussian fit, the mean percentage difference was -1.3 ± 2.3%, 0.8 ± 2.2%, and -0.2 ± 2.9%, for measurements at 0.5 cm, 1.0 cm, and combined depth. Thus, 95% of all the measurements were within ±6% of agreement. SD = standard deviation.

(range 68.1 cGy) for the measured dose, and a minimum and maximum of 167.4 and 208.5 (range 41.1 cGy) for the corresponding calculated dose. There were no statistical differences observed for any comparisons across depth and measurement locations.

The mean of the dose plane for the calculation and measurement for different treatment techniques and treatment modalities can be seen in Figure 5. All the complexities, except for IMRT, had means of the distribution greater than the prescription of 200 cGy. The percentages of points greater than the prescription for the calculation (measurement) were 91.7 (95.0), 74.1 (79.6), 75.0 (69.4), and 55.6% (44.4%) for open, FiF, FiF mixed energy, and IMRT, respectively. Similarly, all of the TPSs, except for TomoHelical, had means of the distribution greater than the prescription of 200 cGy. The percentages of points greater than the prescription for the calculation (measurement) were 68.5 (66.7), 66.7 (60.4), 83.3 (77.1), 100 (100), and 16.7% (50.0%) for Eclipse, Pinnacle, RayStation, TomoDirect, and TomoHelical, respectively. The percentage difference between the mean of the measured and calculated dose for different TPSs, treatment techniques, and field energy are shown in Figure 6, and detailed statistics are in Table 2. The smallest mean percentage difference was observed for the Eclipse TPS (0.0 ± 4.7%), FiF (Single Energy) treatment technique (0.0 ± 4.6%), and high energy (0.3 ± 2.5%). A mean percentage dose difference that was significantly different from zero was not observed.

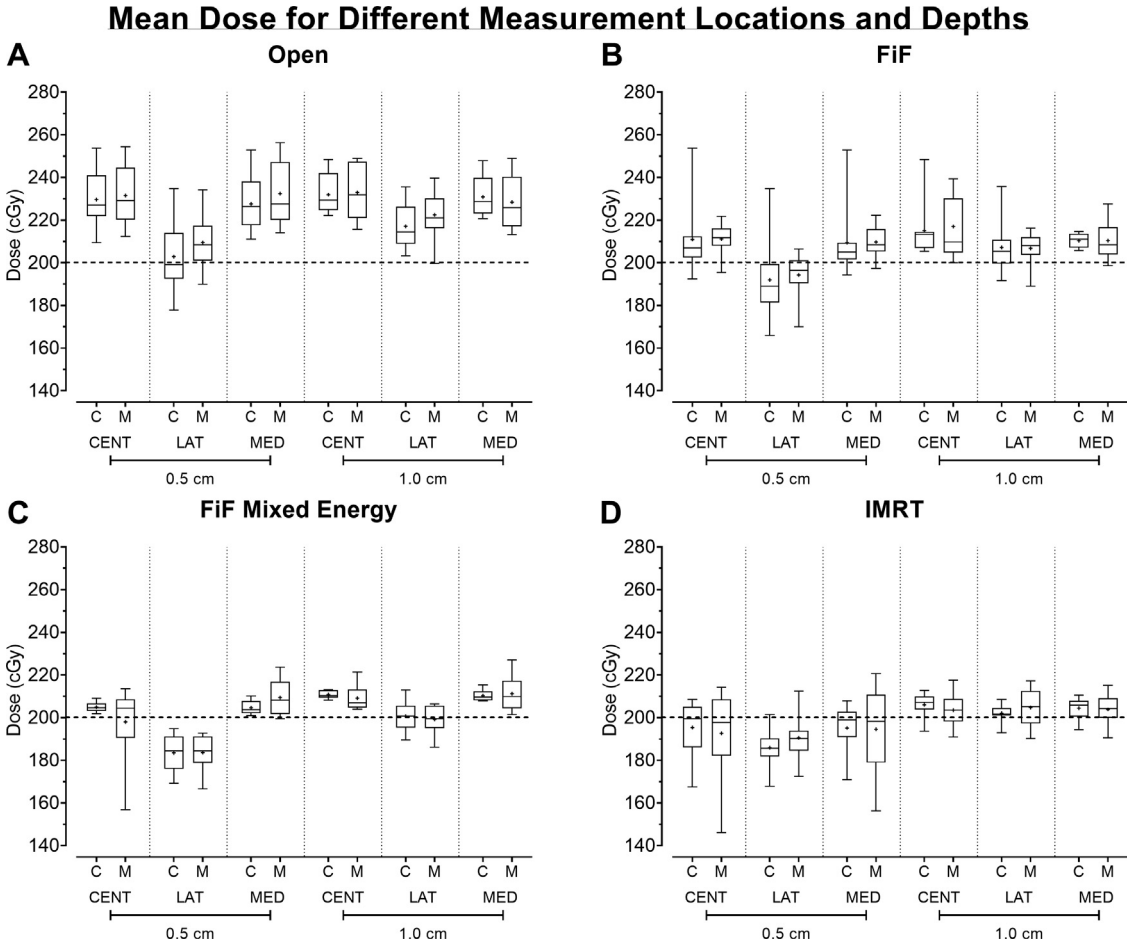

Figure 4 Mean of the dose planes for (A) open fields, (B) field-in-field (FiF), (C) FiF mixed energy, and (D) intensity modulated radiation therapy (IMRT) for the calculation (C) and measurement (M). The data are further divided by laterality (medial [MED], central [CENT], and lateral [LAT]) and depth. The range of the distribution is from minimum to maximum, and the mean of the distribution is denoted by "+." There was good agreement between the mean of the calculation and measurement for all treatment modalities, laterality, and depths. The dotted dashed line represents the prescription dose.

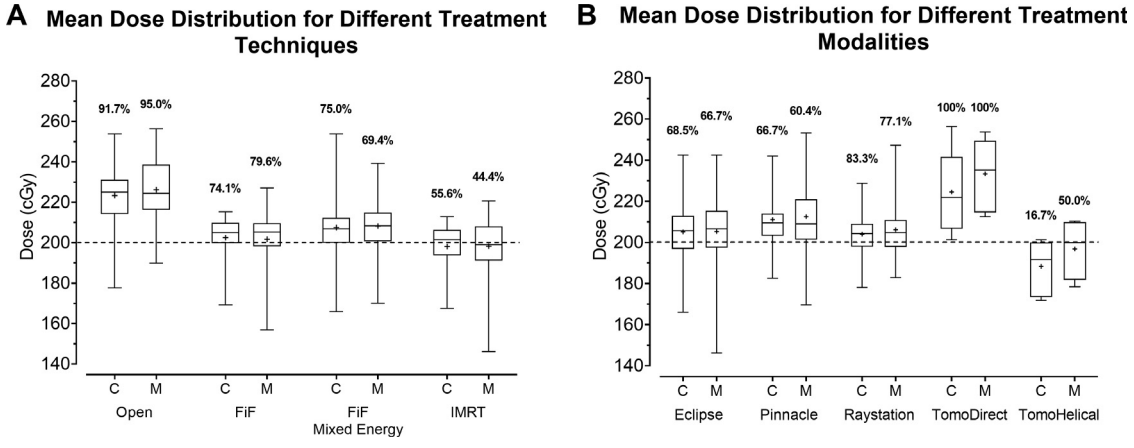

Figure 5 The mean of the dose planes for the calculation (C) and measurement (M) is shown for different (A) treatment techniques and (B) treatment modalities. The range of the distribution is from minimum to maximum, and the mean of the distribution is denoted by "+." The prescription dose is denoted by a horizontal dotted line at 200 cGy, and the percentage of points with mean values greater than the prescription dose is shown above each box.

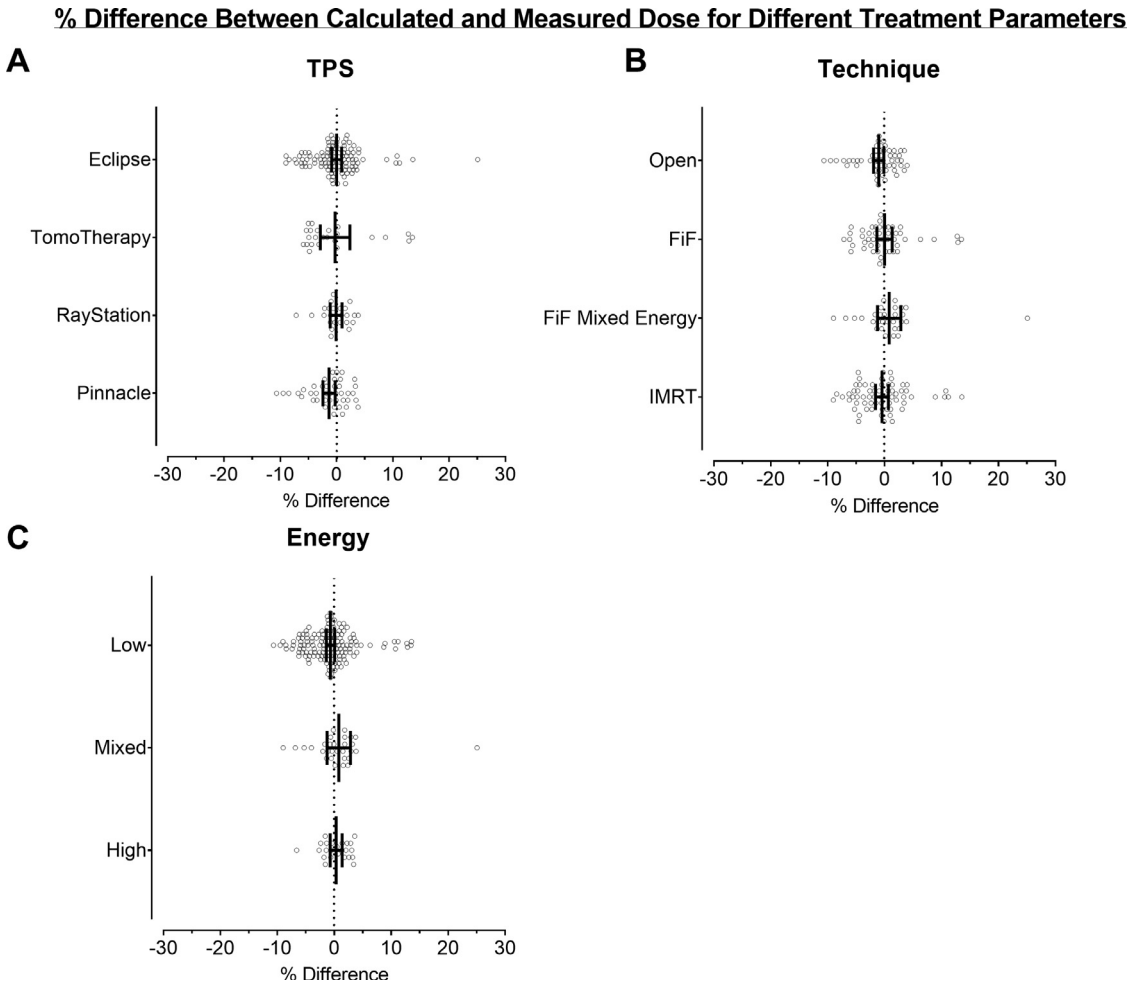

Figure 6 Percentage difference in mean dose between the calculation and measurement for different (A) treatment planning systems (TPSs), (B) treatment techniques, and (C) beam energy. Error bars are overlaid with the distribution and represent the mean with the 95% confidence interval of the distribution. *Abbreviations:* FiF = field-in-field; IMRT = intensity modulated radiation therapy.

Table 2 The percentage difference in mean dose between the calculation and measurement for different TPSs, treatment techniques, and beam energy

Characteristic		Mean % difference	Standard deviation	Lower 95% CI	Upper 95% CI
TPS	Eclipse	0.0	4.7	−0.9	0.9
	Pinnacle	−1.3	3.5	−2.4	−0.3
	Raystation	−0.1	2.5	−1.1	1.0
	Tomotherapy	−0.3	6.2	−2.9	2.4
Treatment technique	Open	−1.0	3.3	−1.9	−0.1
	FiF (single energy)	0.0	4.6	−1.4	1.3
	FiF (mixed energy)	0.8	5.5	−1.3	2.9
	IMRT	−0.5	4.6	−1.6	0.7
Energy	Low	−0.7	4.5	−1.4	0.1
	Mixed	0.8	5.5	−1.3	2.9
	High	0.3	2.5	−0.7	1.4

Abbreviations: FiF = field-in-field; IMRT = intensity modulated radiation therapy; TPS = treatment planning system.
Detailed statistics of Figure 6.

Discussion

This work addresses a gap in knowledge of dose accuracy in the near-surface dose region and may serve as a foundational piece for understanding the relationship between dose and skin-related toxicities in breast radiation therapy. We investigated the correlation between the calculated dose from different TPSs, treatment techniques, energy, and measurement laterality in the near-surface region. The results presented are important to support subsequent evaluations of TPS dose in the near-surface region and its potential correlation to patient toxicities. One of the prominent insights gained is that accurate near-surface doses can be calculated using the various TPSs, treatment complexities, beam energies, and different delivery systems that were used in this investigation.

A novel phantom was used in this multi-institutional study to evaluate the accuracy of the different TPSs in a breast-specific geometry. The approach used is somewhat similar to that by Low et al when adding a pelvis shell to a phantom designed to represent the head.[33] One of the advantages of our phantom is that the overlay can be added to each institution's solid water with a snug fit, which allowed all participants to use film from a common source and in the same measurement geometry.[34] Additionally, the custom phantom used in our work provided the necessary ease of transportation and setup at each site while meeting the geometric requirements for clinically relevant measurements. We were able to have film measurements done at each site with analysis at the MROQC coordinating center. Others have developed a series of 3D-printed custom breast phantoms with ionization chamber and radiochromic film inserts, which can be attached to a commercial anthropomorphic thorax phantom.[35] That phantom was developed specifically for performing end-to-end testing of a whole breast volumetric arc radiation therapy treatment technique.

This work demonstrates that several TPSs reliably calculate dose in the near-surface region to within ±6%. Although this is slightly higher than published recommendations of ±5%,[36] our results are reasonable when considering these measurements are conducted across several institutions, using several different TPSs, treatment planning techniques, and delivery systems, measured in the near-surface region with differences calculated across a 10-cm^2 area rather than a point measurement. Because patients are treated at institutions with different planning systems with customized beam fits, different treatment techniques, and on linear accelerators from different manufacturers, the final results of accuracy within ±6% provide a crucial foundation when further investigating the relationship of toxicity with dose as a function of treatment technique.

There was no statistical difference between the calculated and measured dose for the experimental conditions investigated; however, the mean of the calculated dose was higher than the prescription dose for all treatment techniques except IMRT (Fig 5A), and for all modalities except TomoHelical (Fig 5B). Research studies have shown that IMRT significantly improved dose distribution and target conformity compared with 3D-CRT, resulting in a lower rate of RISTs.[9,23-25] However, no differences were found in the local-regional tumor control between IMRT and 3D-CRT.[25] Thus, IMRT may give physicians the flexibility for skin sparing and a reduced rate of RISTs, but with the disadvantages of higher costs for patients, higher plan complexity, a need for more complex image guidance, and added quality assurance in the treatment workflow.[37-40]

Of all the conditions tested, TomoTherapy had the least amount of data owing to the limited use of the technique across the participating centers. A research study measured surface dose with a wide variety of detectors and demonstrated that Helical TomoTherapy TPSs overestimate the surface dose.[41] However, these detectors were limited by their water-equivalent depth, which was greater than the 70-μm depth needed to measure dose relevant to skin toxicities. An exception to this limitation is a noncommercial MOSFET-based detector with a water equivalent depth of 70 μm used in a separate study.[42] Results from a phase 3 randomized controlled trial provided evidence that Helical TomoTherapy could have a lower rate of RISTs compared with a FiF TomoTherapy, although the study was not powered to detect late toxicities from the Helical TomoTherapy arm. However, both techniques resulted in a 5-year survival rate >96%.

There are limitations to this work that stem from the nature of the phantom and our radiation measurement device. Disadvantages of the breast phantom are that it is stationary without the ability to simulate breathing motion. As such, our setup is more comparable to a breath-hold treatment, and our results may not be representative of the dosimetric accuracy in the near-surface region during a free breathing treatment. Additionally, owing to the use of each institution's solid water as the base for the breast phantom, there are slight differences in electron density that were not accounted for in the planning, which may have contributed to the observed differences in planned versus delivered dose. Another limitation is the phantom setup and, specifically, the use of predetermined shifts versus a CBCT-based alignment. Three of 8 participating centers used predetermined shifts and laser marks for their alignment. There were outliers observed (see Fig 6) for 1 institution's FiF Mixed Energy and IMRT plans, which were both delivered after predetermined shifts of the phantom. Although our setup is not complicated and our outliers not numerous, dosimetric studies such as the one presented in this work may benefit from a consistent and

robust setup alignment that CBCTs provided. Overall, the results presented are interesting because they demonstrate that the dose in the near-surface buildup region of most clinical beams can be accurately calculated.

Conclusion

To the best of our knowledge, this work represents the first multi-institutional comparison of the accuracy of near-surface doses delivered using whole breast plans designed from multiple TPSs and then measured in a clinically relevant phantom geometry for different delivery systems. This study confirms that patient doses calculated in the near-surface region by the respective TPSs in our multi-institutional consortium agree to within ±6% of measured doses from their respective delivery systems. Because of the demonstrated accuracy of the different systems, this work sets a foundation for future investigations to evaluate potential correlations between skin toxicities and near-surface dose. It may also inform future investigations evaluating skin toxicity in clinical trial settings.

Acknowledgments

Special thanks to Lisa Benedetti, Derek Bergsma, Charlie Bucki, Denise Crouch, Mike Dominello, Ahmad Hammoud, Nathan Jones, Wayne Kearsley, Kelley Lovall, Alan Mayville, Pat McDermott, Loree Nisley, Jeff Radawski, Danielle Redner, Ian Reineck, Kailee Walters, Mark Zaki for their participation in this project. MROQC is financially supported by Blue Cross Blue Shield of Michigan and the Blue Care Network of Michigan as part of the BCBSM Value Partnerships Program.

References

1. Kole AJ, Kole L, Moran MS. Acute radiation dermatitis in breast cancer patients: Challenges and solutions. *Breast Cancer (Dove Med Press)*. 2017;9:313-323.
2. Ryan JL. Ionizing radiation: The good, the bad, and the ugly. *J Invest Dermatol*. 2012;132(3):985-993. Pt 2.
3. Archambeau JO, Pezner R, Wasserman T. Pathophysiology of irradiated skin and breast. *Int J Radiat Oncol*. 1995;31:1171-1185.
4. Bray FN, Simmons BJ, Wolfson AH, Nouri K. Acute and chronic cutaneous reactions to ionizing radiation therapy. *Dermatol Ther*. 2016;6:185-206.
5. 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*. 2001;91:2273-2281.
6. 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*. 1995;31:1341-1346.
7. National Cancer Institute. Common terminology criteria for adverse events: (CTCAE). Available at: http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf. Accessed 28 January 2021.
8. Fernando IN, Ford HT, Powles TJ, et al. Factors affecting acute skin toxicity in patients having breast irradiation after conservative surgery: A prospective study of treatment practice at the Royal Marsden Hospital. *Clin Oncol*. 1996;8:226-233.
9. Pignol J-P, Olivotto I, Rakovitch E, et al. A multicenter randomized trial of breast intensity-modulated radiation therapy to reduce acute radiation dermatitis. *J Clin Oncol*. 2008;26:2085-2092.
10. Fisher J, Scott C, Stevens R, et al. Randomized phase III study comparing best supportive care to biafine as a prophylactic agent for radiation-induced skin toxicity for women undergoing breast irradiation: Radiation therapy oncology group (RTOG) 97-13. *Int J Radiat Oncol*. 2000;48:1307-1310.
11. Back M, Guerrieri M, Wratten C, Steigler A. Impact of radiation therapy on acute toxicity in breast conservation therapy for early breast cancer. *Clin Oncol*. 2004;16:12-16.
12. Twardella D, Popanda O, Helmbold I, et al. Personal characteristics, therapy modalities and individual DNA repair capacity as predictive factors of acute skin toxicity in an unselected cohort of breast cancer patients receiving radiotherapy. *Radiother Oncol*. 2003;69:145-153.
13. Rutter CE, Qin L, Higgins SA, Moran MS, Evans SB. Dosimetric and clinical predictors of the development of moist desquamation in breast cancer irradiation. *J Radiat Oncol*. 2014;2:147-152.
14. Wright JL, Takita C, Reis IM, Zhao W, Lee E, Hu JJ. Racial variations in radiation-induced skin toxicity severity: Data from a prospective cohort receiving postmastectomy radiation. *Int J Radiat Oncol*. 2014;90:335-343.
15. Pollard JM, Gatti RA. Clinical radiation sensitivity with DNA repair disorders: An overview. *Int J Radiat Oncol*. 2009;74:1323-1331.
16. Sharp L, Johansson H, Landin Y, Moegelin I-M, Bergenmar M. Frequency and severity of skin reactions in patients with breast cancer undergoing adjuvant radiotherapy, the usefulness of two assessment instruments—a pilot study. *Eur J Cancer*. 2011;47:2665-2672.
17. Porock D, Kristjanson L, Nikoletti S, Cameron F, Pedler P. Predicting the severity of radiation skin reactions in women with breast cancer. *Oncol Nurs Forum*. 1998;25:1019-1029.
18. O'Grady F, Barsky AR, Anamalayil S, et al. Increase in superficial dose in whole-breast irradiation with halcyon straight-through linac compared with traditional C-arm linac with flattening filter: In vivo dosimetry and planning study. *Adv Radiat Oncol*. 2020;5:120-126.
19. Jagsi R, Griffith KA, Boike TP, et al. Differences in the acute toxic effects of breast radiotherapy by fractionation schedule: Comparative analysis of physician-assessed and patient-reported outcomes in a large multicenter cohort. *JAMA Oncol*. 2015;1:918.
20. Jagsi R, Griffith KA, Heimburger D, et al. Choosing wisely? Patterns and correlates of the use of hypofractionated whole-breast radiation therapy in the state of Michigan. *Int J Radiat Oncol*. 2014;90:1010-1016.
21. Whelan TJ, Pignol J-P, Levine MN, et al. Long-term results of hypofractionated radiation therapy for breast cancer. *N Engl J Med*. 2010;362:513-520.
22. Haviland JS, Owen JR, Dewar JA, et al. The UK Standardisation of Breast Radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials. *Lancet Oncol*. 2013;14:1086-1094.
23. Freedman GM, Li T, Nicolaou N, Chen Y, Ma CC-M, Anderson PR. Breast intensity-modulated radiation therapy reduces time spent with acute dermatitis for women of all breast sizes during radiation. *Int J Radiat Oncol Biol Phys*. 2009;74:689-694.
24. Harsolia A, Kestin L, Grills I, et al. Intensity-modulated radiotherapy results in significant decrease in clinical toxicities compared with conventional wedge-based breast radiotherapy. *Int J Radiat Oncol*. 2007;68:1375-1380.

25. Choi KH, Ahn SJ, Jeong JU, et al. Postoperative radiotherapy with intensity-modulated radiation therapy versus 3-dimensional conformal radiotherapy in early breast cancer: A randomized clinical trial of KROG 15-03. *Radiother Oncol*. 2021;154:179-186.
26. Court LE, Tishler RB, Allen AM, Xiang H, Makrigiorgos M, Chin L. Experimental evaluation of the accuracy of skin dose calculation for a commercial treatment planning system. *J Appl Clin Med Phys*. 2008;9:29-35.
27. Akino Y, Das IJ, Bartlett GK, Zhang H, Thompson E, Zook JE. Evaluation of superficial dosimetry between treatment planning system and measurement for several breast cancer treatment techniques. *Med Phys*. 2013;40: 011714.
28. Niroomand-Rad A, Chiu-Tsao S-T, Grams MP, et al. Report of AAPM task group 235 radiochromic film dosimetry: An update to TG-55. *Med Phys*. 2020;47:5986-6025.
29. Lewis D, Micke A, Yu X, Chan MF. An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan: Efficient protocol for radiochromic film dosimetry. *Med Phys*. 2012;39:6339-6350.
30. Lewis D, Devic S. Correcting scan-to-scan response variability for a radiochromic film-based reference dosimetry system. *Med Phys*. 2015;42:5692-5701.
31. Chan MF, Lewis D, Yu X. Is it possible to publish a calibration function for radiochromic film? *Int J Med Phys Clin Eng Radiat Oncol*. 2014;3:25-30.
32. Jagsi R, Griffith KA, Moran JM, et al. Comparative effectiveness analysis of 3D-conformal radiotherapy versus intensity modulated radiotherapy (IMRT) in a prospective multicenter cohort of breast cancer patients. *Int J Radiat Oncol*. 2022;112:643-653.
33. Low DA, Gerber RL, Mutic S, Purdy JA. Phantoms for IMRT dose distribution measurement and treatment verification. *Int J Radiat Oncol*. 1998;40:1231-1235.
34. Létourneau D, McNiven A, Jaffray DA. Multicenter collaborative quality assurance program for the province of Ontario, Canada: First-year results. *Int J Radiat Oncol*. 2013;86:164-169.
35. Delombaerde L, Petillion S, Weltens C, De Roover R, Reynders T, Depuydt T. Technical note: Development of 3D-printed breast phantoms for end-to-end testing of whole breast volumetric arc radiotherapy. *J Appl Clin Med Phys*. 2020;21:315-320.
36. Determination of absorbed dose in a patient irradiated by beams of X- or gamma-rays in radiotherapy procedure. *Rep Int Comm Radiat Units Meas*. 1976;os-13(1):1-3.
37. Saibishkumar EP, MacKenzie MA, Severin D, et al. Skin-sparing radiation using intensity-modulated radiotherapy after conservative surgery in early-stage breast cancer: A planning study. *Int J Radiat Oncol*. 2008;70:485-491.
38. Ezzell GA, Burmeister JW, Dogan N, et al. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM task group 119. *Med Phys*. 2009;36:5359-5373.
39. Klein EE, Hanley J, Bayouth J, et al. Task group 142 report: Quality assurance of medical accelerators. *Med Phys*. 2009;36:4197-4212.
40. Miften M, Olch A, Mihailidis D, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM task group no. 218. *Med Phys*. 2018;45:e53-e83.
41. Snir JA, Mosalaei H, Jordan K, Yartsev S. Surface dose measurement for helical tomotherapy. *Med Phys*. 2011;38:3104-3107.
42. Jong WL, Ung NM, Wong JHD, et al. In vivo skin dose measurement using MOSkin detectors in tangential breast radiotherapy. *Phys Med*. 2016;32:1466-1474.