5-2019

DEVELOPMENT OF AUTOMATED RADIOTHERAPY TREATMENT PLANNING FOR CERVICAL AND BREAST CANCER FOR RESOURCE-CONSTRAINED CLINICS

Kelly Kisling

DEVELOPMENT OF AUTOMATED RADIOTHERAPY TREATMENT PLANNING FOR CERVICAL AND BREAST CANCER FOR RESOURCE-CONSTRAINED CLINICS

By

Kelly Kisling, M.S.

APPROVED:

_____
Laurence E. Court, Ph.D.
Advisory Professor

_____
Peter A. Balter, Ph.D.

_____
Rebecca M. Howell, Ph.D.

_____
Anuja Jhingran, M.D.

_____
Kathleen Schmeler, M.D.

APPROVED:

_____
Dean, The University of Texas
MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

DEVELOPMENT OF AUTOMATED RADIOTHERAPY TREATMENT PLANNING FOR CERVICAL AND BREAST CANCER FOR RESOURCE-CONSTRAINED CLINICS

A

DISSERTATION

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Kelly Kisling, M.S.
Houston, Texas

May, 2019

Acknowledgments

I am grateful to have so many people to acknowledge who have contributed to this work and to my development as a researcher. First, I would like to thank my advisor, Dr. Laurence Court for his unwavering optimism and encouragement that undoubtedly kept me moving forward through the difficult aspects of this project. He has a gift for cultivating an environment of scientific curiosity and critical thinking. Also, his passion for creating meaningful change is a source of inspiration.

I would also like to thank my advisory committee (Drs. Peter Balter, Rebecca Howell, Anuja Jhigran, and Kathleen Schmeler) whose guidance and interest in my research have been critical. I would like to especially acknowledge Dr. Howell, who was also my advisor for my MS research. She is a fantastic role model, teacher, and advisor who always pushes for the best interest of her students and for the best work out of them.

I would also like to acknowledge the current and former members of my lab group, from whom I have learned some of my most valuable lessons of grad school: Brian Anderson, Dr. Carlos Cardenas, Dr. Xenia Fave, Dr. David Fried, Skylar Gay, Dr. Rachel Ger, Dr. Scott Ingram, Dr. Yuting Li, Dr. Dennis Mackin, Dr. Rachel McCarroll, Tucker Netherton, Dr. Calli Nguyen, Dr. Josh Niedzielski, Constance Owens, Deborah Raffetto, DJ Rhee, Dr. Yvonne Roed, Dr. Ashley Rubinstein, and Dr. Jinzhong Yang. I have been inspired by their curiosity, insistence on scientific rigor, and staunch opinions on the best way to communicate research.

There are many other staff and faculty at MD Anderson with whom I have collaborated in this work and who have provided me with mentorship. There are many wonderful medical physicists, physicians, and dosimetrists in the Department of Radiation Oncologist who have taken time to discuss clinical matters with me and thoughtfully answered my multitude of questions. Namely, Dr. Simona Shaitelman who has graced me with not only her wisdom in radiotherapy of breast cancer, but also in career advice. Also Dr. Anuja Jhingran who has mentored me in pursuing research in global health and shared her expertise in radiotherapy of cervical cancer. I would like to acknowledge Dr. Beth Beadle for always

giving me thoughtful feedback on the tools I am developing, guiding the integration of my work into the bigger picture of automated planning, and being a great travel partner. I would especially like to thank Dr. Joy Zhang, who has patiently mentored me in the development of software and provided encouragement and friendship from the very start of my PhD. She is a brilliant woman whose work ethic is unmatched, and I was lucky to have gotten to collaborate with her so closely on this project!

I have been fortunate to be able to work on this project with collaborators across the globe. Without them, this project would have gone nowhere and would not have the promise of making a real, meaningful impact on the state of radiotherapy globally. I would like to thank Dr. David Anderson, Kobus Botha, Hester Burger, Monique du Toit, Dr. Nazia Fakie, Dr. Hannah Simonds, Christoph Trauernicht, and many more people at Groote Schuur and Tygerberg Hospitals.

I owe special thanks to my wonderful family for their love and support throughout my life. Thank you to my parents, Doug and Karrie Kisling, who have encouraged us to be independent, strong-willed, and adventurous women, and who have set an example for how hard work truly pays off. And thank you to my big sisters, Amy Kisling and Carter Bisso, for being fantastic role models and friends.

Finally, I cannot fully express the appreciation and gratitude I have for Tom Appleman, my partner in life and outdoor adventures. From the beginning has he provided me with unfailing support for my career and life aspirations and has been a source of strength during the hard days of my PhD. Tom, along with our dog Maple, has made all the small and large moments in between working on this PhD so wonderful. I couldn't imagine a better teammate in life, and am excited for our next chapter, and all the rest to follow. Thank you, Tom.

Acknowledgment of Support: I would like to thank the NIH for funding this work (National Institutes of Health/National Cancer Institute grants UH2-CA202665 and UH3-CA202665) and Varian and Mobius Medical systems for providing support for this work.

DEVELOPMENT OF AUTOMATED RADIOTHERAPY TREATMENT PLANNING FOR CERVICAL AND BREAST CANCER FOR RESOURCE-CONSTRAINED CLINICS

Kelly Kisling, M.S.

Advisory Professor: Laurence Court, Ph.D.

Globally, cancer rates are on the rise, especially in low- and middle-income countries (LMICs). However, many of these countries lack access to radiotherapy, which is due in part to a substantial shortage of the staff necessary to deliver safe and effective radiotherapy. This staff shortage could be mitigated by the automation of the radiation treatment planning process. To this end, we developed automated planning for cervical and postmastectomy breast cancer radiotherapy, the two most common types of cancer in women in many LMICs.

For radiotherapy of cervical cancer in resource-constrained clinics, the recommended treatment technique is a four-field box. We created algorithms to plan four-field box treatments with homogenous dose distributions by automatically determining the beam apertures and relative beam weights. Using our techniques we automatically planned 150 four-field-box treatments and 89% were scored acceptable by radiation oncologists. The dose distributions were more homogenous ($p<0.001$) using automatically optimized beam weights compared with equal beam weights. We also created an automatic quality assurance (QA) technique to verify the clinical acceptability of the beam apertures, which flagged 90% of the unacceptable beam apertures (false-positive rate: 16%).

For radiotherapy of node-positive, postmastectomy breast cancer, it is recommended to treat the chest wall and ipsilateral nodes, while reducing the dose to normal tissues, such as the heart and lungs. We created algorithms to plan three-field treatments (mono-isocentric tangential and supraclavicular fields) on free-breathing patient CTs. The dose distribution was automatically optimized by using mixed energy photon beams and field-in-field dose modulation. Using these algorithms, we automatically planned radiotherapy treatments for 10 left-sided, postmastectomy patients. The plans were evaluated

quantitatively based on their dose distributions, and 90% of the plans met constraints for lung dose, heart dose and target coverage. Physicians accepted all plans either as-is (50%) or with only minor changes (50%). Automatic QA of the plans flagged 92% of the changes requested by physicians.

To assess the risk of failure in our automated treatment planning workflow, we performed Failure Modes and Effects Analysis (FMEA). FMEA showed that a specially-designed automated QA program reduced the risk of automated treatment planning. Additionally, we found that human error is still a prominent cause of potential failures and that manual plan reviews of automatically generated plans are still vital for safe delivery of radiotherapy.

In conclusion, automated treatment planning and QA for radiotherapy of cervical and breast cancers were clinically viable for a majority of patients tested. Our algorithms will be implemented clinically at our partner hospitals in South Africa in the next year.

Table of Contents

List of Illustrations

List of Tables

Chapter 1 : Introduction

1.1 Prevalence of cervical and breast cancer

The majority of the global cancer burden is in low- and middle-income countries (LMICs), where two-thirds of all cancer deaths in the world occur [1]. This amounts to more than eight million cases of cancer and five million deaths from cancer each year in a part of the world that has few resources to fight the disease [1]. Furthermore, cancer rates are on the rise, and the number of cases worldwide is predicted to rise from 14 million to 20 million cases of cancer each year by 2025 [2]. This rate of increase is expected to be the most rapid in LMICs [2].

Breast and cervical cancers are extremely prevalent in LMICs, where they are the first and second most common types of cancer occurring in women [1]. They are also the first and third most common cause of cancer death for women in these regions [1]. More than half of all breast cancers occur in the developing world, where a large proportion of cases are detected in late stages [3]. A large majority of cervical cancer deaths (87%) occur in LMICs, where the risk of dying from the disease is three times higher than in more developed regions [1].

1.2 Radiotherapy for cervical cancer

For locally advanced cervical cancer (stages IB2–IVA), the standard of care treatment is concurrent chemotherapy and radiation, including both external beam radiotherapy and brachytherapy [4]–[9]. The goal of the external beam radiotherapy is to treat the gross tumor and any at-risk tissues in the pelvis, including lymph nodes and paracervical tissue, which reduces the risk of regional spread and the volume of the gross tumor prior to brachytherapy. Overall, radiotherapy has been shown to greatly improve local control of cervical cancer [4]–[9].

There are several primary techniques used to deliver external beam radiotherapy for cervical cancer: conventional, using two to four fields based on the bony anatomy; conformal, using two to four

fields that are designed on the basis of soft tissue contours within the pelvis; and intensity-modulated radiotherapy (IMRT), using five or more fields that are inversely-optimized on the basis of soft tissue contours within the pelvis. The selection of which treatment technique to use depends on the clinical resources available, the dose to normal tissues, and organ motion within the treatment area. The conventional technique is recommended for the treatment of locally advanced cervical cancer in resource-constrained clinics by both the American Society of Clinical Oncology (ASCO) and the International Atomic Energy Agency (IAEA) [10], [11]. The conventional technique uses considerably less resources than both the conformal and IMRT techniques which rely on accurate contouring.

1.3 Radiotherapy for breast cancer

For many stages of breast cancer, radiation is essential for treatment after surgery (lumpectomy or mastectomy). In more advanced stages of the disease, as is common in LMICs, the surgical procedure is usually a mastectomy. If the axillary lymph nodes are found to be involved, radiotherapy of the chest wall and at-risk lymph nodes is the standard treatment and has been shown to reduce the risk of local recurrence and improve survival among these patients [12]–[15]. The targeted lymph nodes include the level III axillary and supraclavicular (SCV) lymph nodes, and sometimes also include the internal mammary chain (IMC) lymph nodes. Including the IMC lymph nodes in the radiotherapy treatment increases the complexity of the radiotherapy and may increase the dose to sensitive normal tissues, such as the heart [16]. Given that the treatment of the IMC lymph nodes remains controversial, they may not be included if there is no indication that these nodes might be involved [17]. In resource constrained settings, the IMC lymph nodes are commonly not treated, as they do not have access to specialized techniques that help reduce heart dose by delivering the radiation to the patient under deep inspiration breath hold.

One of the most common techniques for the treatment of node-positive, postmastectomy breast cancer is a tangential pair of fields, for treating the chest wall, which is matched to an anterior-oblique

field, for treating the SCV lymph nodes. The goal of this technique is to treat the entire chest wall and at-risk lymph nodes (SCV and level III axillary), while minimizing dose to sensitive normal tissues, such as the lung and heart [18]–[21]. Additionally, studies have shown that using radiotherapy techniques which achieve a homogenous dose can significantly reduce painful acute skin reactions [22]. There are a few techniques employed in order to achieve a homogenous dose, including using a combination of high and low energy photon beams and various field modulation techniques, such as wedges or field-in-field (FIF) segments.

1.4 Radiotherapy resources in low- and middle-income countries

Radiation therapy is an essential aspect of the treatment for cancer: approximately half of all cancers require radiation therapy as part of their treatment [23]. Yet in low- and middle-income countries, radiation therapy is only accessible to half the number of people who need it [23]. One of the major contributing factors to the lack of radiotherapy access in less-developed regions is an enormous shortage of specialized radiotherapy staff [24]. Staff shortages in resource-constrained radiotherapy clinics have led to the use of simplified treatment techniques that require less time to plan than modern treatments, but can result in inferior outcomes. The time that staff spend planning treatments has been identified as a major contributor to the cost of radiation therapy for low- and middle-income countries [25]. Shortening, or eliminating, the time required to plan treatments could reduce this huge deficit in radiotherapy staff. As a result, we could lower the cost of radiation therapy.

1.5 Automation in Treatment Planning

Currently, treatment planning is a time-intensive, manual process that is necessary for delivering high quality radiation therapy to every patient. The process of treatment planning can be broken down into three general steps: (1) designing the treatment beam geometry, (2) optimizing the dose distribution delivered by the beams, (3) and quality assurance checks verifying the appropriateness and

safety of each treatment plan. These steps involve major human input from specialized staff, and in order to automate this process, we need to develop and validate novel automation techniques that will replace every manual step of treatment planning.

While work has been done towards automating some of the individual steps of planning, the focus has been to ease workflows in high-resource clinics such as those in the United States. Many of these have focused on automated optimization of inversely-planned treatments, such as IMRT [26]–[29]. However, there are many other manual steps that would need to be completed in order to plan these treatments. For breast cancer, there has been some work towards automation of planning for individual process steps or for a simpler radiotherapy technique which is used for earlier stages of the disease [30]–[34]. However, given the prevalence of more advanced stages of breast cancer in LMICs, a more complex treatment technique is needed but has not yet been automated. There has never before been a concerted effort to develop a fully-integrated automated planning solution that requires no additional inputs from staff for cervical cancer or node-positive, postmastectomy breast cancer.

1.6 Study Goal

The overall objective of this project was to automate treatment planning of radiotherapy for patients with locally advanced cervical cancer and node-positive, postmastectomy breast cancer, including automated verification of plan quality. We validated the techniques that we developed and integrated them as a fully automated planning tool. Our goal was to require only a computed tomography (CT) scan and prescription as input, and with no other human interaction, generate a patient-specific treatment plan.

In order to determine if the plans generated by the fully automated planning tool were acceptable for treatment, we tested the system retrospectively using patient CT datasets from several institutions, including MD Anderson and our partner hospitals located in South Africa. The resulting treatment plans were evaluated by radiation oncologists from several institutions in both South Africa and the United

States. Finally, we created and test techniques to reduce the risk of automated planning by automatically detecting when plans are not of acceptable quality.

Chapter 2 : Purpose and Central Hypothesis

Central Hypothesis:

We hypothesized that 90% of the treatment plans created with our automated planning techniques will be rated as clinically acceptable by radiation oncologists and that our automated quality verification techniques will detect 90% of unacceptable plans.

Specific Aim 1:

Aim: Develop automated treatment planning of radiotherapy for patients with locally advanced cervical cancer.

Hypothesis: Treatments planned by our automation techniques would be rated as clinically acceptable by radiation oncologists for at least 90% of plans.

Aim 1.1: Automated design of the beam apertures.

Aim 1.2: Automated optimization of the beam weights.

The work towards Specific Aim 1 is presented in Chapter 3: Automated Treatment Planning for Radiotherapy of Locally Advanced Cervical Cancer.

Specific Aim 2:

Aim: Develop automated treatment planning of radiotherapy for patients with node-positive, postmastectomy breast cancer.

Hypothesis: Treatments planned by our automation techniques would be rated as clinically acceptable by radiation oncologists for at least 90% of plans.

Aim 2.1: Automated design of the treatment field geometry.

Aim 2.2: Automated optimization of the dose distribution.

The work towards Specific Aim 2 is presented in Chapter 4: Automated Treatment Planning for Radiotherapy of Postmastectomy Breast Cancer.

Specific Aim 3:

Aim: Assess the risk of automated treatment planning and develop automated quality verification techniques for treatment plans resulting from Aims 1 and 2.

Hypothesis: Our quality verification techniques would automatically detect 90% of plans with unacceptable plan quality.

Aim 3.1: Assess the risk of automated planning.

Aim 3.2: Develop automated quality assurance (QA) for cervical cancer automated planning.

Aim 3.3: Develop automated QA for breast cancer automated planning.

The work towards Aim 3.1 is presented in Chapter 5: Assessment of Risk in Automated Treatment Planning.

The work towards Aim 3.2 is presented in Chapter 6: Automated Plan Quality Verification for Cervical Cancer Radiotherapy.

The work towards Aim 3.3 is presented in Chapter 4: Automated Treatment Planning for Radiotherapy of Postmastectomy Breast Cancer.

Chapter 3 : Automated Treatment Planning for Radiotherapy of Locally Advanced Cervical Cancer

This chapter is based upon the following article:

Kisling K, Zhang L, Simonds H, Fakie N, Yang J, McCarroll R, Balter P, Burger H, Bogler O, Howell R, Schmeler K, Mejia M, Beadle BM, Jhingran A, Court L. Fully automatic treatment planning for external beam radiation therapy of locally advanced cervical cancer – a tool for low-resource clinics. *J Glob Oncol*, no. 5. 2019:1-9. doi:10.1200/JGO.18.00107

3.1 Introduction

Global cancer rates are increasing, especially in low- and middle-income countries [1]. By 2025, 20 million cancer cases are predicted worldwide annually [2], of which half would benefit from treatment with radiation therapy [23], [35], [36]. However, many countries lack adequate radiation therapy capabilities [23]; this is due, in part, to staff shortages in these regions [24]. Automating radiation treatment planning could help mitigate this limitation by allowing technology to do a large part of the required work to begin treatment of patients. In addition, an expedited planning process could enable patients to be treated much sooner after diagnosis. According to ASCO and the International Atomic Energy Agency, the recommended treatment of invasive cervical cancer in low-resource settings is a radiotherapy technique known as a “four-field box” [10], [11]. This technique uses four orthogonal beams to treat the gross tumor and at-risk tissues in the pelvis. The beam apertures are based on bony anatomy that is visible in a digitally reconstructed radiograph from each of the beam angles: anteroposterior, posteroanterior, right lateral, and left lateral. Examples of beam apertures are shown in Figure 1.

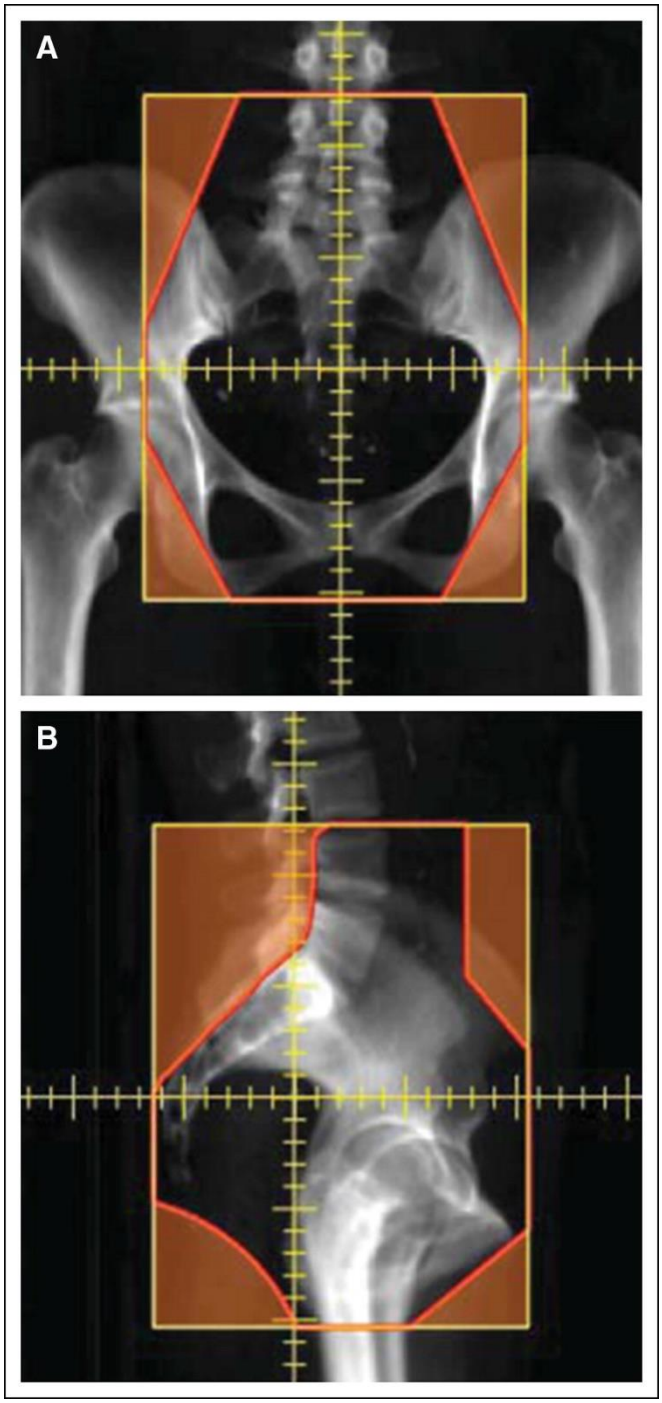

Figure 1. Automatically created treatment fields. Beam's eye view of the (A) anteroposterior (AP) and (B) right lateral beam angles. The beam apertures are designed on the basis of the bony anatomy and will be collimated using the multileaf collimator.

Through a collaboration with hospitals in South Africa and the United States, we have developed a fully automatic treatment planning tool, the Radiation Planning Assistant (RPA) [37]. The RPA designs patient-specific four-field box radiation treatments for locally advanced cervical cancer, one of the most prevalent forms of cancer in low-resource settings [1]. To build the RPA, we developed algorithms to automate every step in the treatment planning process. The RPA has been integrated with a commercial treatment planning system (TPS) to plan three-dimensional treatments on planning CT scans with no human input.

The objective of this study was to validate the individual algorithms of the RPA and to test the fully integrated system on patient CT scans. We retrospectively tested the RPA using patient CT scans from cancer hospitals in the United States and in South Africa. We have also implemented a semiautomated version of the RPA into the clinical workflow at The University of Texas MD Anderson Cancer Center (Houston, TX; hereafter, MD Anderson).

3.2 Methods

All studies and patient data were handled in accordance with the corresponding approved institutional review board protocol, and where required, patient consent was obtained.

3.2.1 Overview of the RPA

To plan a patient-specific treatment with the RPA, the following inputs are used: (1) a CT scan of the patient in the treatment position and (2) a plan order from the physician, which includes basic patient information, including the prescription. With no further human input, the RPA automatically creates a treatment plan that is ready for physician review, along with plan documentation. This documentation is for the patient's medical record and for performing quality assurance checks that are vital to delivering safe radiotherapy [38]. The documentation includes all dose distributions, allowing the physician to review the quality of the plan, including target coverage.

Algorithms have been developed to automate each manual step of treatment planning and have been integrated with the Eclipse TPS (Varian Medical Systems, Palo Alto, CA) using its Application Programming Interface to form the fully automatic TPS. The algorithms that automate each step are described in the following section.

3.2.2 In-House Automation Algorithms

3.2.2.1 Delineation of the body contour

The body contour (Figure 2) is important for accurate dose calculation in the Eclipse TPS. The first step in this algorithm is to identify the location of the couch using the sum projection signal along the lateral direction and then searching for the most representative peak. The couch is then removed from the image by setting all pixels posterior to this line to the CT number of air. The RPA then searches for the body contour by thresholding the CT image intensity (with the couch removed) into a binary mask; it then uses postprocessing to ensure the topologic characteristics and smoothness.

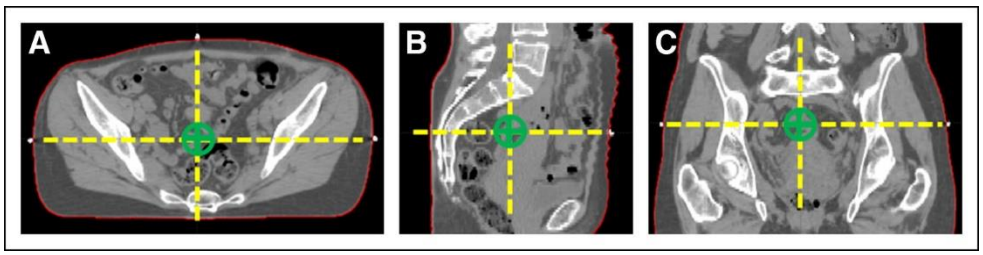

Figure 2. Body contour and marked isocenter. (A-C) Three views, (A) axial, (B) sagittal, and (C) coronal, of the computed tomography scan of a patient. The automatically segmented body contour is outlined in red. The views intersect at the location of the marked isocenter (green), which is determined on the basis of the radiopaque external fiducials. The intersecting planes are denoted by the dashed yellow line.

3.2.2.2 Detection of the marked isocenter

The next step in the RPA is to automatically detect the marked isocenter, as indicated by the intersection of three radiopaque fiducials placed on the patient's skin during the planning CT scan (Figure 2). The RPA automatically detects the marked isocenter by defining a search domain within the

bandwidth of the body contour. Potential fiducial candidates within the search domain are identified on the basis of the CT number. Any false candidates are removed using several criteria, including size, location, and geometry. Finally, the intersection of the selected cluster of three fiducials is used to define the marked isocenter.

3.2.2.3 Design of the treatment field apertures

The RPA then automatically designs the four orthogonal treatment beams, which intersect at the marked isocenter. First, the RPA automatically segments the following bony anatomy on the CT image: bony pelvis, femoral heads, sacrum, and fourth and fifth lumbar vertebral bodies. The RPA uses a deformable, multiatlas technique for automatic segmentation [39], [40]. Next, the RPA projects the segmented anatomy into each beam's eye view (BEV). The RPA automatically identifies anatomic landmarks in the BEV, such as the widest extent of the pelvic inlet, and sets the beam aperture on the basis of these landmarks according to a set of defined rules (eg, 2 cm wider than the pelvic inlet). A representation of this workflow is shown in Figure 3.

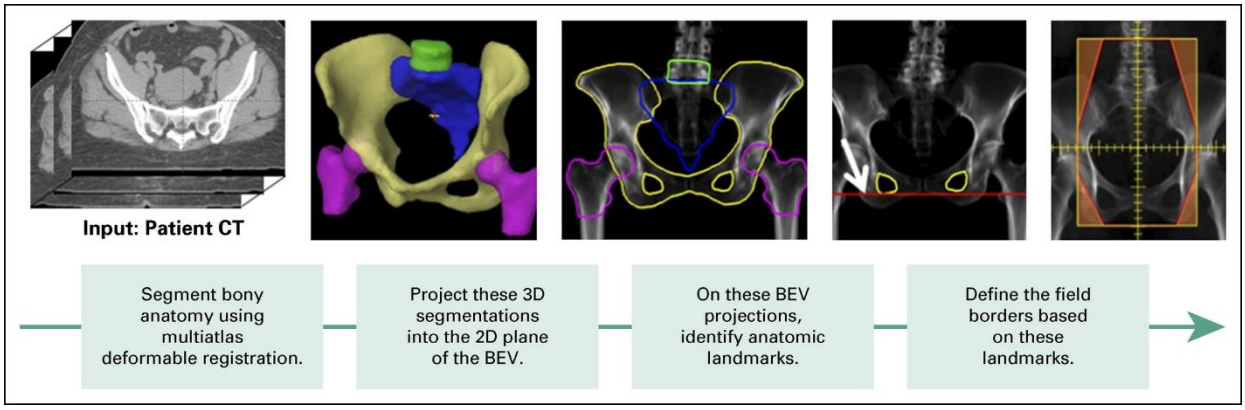

Figure 3. Workflow of the algorithm that automatically designs four-field box treatment beams. For automated planning, the only input is a computed tomography scan and a prescription. No other human input is required, and a plan is presented for physician review. 2D, two dimensional; 3D, three dimensional; BEV, beam's eye view

3.2.2.4 Optimization of the dose distribution

Next, the RPA creates the treatment beams in the Eclipse TPS using the automatically defined beam apertures set at the automatically located isocenter. The RPA then automatically calculates the dose delivered by each beam using 18-MV photons. To achieve a homogenous dose distribution within the treated volume, the RPA automatically determines the weighted contribution of each beam. The RPA uses a least-squares fitting to determine the beam weights that minimize the dose heterogeneity inside the treated volume. The treated volume is defined as the volume intersected by all beams, contracted by a 0.5-cm margin to exclude the rapid dose drop-off at the field edge.

3.2.3 Retrospective Testing of RPA Algorithms

We first tested each algorithm retrospectively on 150 pelvic CT scans of female patients at MD Anderson. Then we tested the fully integrated RPA on 10 CT scans of female patients from Tygerberg Hospital (Cape Town, South Africa) and four CT scans from Groote Schuur Hospital (Cape Town, South Africa). All CT scans had been acquired for radiotherapy planning, with the patients supine.

3.2.3.1 Delineation of the body contour

The automatically delineated body contour from the RPA was compared with the body contour resulting from Eclipse's semiautomated body contour tool, with manual edits where necessary. The two body contours were compared quantitatively using the Dice similarity coefficient, mean surface distance, and Hausdorff distance [39].

3.2.3.2 Detection of the marked isocenter

The automatically localized marked isocenter was compared with an isocenter that had been manually placed at the intersection of the three fiducial markers. The absolute distance between these two points was calculated and used for comparison.

3.2.3.3 Design of the treatment-field apertures

The automatically created treatment-field apertures were reviewed by two physicians specializing in gynecologic radiation oncology, one from MD Anderson (A.J.) and one from Tygerberg Hospital in

South Africa (H.S.). They rated each field as “acceptable” or “not acceptable” for treatment, on the basis of whether they would treat the patient using that field. For a plan to be acceptable, all four fields must have been rated as acceptable.

3.2.3.4 Optimization of the dose distribution

The dose distributions were calculated using automatically optimized beam weights and were compared with nonoptimized dose distributions, which used equally weighted beams (ie, each beam contributed the same dose to the calculation point). The maximum dose, defined by the hottest 1 cc of tissue, was evaluated. We also assessed the coverage, defined by the percentage of the treated volume covered by at least 95% of the prescription dose. The values with and without automated beam weight optimization were compared using a Wilcoxon signed-rank test.

3.2.3.5 Running time of the RPA system

The time for the RPA to automatically plan a treatment was recorded. This included every step, beginning from the import of the CT scan and plan order into the RPA and ending with the optimized and calculated treatment plan in the TPS, ready for physician review.

3.2.3.6 Running the RPA remotely on patients at two South African hospitals

The fully integrated RPA was tested on-site at Tygerberg Hospital and Groote Schuur Hospital. The resulting treatment plans and dose distributions were reviewed by physicians specializing in gynecologic oncology at the corresponding hospital (H.S., N.F.) and rated as acceptable or not acceptable for treatment.

3.2.4 Clinical Deployment at MD Anderson

A semiautomated version of the RPA was created and deployed into the clinical workflow at MD Anderson for patients with cervical cancer in July 2016. This version was integrated with the Pinnacle TPS (Philips Healthcare, Andover, MA). The workflow of this system differs from the fully automated workflow in that the physician manually contours the soft-tissue target volumes on the CT scan. After the CT scan is imported into the TPS, the dosimetrist exports the CT scan to the RPA, The RPA then

automatically detects the marked isocenter and designs the treatment-field apertures (still based on the bony anatomy). Once complete, the RPA automatically sends an e-mail indicating that the plan is ready, and the dosimetrist imports the uncalculated treatment beams. The physician reviews the beams, making any necessary edits on the basis of the contours of the target and critical structures, and planning continues.

We assessed any manual changes to the location of the marked isocenter. We also quantitatively compared the extent of the physician edits to the automatically planned beam apertures, using the mean surface distance and Hausdorff distance.

3.3 Results

3.3.1 Retrospective Testing of RPA Algorithms

3.3.1.1 Delineation of the body contour

A typical result of the automatically delineated body contour is shown in Figure 2. This body contour agreed well with the contour generated using Eclipse's semiautomatic tool with manual edits. The median Dice similarity coefficient was 0.996 (standard deviation [SD], 0.001; range, 0.988 to 0.997). The median mean surface distance was 0.6 (SD, 0.2; range, 0.4 to 1.9) mm. The median Hausdorff distance was 22.3 (SD, 18.6; range, 5.7 to 122.7) mm.

The largest discrepancies were found when the patient's arm was included in only one of the contours. Although these differences may seem large in some patients, they result from differences in how each technique handled the inclusion of the patients' arms. These discrepancies are outside the treatment area and would not affect the dose delivered.

3.3.1.2 Detection of the marked isocenter

The distances between automatically and manually placed marked isocenters were small (average, 1.1 mm; SD, 0.7; range, 0.1 to 2.9 mm). The largest discrepancies were found when the fiducials did not

all appear on the same axial slice of the patient's CT scan. This sometimes led to the isocenters being located on adjacent slices.

3.3.1.3 Design of the treatment-field apertures

An example of the treatment fields generated by the RPA are shown in Figure 1. Of the 150 treatment plans (n = 600 fields), one physician rated 136 (91%) as acceptable for treatment. The second physician found that the image quality of the BEV was too poor in six of 150 plans (four had large amounts of bowel contrast that partially obstructed the bony anatomy in the BEV) and did not rate these six plans. Of the remaining 144 plans, the physician rated 126 (88%) as acceptable. Of the plans marked as unacceptable by at least one physician (n = 23), 19 (83%) had incorrectly placed superior borders as a result of inaccurate contouring of the vertebral bodies during the automatic segmentation step. To overcome this, we will incorporate an option to manually adjust this border in the workspace of the RPA where the physician reviews the treatment plan.

3.3.1.4 Optimization of the dose distribution

Figure 4 shows a comparison of the maximum dose for each patient with automatic beam-weight optimization versus without optimization. The maximum dose was significantly lower using automatically optimized beam weights, with a median change of −1.9% ($P < 0.001$, range −10.0% to +0.4%). In addition, there was a small yet statistically significant increase in the coverage of the treated volume. The median percentage of the volume covered by 95% of the prescription increased by 0.6% ($P < .001$, range: −2.8% to +2.8%).

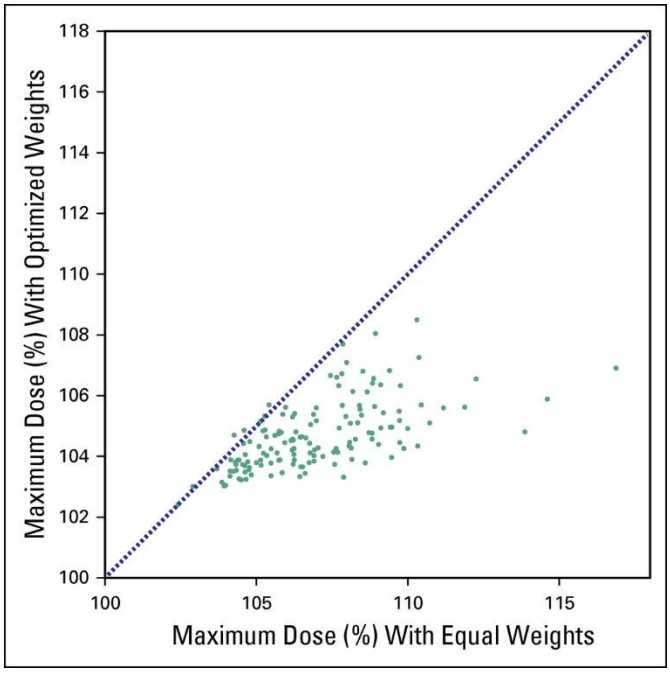

Figure 4. Maximum dose was reduced using automatic beam-weight optimization. The maximum dose (hottest 1 cc) is shown for each patient (n = 149) as a percentage of the prescription dose for optimized versus equal beam weights (nonoptimized). The dotted line represents no change in the maximum dose, and all points below this line showed a reduction in the maximum dose. The reduction was especially large for patients who had very high maximum doses using equal beam weights.

The use of automatic beam-weight optimization was especially beneficial for patients with high maximum dose ($\geq$ 107% of the prescription dose) without optimization. These patients' plans experienced a larger median change in maximum dose ($-$3.5%). Furthermore, the percentage of patients with high maximum doses was reduced from 44% without optimization to 3% with optimization. Figure 5 shows the dose distribution of one axial slice from one patient whose very high maximum dose was greatly reduced using optimized beam weights.

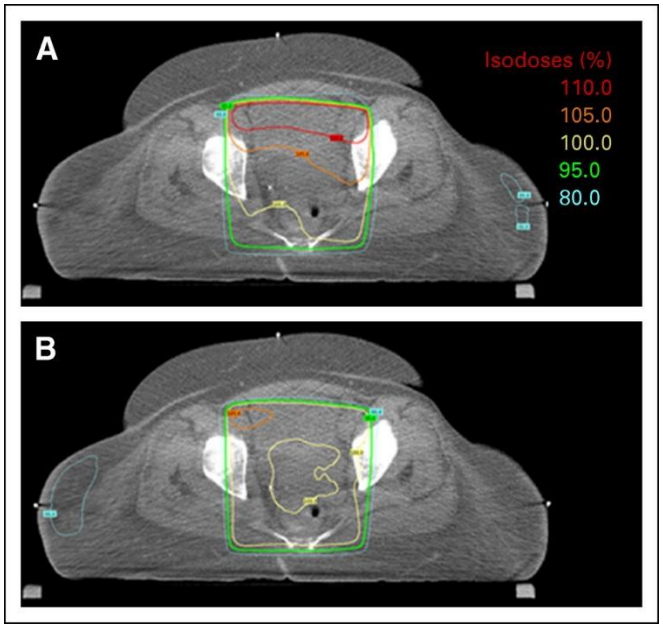

Figure 5. Patient plans with high maximum doses experience a substantial reduction in the maximum dose with automatic beam-weight optimization. The resulting dose distribution for an (A) automatically planned four-field box with equal beam weights (nonoptimized) and (B) automatically optimized beam weights. The maximum dose was reduced from 117% to 107% of the prescription dose for this patient.

3.3.1.5 Running time of the RPA system

Once the planning CT scan and plan order were imported, the fully integrated RPA created a plan in Eclipse ready for physician review in a median of 11.0 minutes (range, 8.2 to 13.6 minutes).

3.3.1.6 Running the RPA remotely in patients at two South African hospitals

Of the 14 treatment plans created on the planning CT scans of patients from the partner hospitals in South Africa, 100% were approved for treatment by the physician (10 plans from Tygerberg Hospital and four from Groote Schuur Hospital).

3.3.2 Clinical Deployment at MD Anderson

Since the clinical version of the RPA was deployed at MD Anderson, it has been used in the planning of 24 patients with cervical cancer. The location of the marked isocenter was not adjusted for 20 patients and was adjusted less than 1 mm for four patients. The physicians edited the automatically

created treatment fields on the basis of their contours of the target and normal tissues. When comparing the fields before and after physician edits, the median mean surface distance was 3.5 mm (SD, 2.4 mm; range, 0.0 to 10.4 mm) and the median Hausdorff distance was 13.9 (SD, 9.1; range, 0.0 to 42.0) mm.

3.4 Discussion

In this work, we validated the RPA's algorithms with physician review of a large cohort of patients and performed remote testing of the fully integrated RPA. This work represents a critical step before implementation of the fully automated system in the clinic. To our knowledge, this is the first work toward automated treatment planning for radiation therapy of cervical cancer.

Before this study, the algorithms for defining the beam apertures were honed over several testing iterations on more than 250 patient CT scans with feedback from physicians at MD Anderson and Tygerberg Hospital and on the basis of the clinical edits made by physicians using the MD Anderson-deployed version of the RPA. The final algorithm, validated in this study, was a consensus of the radiation oncologists for patients whose disease extent was limited to the upper two-thirds of the vagina and with only pelvic lymph node involvement. In the future, we can extend this work for patients with more advanced disease (eg, involvement of the distal vagina or paraaortic nodes) by including variations on the beam-aperture definitions. Within the RPA workflow, the rules by which the beam apertures are defined can be adjusted for a range of disease stages, as long as these rules are based on automatically segmented bony anatomy.

In addition to extensive retrospective testing at MD Anderson, we conducted a successful retrospective test of remote, fully automatic treatment planning at two clinics in South Africa. Moving forward, we will prepare for clinical deployment and testing, beginning with our two partner clinics in South Africa. We will monitor the prospective use of the RPA and evaluate its effect on clinical workflow, including the time staff spend planning and the time from CT simulation to first treatment.

During this testing, we expect to address challenges on the basis of differences in clinical workflow and software and hardware platforms. Ultimately, our goal is to deploy in clinics with fewer resources, which will likely introduce new challenges in terms of staffing, workflow, and equipment. We also are evaluating options to make this tool accessible to low-resource clinics, considering that there may be limited financial resources available. In addition, we are developing automated treatment planning for head-and-neck cancer radiation therapy [41] and postmastectomy chest-wall radiation therapy [37]. The treatment technique planned by the RPA is recommended for cervical cancer in low-resource clinics, according to the International Atomic Energy Agency and ASCO [10], [11]. Although treatment apertures on the basis of soft-tissue contours would be preferable for curative treatments, the bony anatomy approach is used as an alternative in low-resource settings where there is a lack of staff to complete the manual contouring necessary for more conformal treatments. With plans created by the RPA, the physician can use the automatically created documentation to review the dose distribution and evaluate the plan's coverage, even without having contoured the soft-tissue disease.

Given the prevalence of cervical cancer, the fully automatic treatment planning offered by the RPA could help alleviate staff shortages in low-resource clinics. In addition, by reducing the back-and-forth handoffs between planners and physicians needed to manually plan a treatment, the automated system could prepare a plan more quickly, presenting a plan for review shortly after the CT scan is acquired. We envision the RPA facilitating same-day treatments, where a patient never has to leave the clinic between CT scan and her first treatment. In contrast, for patients with gynecologic pelvic disease in our clinical practice, the median planning time is 21 hours (interquartile range, 7 to 47 hours) from CT simulation to when the plan is ready for physician review, including handoffs and time when the plan is not being actively worked on (unpublished data). Furthermore, handoffs between staff have been identified as a weakness in radiotherapy safety, and any reduction in the number of handoffs may result in an improvement in the safety of radiation therapy [42], [43].

The results of this study indicate that fully automatic treatment planning for cervical cancer is achievable. More prospective studies are necessary and are ongoing in the United States and planned with our international partner clinics. By reducing the work required by trained staff, the RPA could ease the burden of creating patient-specific treatment plans in resource-constrained clinics. As a result, using the RPA to automatically plan treatments could help reduce some of the barriers to establishing radiation therapy programs.

Note: A statement of appropriate use of the automated planning tool for cervical cancer can be found in Appendix A.

Chapter 4 : Automated Treatment Planning for Radiotherapy of Postmastectomy Breast Cancer

4.1 Introduction

Breast cancer is the most common cancer in women worldwide, including many low- and middle-income countries (LMICs) [1]. Generally, breast cancer is diagnosed at more advanced stages in LMICs compared with more developed countries [3]. For breast cancer with four or more positive lymph nodes, the standard of care is postmastectomy radiotherapy (PMRT) to the chest wall and ipsilateral lymph nodes, which reduces the risk of local recurrence and improves survival [12]–[14]. There are also increasing indications to deliver PMRT to patients with one to three positive lymph nodes or those with high-risk node negative disease, as radiation in such situations is associated with improvement in disease-free survival [15]. Planning PMRT treatments can be difficult and time-consuming, as it involves using a complex combination of matched fields and various techniques to reduce the dose to organs at risk (OARs) and improve the homogeneity of dose to the targets. Planning such treatments is further complicated by the lack of access to technologies that facilitate deep-inspiration breath-hold techniques that reduce the dose to the heart, which is often the case in resource-constrained clinics in LMICs. These countries also have insufficient access to radiotherapy [23] in part because of a sizable shortage in the trained staff needed to plan and deliver radiation treatment [24]. Treatment planning constitutes a substantial amount of radiotherapy staff workload, and that workload could be reduced by increased automation.

To date, most work on automating treatment planning for breast cancer has focused on a tangential field-only treatment technique or on specific steps in treatment planning, such as the inverse-planning of tangential intensity-modulated radiotherapy [30]–[34]. Many of these techniques are effective and have improved the efficiency of treatment planning. Expanding from these efforts, we

have developed a tool that automates the entire planning process for PMRT, which is necessary for treating more advanced breast cancers. The PMRT technique differs from the previously automated tangential field-only technique in that it includes a supraclavicular (SCV) field that is matched to tangential fields via a non-divergent border at the match line. Another unique feature of our current automation technique is that previous automation techniques require a particular placement of external fiducial markers, which is not standardized among clinics; the techniques and materials used for placing these markers vary greatly. Our goal in the present study was to develop an automated technique for planning PMRT that can be widely used at multiple institutions around the world. Thus, we designed a tool that does not require placement of external fiducial markers to determine the borders of the treatment fields.

Herein we describe the automated treatment planning tool we developed, including the techniques used for automation and the results of a planning study for patients with breast cancer who underwent PMRT. We developed this automated planning tool in a collaboration between institutions in the United States and South Africa, and it is intended for use in resource-constrained settings for radiotherapy for locally advanced breast cancer after mastectomy.

4.2 Materials and Methods

4.2.1 ***Overview of the automated planning tool***

The automated planning tool tested in this work designs PMRT treatments using a monoisocentric tangential and SCV field technique. To reduce the dose to OARs and improve dose homogeneity in the targets, the algorithms in this tool optimize the use of mixed high- and low-energy photon beams and, for the tangential fields, the use of field-in-field (FIF) segments. The automated planning tool was developed assuming the radiation treatments would be planned on a free-breathing computed tomography (CT) scans of patients in the head-first, supine position owing to resource limitations. The only external fiducial markers required are those indicating the position of the marked isocenter (i.e.,

wires are not necessary for determining the beam geometry). The external fiducial markers indicating the marked isocenter are automatically detected as described previously [44]. The initial version of this tool was developed for left-sided treatments only. Given the additional complexity of left-sided treatments because of the heart's proximity to the targets in these treatments, translating this approach to right-sided treatments should be easier than translating it in the opposite direction. Patient treatments were prescribed for a hypofractionated regimen of 40.05 Gy delivered in 15 fractions [45].

The techniques for PMRT planning automation used algorithms developed in-house that were integrated with the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA, USA) via its application programming interface. This PMRT planning tool is part of a suite of automated planning tools called the Radiation Planning Assistant that is being developed to automate planning processes for resource-constrained clinics [37] and currently includes treatment planning for radiotherapy of cervical and head and neck cancers [41], [44]. The inputs required by the Radiation Planning Assistant are a plan directive from the physician specifying the prescription and a CT scan for treatment planning. Another input needed for automatically planning PMRT is the location of the inferior border of the tangent. This additional user input is currently required because an automated technique for identifying this border that is sufficiently reliable has yet to be found, largely because of substantial variability between patients. Therefore, the user is required to identify the CT slice of the tangent fields' inferior border before automated planning is initiated.

The overall workflow for our automation technique for PMRT is illustrated in Figure 6. After automatic segmentation of the targets, OARs, and additional planning structures are two main automated planning steps: (1) setting up the treatment beams and (2) optimizing the dose distributions. These steps are described in detail below. Once automated planning is completed, the user is presented with a composite treatment plan consisting of the tangential and SCV field plans, calculated dose, and heart and ipsilateral (left) lung contours, all of which are created automatically.

We also implemented a technique to automatically verify the quality of the treatment plan and determine when it deviates from standard quality metrics to flag these deviations to the reviewer of the plan.

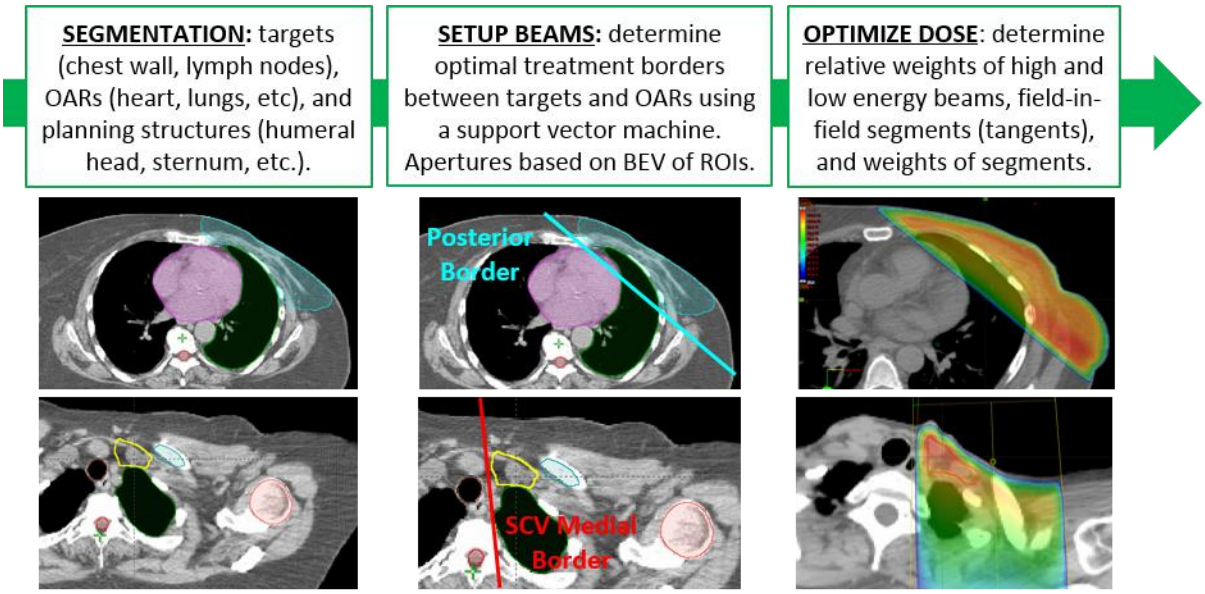

Figure 6. Overview of the method for automated planning of postmastectomy radiotherapy. OARs, organs at risk; BEV, beam's eye view; ROIs, regions of interest.

4.2.1.1 Automation of segmentation

The first step in our planning process is to automatically segment various anatomic structures, including the targets (chest wall and lymph nodes), several OARs (e.g., both lungs, heart, spinal canal), and other structures useful in defining beam geometry (e.g., sternum, clavicle, trachea). Automated segmentation was done by deforming multiple atlases of patient contours to the target patient and combining the deformed contours by using fusion based on the Simultaneous Truth and Performance Level Estimation [46], [47]. These atlases were created by our group and consisted of 11 patient CT scans with contours. This multi-atlas segmentation approach has been successfully used for many other anatomic sites [39]–[41].

4.2.1.2 Automation of beam setup

Once automatic segmentation was complete, the next step was to determine the CT slice of the isocenter, representing the match line between the tangential fields and the SCV field. This slice was initially placed at the inferior aspect of the clavicular head. From there, the posterior, non-divergent border of the tangential fields (the principal border of the tangents separating the chest wall from the OARs) was determined by using support vector machine classification. This technique was adapted from the work described by Zhao et al. [33] and assigned points within the contours to one of two classes: target (chest wall) and avoidance (heart, lungs, and contralateral breast). These classified points were used as inputs to the support vector machine to determine the optimal three-dimensional plane separating the two classes. This plane represents the posterior border of the tangential fields and can be used to derive the beam parameters for the medial and lateral tangential fields, including the gantry angles, collimator angles, and jaw/multileaf collimator (MLC) positions defining the posterior border. The collimator angle is set to zero, and the posterior border is defined using the MLC. The anterior jaw is defined to provide 2 cm of flash from the projection of the body contour. The inferior jaw is defined at the projection of the slice of the body contour at the inferior border (which was previously defined manually).

The first step in defining the SCV beam parameters after automatic segmentation was to determine the optimal medial border separating the targets (SCV lymph nodes) from the avoidance tissues (trachea and spinal canal). Again, a support vector machine was used to find the optimal plane, and this plane was used to define the gantry angle and medial jaw/MLC positions of the SCV field. Finally, the superior jaw was determined based on the beam's eye view (BEV) projection of the cricoid cartilage, and the lateral jaw and MLC positions for humeral head blocking were determined based on the BEV projection of the humeral head.

In some cases, the location of the match line (and isocenter in the superior-inferior direction) had to be automatically adjusted based on the patient's anatomy and beam geometry. If the tangential

field length exceeded the machine capabilities (>20 cm for a Varian C-arm linear accelerator) or the part of the SCV field inferior to the humeral head was insufficient (<2 cm), the location of the match line was automatically adjusted toward the inferior direction, and planning continued. In some cases with excessive lung exposure in the SCV field (>4 cm based on the projection of the lung in the BEV), the match line was moved more toward the superior direction. In the latter case, two plans were created: the original plan with the match line at the inferior aspect of the clavicular head and an alternative plan with a more superior match line to reduce the amount of lung in the SCV field. The rationale for creating two plans was that several clinical factors contributed to the decision to move the match line, including the location of the level III axillary nodes and the possibility of the tangential fields intersecting part of the patient's arm. If upon reviewing the original plan the physician decides that moving the match line in the superior direction is advantageous, he or she can review the alternative match line plan. The alternative match line plans were created while still considering the constraints of maximum tangential field length and proximity of the match line to the humeral head. Figure 7 illustrates an example of the resulting SCV field BEV for the original and alternative match line plans for one of the test patients.

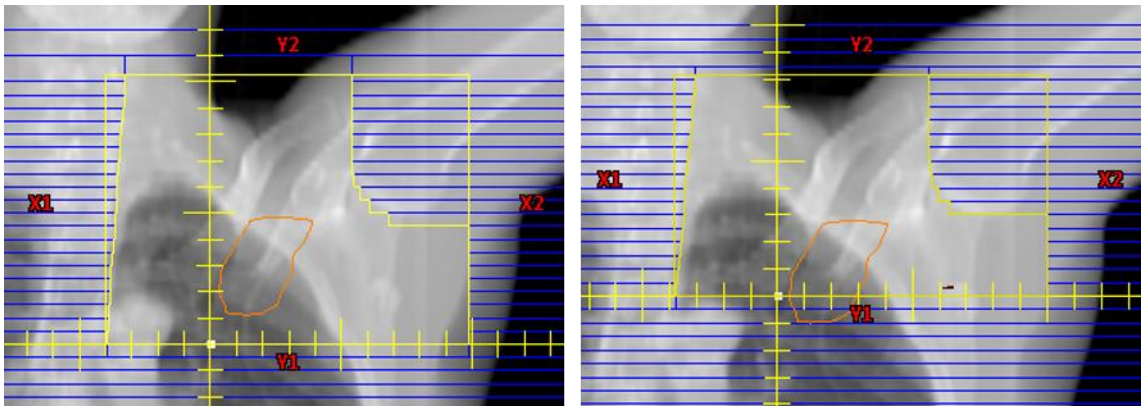

Figure 7. Beam's eye view of supraclavicular fields for the original match line plan (left) and alternative match line plan (right) for the same patient. The contour of the level III axillary nodes are shown projected in orange.

4.2.1.3 Automation of dose optimization

Once the parameters for the open beams were determined, the three beams (medial and lateral tangents and SCV) were automatically generated in the treatment planning system and calculated for both 6- and 18-MV photons. Next, the dose per beam was exported and used as inputs for optimizing the dose distribution to reduce the maximum dose in the plan and the doses to OARs while maintaining target coverage. First, a normalization volume was created for each plan (tangential and SCV). These normalization volumes were derived from the automatic segmentations of the target structures (chest wall for the tangential plan and SCV and level III axillary lymph nodes for the SCV plan) that fell within the limits of the treatment fields. The normalization volumes were used to ensure that coverage was maintained throughout the optimization such that 95% of the normalization volume received 95% of the prescribed dose.

In the optimization of the dose distribution, the relative weights of the high- and low-energy beams were determined first for the open fields by using a brute-force search strategy. Next, for the tangential fields, one or two FIF segments were created iteratively by blocking the dose cloud of the hot spots within each BEV, mimicking a forward-planning methodology. The dose of the hot spot was determined adaptively based on the current maximum dose in the optimization and the relative size of the dose cloud. The beam energy of each FIF segment and relative weighting were also determined by using a brute-force search strategy. The final resulting plans could have at most two SCV fields (high- and low-energy fields, no FIF segments) and four tangential field segments per beam angle (high- and low-energy open fields, two FIF segments). Combining the FIF segments with the open fields would result in a maximum of six treatment beams (two SCVs, two medial tangents, and two lateral tangents).

4.2.2 Evaluation of automated treatment planning

The performance of the automated planning tool was evaluated by using scans from a sample of 10 patients who underwent left-sided mastectomy and a free-breathing CT scan for radiotherapy planning at a partner hospital in South Africa. These CT scans were acquired with the patient in the head-first,

supine position on a breast board with both arms raised over the head. All patients had external fiducial markers indicating their marked isocenters. Some patients had additional wires placed for treatment planning, although the wires were not used in this study. These patients' CT scans were not used during algorithm development or preliminary testing of the automatic planning tool. All patient data used in this study were handled in accordance with an approved institutional protocol.

4.2.2.1 Validation of automatic segmentation

Although our automated techniques for PMRT planning make use of several automatically segmented structures for creating treatment plans, only the heart and ipsilateral (left) lung are presented with the plan. Therefore, the accuracy of segmentation of these contours was validated by comparing the automatically generated contours with physician-approved, manually edited contours. The physician-approved contours were created by using our automatic segmentation tool and then edited manually and reviewed and approved by a radiation oncologist (S.F.S.) with expertise in treatment of breast cancer. The contours were created according to the guidelines provided in the Breast Cancer Atlas for Radiation Therapy Planning from the Radiation Therapy Oncology Group [48]. The contours were compared geometrically by using the Dice similarity coefficient, mean surface distance, and Hausdorff distance and dosimetrically by using differences in dose metrics for each set of contours from the automatically planned treatments.

4.2.2.2 Assessment of the automatically created treatment plans

Once the automatically created plans were ready in the treatment planning system, they were reviewed for acceptability for treatment by two radiation oncologists with expertise in the treatment of breast cancer (D.A. and T.T.) at Groote Schuur Hospital (Cape Town, South Africa). The plans were rated on a three-tiered scale: use plan as is, use plan with minor changes, and plan requires major changes. The specific changes requested for each plan were recorded. If a physician requested to see the alternative match line plan for a patient, that plan was shown to the physician (if it had been created),

and the physician selected the preferred plan. The physicians' final plan ratings were reported for their selected preferred plans.

The selected plans were also assessed quantitatively for compliance with dose objectives for target coverage, OARs, and maximum plan dose. These dose objectives were evaluated by using the physician-approved, manually edited contours. Coverage of the following target structures was assessed: the chest wall, SCV lymph nodes, and level III axillary lymph nodes. The preferred dose objectives and acceptable dose limits used for evaluation of the targets and OARs are presented in Table 1. These objectives were determined according to several sources, including recommendations from The Royal College of Radiologists [45], objectives from the Alliance A221505 clinical trial of hypofractionated PMRT (unpublished protocol) [49], and clinical constraints from collaborating institutions. Maximum doses were assessed separately for tangential and SCV field plans (preferred maximum dose <112% of the prescription) as well as for the composite plan (preferred maximum dose <116% of the prescription).

Table 1. Hypofractionated PMRT dose objectives for target coverage and organs-at-risk

Structure	Dose metric	Preferred objective	Acceptable limit
Targets	Volume >95% Rx	>95%	n/a
Targets	Volume >90% Rx	n/a	>90%
Heart	Mean dose	<4 Gy	<6 Gy
	Volume >25 Gy	<7%	<10%
Ipsilateral lung	Volume >17 Gy	<35%	<40%

Rx: prescription; n/a: not applicable.

4.2.3. Automated verification of PMRT plan quality

Preliminary testing of automated PMRT planning techniques demonstrated that plans were most commonly rejected because of their dosimetric properties rather than the geometric design of the beam setup. As a result, automated verification of the dose distribution was integrated into the automated PMRT planning tool to alert the plan reviewer to these potential dosimetric deviations, with alert thresholds set based on published dose objectives or clinic-specific objectives. These thresholds included maximum plan doses and doses to OARs. Additional verifications included those of the amount of lung projected in the BEV of the SCV and beam properties, such as the SCV gantry angle (Table 2). The ability of these verification tests to detect potential deviations in the quality of the automatically planned treatments was evaluated for the 10 patient CT scans evaluated in this study.

Table 2. Thresholds for automated verification tests of automated PMRT plan quality

Test object	Test metric	Threshold
Maximum dose (composite)	Point dose	>116% of prescription
Maximum dose (tangential plan)	Point dose	>112% of prescription
Maximum dose (SCV plan)	Point dose	>112% of prescription
Heart dose	Mean dose	<4 Gy
	Volume >25 Gy	<7%
Ipsilateral lung dose	Volume >17 Gy	<30%
Lung in SCV field	Projection height	<4 cm
SCV gantry angle	Angle off vertical	>15˚

4.3 Results

4.3.1 Validation of automatic segmentation

The results of the geometric comparison of the automatically generated and physician-approved contours of the heart and ipsilateral lung are shown in Figure 8. The Dice similarity coefficient values were all at least 0.85 for the heart and 0.93 for the ipsilateral lung, indicating very good agreement. We also observed good agreement of the mean surface distance, with all values less than 0.5 cm for all contours. The greatest differences between the automatically generated and physician-approved contours, as indicated by the larger Hausdorff distances, occurred when one contour included more slices in the superior or inferior direction than the other.

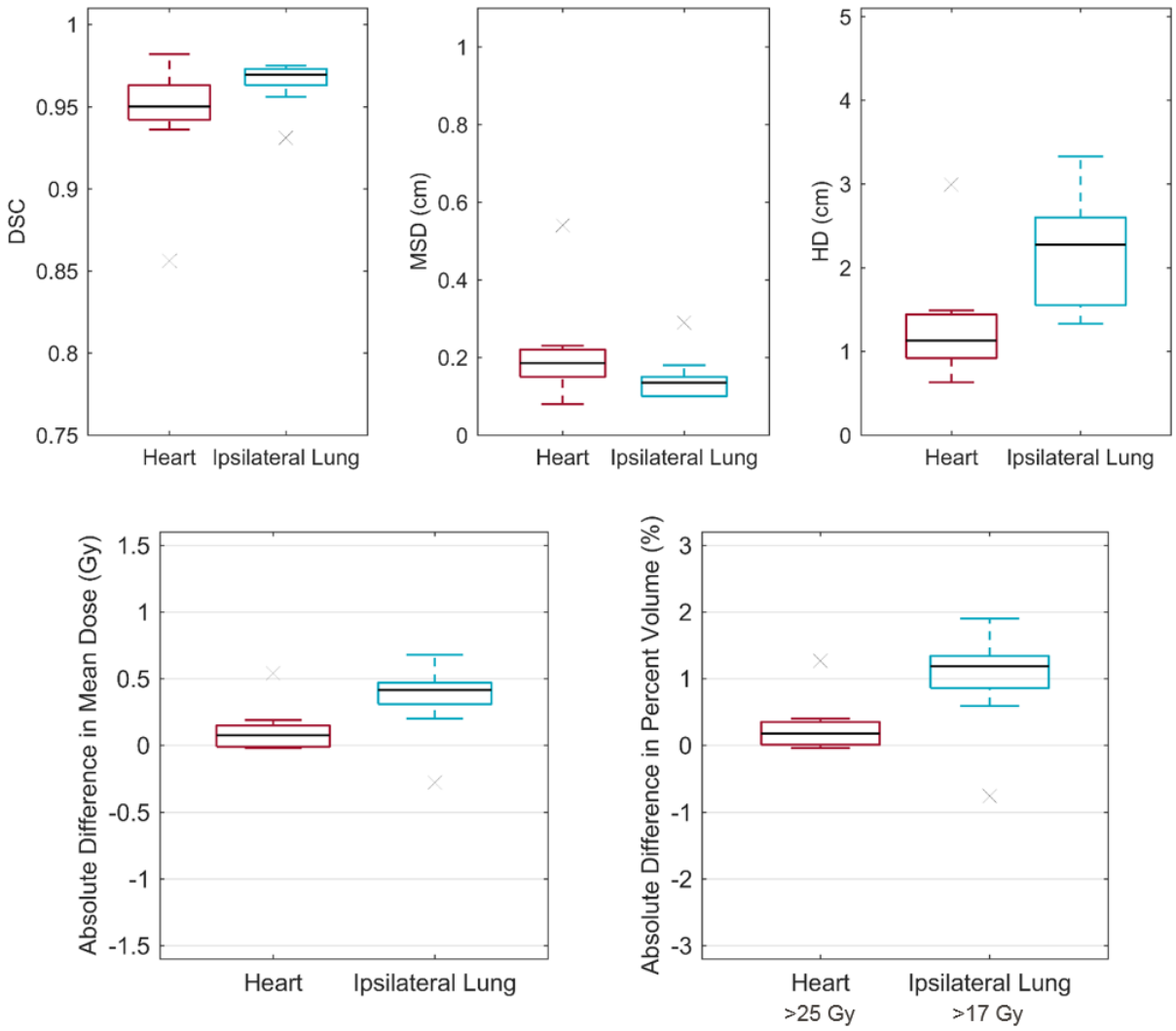

Figure 8. Comparison of the automatically generated and physician-approved contours of the heart and ipsilateral lung. The top row shows the following results of the geometric comparisons: Dice similarity coefficient (DSC; left), mean surface distance (MSD; center), and Hausdorff distance (HD; right). The bottom row shows the following results of the dosimetric comparisons: mean OAR dose (left) and dose-volume histogram metric (right). The absolute difference in metrics was the results for the physician-approved contour minus that for the automatically generated contour for the same plan.

The differences in the dose metrics when evaluated on the physician-approved contours compared with the automatically generated contours of the heart and ipsilateral lung are also shown in Figure 8. The differences were all small, resulting in less than a 1 Gy difference in mean dose and less than a 2% difference in volume for the dose-volume histogram metrics. Although the dose to the physician-approved contours did trend higher than the automatically generated contours, it did not affect whether a plan met the dose objectives for planning. This demonstrates that using the doses to the automatically generated contours is appropriate when reporting plan quality to the user. (Note: the doses reported in the sections below are those to the manually edited contours.)

4.3.2 Assessment of the automatically created treatment plans

Upon physician review of the final 10 automatically planned treatments, the physicians rated all plans either acceptable as is (50%) or with only minor changes (50%). Of these plans, four were the alternative plans, in which the physicians preferred that the match line be placed more superior to the original match line. For the 10 original plans, physicians requested to see an alternative match line plan for five patients. One of these patients was constrained by the tangential field length limit (>20 cm), so alternative plans were presented for the remaining four, all of which were preferred by the physicians. Figure 9 shows an example of the dose distributions for a representative plan rated acceptable as is. Physicians requested a total of seven changes for the five plans: adding a heart block to reduce the heart dose in two patients, reducing the maximum dose in the tangential plan for one patient with a large separation, and reducing the depth of nodal coverage to reduce the maximum dose in the SCV plan, reducing the lung dose, adjusting the superior border of the SCV field, and adjusting the SCV angle in one patient each.

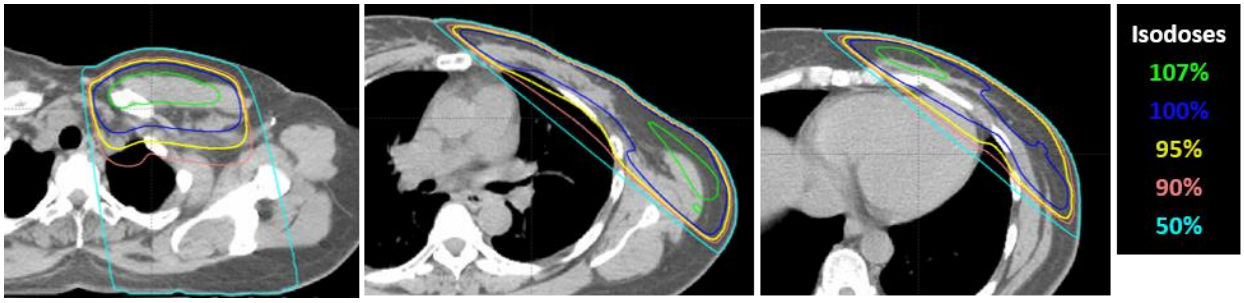

Figure 9. The resulting dose distribution for a test patient for whom the PMRT plan was rated use as is. The dose is shown on three different axial slices of the CT scan: in the supraclavicular field (left), in the superior half of the tangential fields (middle), and in the inferior half of the tangential fields (right).

The resulting dose metrics for the heart and ipsilateral lung, coverage of the targets, and maximum doses are shown in Figure 10. Ninety percent of the plans met the acceptable objective for the heart, with 60% meeting preferred objectives, and all of the plans met the preferred objectives for the ipsilateral lung. Regarding target coverage, all of the plans met the preferred objective, in which 95% of the volume was covered by 95% of the prescribed dose. Eighty percent of the plans met the preferred objective for the maximum dose (<116% of the prescribed dose). The two patients' plans that exceeded this value had maximum doses less than 118% of the prescription dose that were caused by large chest wall separation and deep lymph node targets. The locations of the maximum doses were checked to ensure that they did not fall within the brachial plexus.

On average, we found that creation of one treatment plan took 38 min (range 28-52 min). For plans for which an alternative match line plan was automatically created based on the amount of lung in the SCV BEV, an average of 24 min (range 17-28 min) was added to the planning time. On average, the majority of the time went toward setting up the beams (19 min), contouring (11 min), and dose optimization (7 min). The automated planning tool tested in this system was implemented on a standalone workstation. In the future, the system that will be deployed for clinical implementation will

consist of a network of servers facilitating parallel processes, a distributed calculation framework, and further improvements in computational speed.

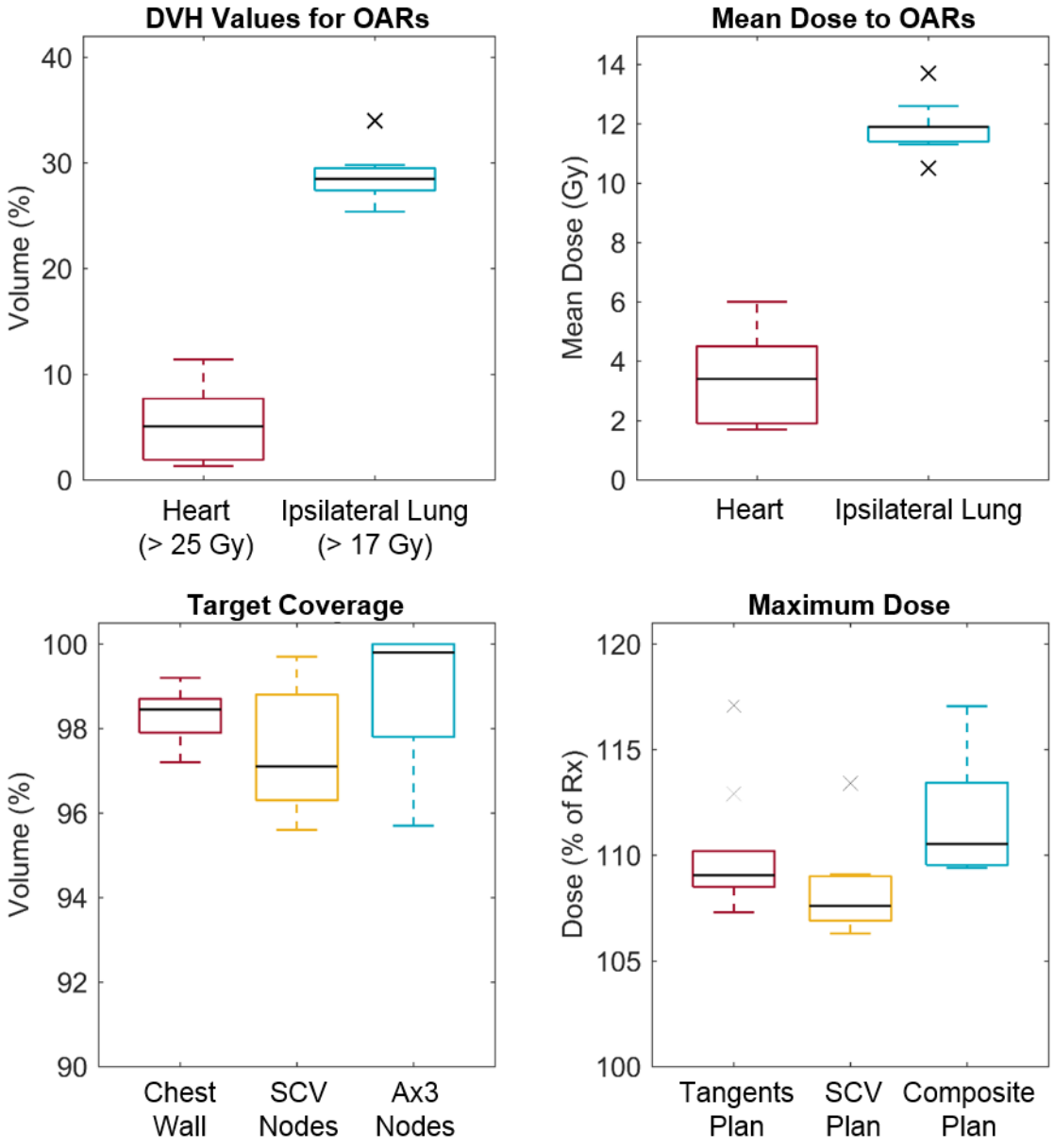

Figure 10. The final dose metrics for the 10 automatically planned treatments. The dose-volume histogram (DVH) metrics (top left), mean dose delivered to the heart and ipsilateral lung (top right), coverage of targets by 95% or the prescription dose (bottom left), and maximum doses for the tangential and SCV plans and for a composite of the two plans (bottom right) are shown. OARs, organs at risk; SCV, supraclavicular; Ax3, level III axillary; Rx, prescription.

4.3.3 Automated verification of PMRT plan quality

In their evaluation of the 10 treatment plans, physicians requested a total of seven changes (described above) plus five requests to review alternative match line plans. Of these 12 requests, 11 (92%) were detected by the automatic plan quality verification tests. The only requested change not detected was adjustment of the SCV field's superior border. The tests detected four additional potential deviations that the physicians did not request to be changed: two plans with slightly high heart doses, one plan with a slightly high ipsilateral lung dose, and one plan with a slightly large amount of lung in the SCV field. All four of these potential deviations were very close to the thresholds set for each test.

4.4 Discussion

In this work, we demonstrated the clinical viability of our automated planning tool for radiotherapy for locally advanced breast cancer after mastectomy. We developed this tool with the goal of reducing the workload on the limited radiotherapy staff of resource-constrained clinics, such as those in LMICs [24]. Such automated tools have the potential to reduce staff workloads and improve the reliability of treatment planning [50]. Investigators have shown that clinical implementation of automated planning for tangential breast irradiation can improve the efficiency and quality of treatment planning, even facilitating same-day treatments [30], [51]. We would expect an even greater gain in efficiency for automation of the complex PMRT necessary for advanced breast cancers, which are more common in LMICs owing to late-stage diagnosis. Some have also suggested that reducing the effort spent in treatment planning through automation could reduce the cost of radiotherapy programs [25]. Automation of PMRT planning is a major step toward improving treatment planning efficiency, especially given that breast cancer is one of the most common cancers in LMICs.

To our knowledge, this is the first automated treatment planning tool designed for PMRT, which uses tangential fields to irradiate the chest wall matched with an oblique en face beam to irradiate the SCV lymph nodes. Although some clinics may also include the internal mammary chain (IMC) lymph

nodes in such treatment, this remains an area of controversy [17]. Treatment of the IMC increases the dose delivered to the heart and the risk of heart disease [16]. In the context of patients undergoing free-breathing treatments because of resource limitations in LMICs, the heart doses and risk of heart disease are even higher. For this reason, clinics may only treat the IMC if these nodes are suspected of being involved. Therefore, we designed our automated planning tool for a technique that irradiates the chest wall and SCV and level III axillary lymph nodes, and not the IMC lymph nodes.

Without using external markers to set up the beams, the automation algorithms determined the appropriate beam angles for the tangential fields in this test cohort. However, not using markers means that one manual input is required for treatment planning: the location of the inferior border of the tangential fields. In the user interface of our automated planning tool, this border is conveniently selected in the same workspace in which the CT scan is approved for automated planning, which should add negligible effort and time to the entire planning process. Also, clinicians may follow their typical process for marking the inferior border, such as placement of a wire, to facilitate their selection of this location. The patient CT scans we tested here as well as those used previously in the development of this automated planning tool came from several different institutions. As a result, a variety of approaches for placing markers and wires were used for these patients, which did not seem to influence the performance of our algorithms in setting up the beams.

We also integrated a method into our automated planning tool to automatically alert reviewers of the PMRT plan to any potential discrepancies in plan quality, which mainly resulted from the plan's dose distribution. Given the proximity of the targets to sensitive normal tissues and the variations in patient anatomy, balancing normal tissue sparing and full target coverage is sometimes a challenge in radiotherapy for breast cancer. We designed our automated treatment planning tool to create plans with full coverage of the target volumes. Therefore, this automatic verification may be useful in alerting the plan reviewer to potential discrepancies in plan quality, such as high heart dose, which may call for adjustment of the plan depending on the acceptable clinical compromises for that particular patient

and plan. The alerts generated by the automatic plan quality verification tests could help expedite adjustment of treatment plans if necessary and improve the safety of automated treatment planning by automatically alerting the plan reviewer to potential issues with the plan.

The current version of this automated planning tool for PMRT designs radiation treatments based on specific clinical practices that may vary by institution, including patient positioning (supine, both arms up), the use of free-breathing scans, the location of the match line or superior border of the SCV field, and a hypofractionated treatment regimen. However, the automation techniques presented herein can be easily adapted to comply with variations in this clinical practice. Some variations in practice may require greater adaptations in the automation technique, but are still feasible. One example is the use of a single-energy photon beam, which is common in many clinics, rather than a mix of high- and low-energy beams. With adaptation, automation of PMRT with a single-energy photon beam is very likely possible, although the dose distributions for many patients would be expected to be hotter than those with mixed-energy photon beams. Another possible adaptation is using tangential and SCV fields for the treatment of intact breast and at-risk nodes. By changing the automatic segmentation to create the breast contour to use as the target for the tangents rather than the chest wall, the automation techniques would function the same as those used for planning PMRT. Before clinical implementation, all of these adaptations would need to be tested thoroughly.

In the evaluation described herein, we retrospectively planned PMRT treatments for a small cohort of patients with locally advanced breast cancer after mastectomy that required only minor changes or were acceptable for treatment as is. We also found that the automatic segmentation of the heart and ipsilateral lung was sufficiently accurate for presentation to the end user. Moving forward, we will subject this automated planning tool to prospective testing with our collaborating institutions before clinical implementation. In this prospective testing, we will compare automatically generated PMRT plans head-to-head with the corresponding manually created treatment plans.

Key considerations for clinical deployment of this automated treatment planning tool is ensuring that patient selection and setup are appropriate for the treatment technique planned by this automated tool. Thorough training will be necessary to ensure safe use and help users understand the specifications of this system. During clinical implementation, we will collect data on planning times to quantify improvements in efficiency when using this automated planning tool.

4.5 Conclusion

We developed and tested an automated planning tool for PMRT and demonstrated its viability for implementation in resource-constrained clinics in LMICs. This tool has the potential to improve efficiency in planning these complex treatments for breast cancer.

Note: A statement of appropriate use of the automated planning tool for cervical cancer can be found in Appendix B.

Chapter 5 : Assessment of Risk in Automated Treatment Planning

This chapter is based upon the following accepted article:

Kisling K, Johnson JL, Simonds H, Zhang L, Jhingran A, Beadle BM, Burger H, du Toi M, Joubert N, Makufa R, Shaw W, Trauernicht C, Balter P, Howell R, Schmeler K, Court L. A Risk Assessment of Automated Treatment Planning and Recommendations for Clinical Deployment. *Med Phys*.

5.1 Introduction

Automation has the potential to improve the consistency and efficiency of radiation treatment planning. Additionally, the automation of treatment planning promises to improve safety by preventing human errors and reducing handoffs between medical staff members, both of which have been shown to be weaknesses in radiotherapy safety [42], [52], [53]. While it is generally assumed that automation leads to elimination of the risks associated with human error, it can introduce new or different types of error that are not part of the routine, manual treatment planning process. For example, automated processes may introduce new or added risks from the lack of active participation by a human user who could catch errors that may go undetected by computer algorithms.

As with any other new technology to be implemented into clinical workflow, it is vital to assess the risk introduced by automated treatment planning in each step of the workflow. Furthermore, Task Group 100 (TG-100) of the American Association of Physicists in Medicine (AAPM) recommends that all new devices undergo a systematic risk analysis [53]. Failure mode and effects analysis (FMEA) is an established technique for methodically and prospectively identifying the risks involved in a process. This method has been used by several other groups to assess the risks of various processes in radiation oncology practice [43], [54]–[57].

Our group recently developed a fully automatic treatment planning tool, the Radiation Planning Assistant (RPA) [37], [44]. In addition to developing algorithms to automate treatment planning, we also implemented a quality assurance (QA) program specific to the RPA in order to enhance the safety

of automated planning. In the present work, we used FMEA to assess the risk of various failure modes in automated planning for cervical cancer radiotherapy with the RPA. We then assessed the impact of the specialized QA program on the identified risks.

5.2 Materials and Methods

5.2.1 Description of the RPA

The workflow of automated planning with the RPA was previously described by Court et al [37]. Currently, the RPA is being implemented as a remote system with which the user interacts via a web interface. A locally installed system is also possible. The input to the RPA is a plan directive from the physician with patient information, including the treatment site and prescription, and the planning computed tomography (CT) images of the patient in the treatment position. In the RPA's user interface, qualified staff can enter, review, and approve the plan directive and planning CT. Once these are approved, the RPA automatically begins planning.

The RPA is being developed for all treatment sites, starting with cancers of the uterine cervix, head and neck, and breast. Here, we focused on RPA treatment planning for cervical cancer using a 4-field box technique with beam apertures based on bony anatomy. The algorithms and validation results of the RPA for cervical cancer have been previously described [44]. Briefly, the RPA uses in-house-developed algorithms that are integrated with the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA) via Eclipse's application programming interface. The marked isocenter is automatically localized according to the positions of 3-point external fiducial markers, and the body contour is automatically segmented. Next, the pelvic bony anatomy is automatically segmented with an auto-segmentation tool using deformable registration of multiple atlases [46]. The pelvic bony anatomy is projected into each beam's-eye view, and the beam apertures are designed on the basis of anatomical landmarks identified on the projections of the bony anatomy. The treatment beam parameters are then automatically set in the treatment planning system, and the dose is calculated

using the Analytical Anisotropic Algorithm. The relative beam weights are optimized to achieve a homogeneous dose distribution within the treated volume.

After planning is complete, the RPA presents the plan as a PDF document for a physician to review. If the physician approves the plan, DICOM-format treatment plan files are transferred to the user. The expectation is that users will import the files into their own treatment planning systems for review before treatment and will perform their standard pretreatment QA.

5.2.2 FMEA of Automated Planning

For the FMEA, a team of subject-matter experts (3 medical physicists and 1 radiation oncologist) first enumerated the steps in the RPA automated treatment planning workflow for patients with cervical cancer, from CT simulation to plan approval by the physician. Then, for each process step, the group of 3 physicists identified potential failure modes and possible causes of each failure mode. The process map, failure modes, and scoring were determined assuming a generic clinic that follows the practices outlined in the American College of Radiology accreditation requirements, including physics plan review [58]. Prior to scoring, the process map and potential failure modes and causes were reviewed by 6 medical physicists at 4 centers in South Africa and Botswana that are prospective users of the RPA. These physicists assessed the applicability of the proposed treatment planning workflow and failure modes to their local clinical practice to identify any substantial discrepancies, and none were found.

Next, the original team of 3 physicists scored the likelihood of occurrence (O), the severity (S), and the likelihood of going undetected (D) for each potential cause of each potential failure mode using a 10-point scoring system. For scoring O and D, we used the values recommended in Table II of the TG-100 report [53]. For scoring S, we used the system recommended by Faught, which augments the original definitions in the TG-100 report with quantitative descriptions of the severity [59]. This system is reproduced here in Table 3. The value for each O, S, and D score was the consensus value as determined by the group. We chose to focus on scoring the causes of failure modes individually in

order to better design the QA program to mitigate the variable risks attributable to different causes of a potential failure mode.

Table 3. The system used for scoring the severity of a potential failure mode, reproduced from Faught [59].

Severity Score (S)	Qualitative Definition	Quantitative Definition
1	No effect	0.0%-2.9%
2	Inconvenience	3.0%-3.9%
3		4.0%-4.9%
4	Minor dosimetric error, suboptimal plan or treatment	5.0%-6.9%
5	Limited toxicity or tumor underdose. Wrong dose, dose distribution, location, or volume	7.0%-8.9%
6		9.0%-9.9%
7	Recordable event. Potentially serious toxicity or tumor under dose	10.0%-14.9%
8		15.0%-19.9%
9	Reportable event. Possible very serious toxicity or tumor under dose. Very wrong dose, dose distribution, location, or volume	20.0%-49.9%
10	Catastrophic	≥50.0%

A risk priority number (RPN) was calculated by multiplying the O, S, and D scores. For potential failure modes with causes that were specifically related to the failure of an RPA algorithm, we were able to determine the O score quantitatively on the basis of our retrospective testing of the RPA using approximately 500 patient CT scans rather than making an estimate of the percent likelihood of occurrence. For example, a specific algorithm failure that occurred twice during testing would receive an O score of 6, representing an occurrence rate of less than 0.5%. If a potential failure mode never

occurred, we estimated the likelihood of occurrence from our experience and knowledge of the algorithms.

5.2.3 Description of the QA Program

To minimize opportunities for error, we implemented a QA program for the RPA that included 3 types of QA to detect errors in the automatically created treatment plans. We applied this QA program, which is intended to supplement users' standard QA processes, to all possible steps in the automated planning process. This QA program was initially developed prior to the FMEA based on our experience with the initial retrospective testing and development. The QA program was updated with additional tests based on the results of this FMEA.

The first type of QA applied was an automatic, independent validation of the results of each step, which is a similar concept to the independent dose verification used in radiotherapy. For example, we use 2 independent methods to automatically detect the marked isocenter based on a 3-point external fiducial setup. The primary method searches within a band around the body contour for the high contrast external fiducials. The secondary method differs in that it searches for high contrast objects that constitute a triangle topology. The result of the primary method is used in the treatment plan, and its agreement with the results of the secondary method is verified. If the 2 methods do not agree, the treatment plan is flagged for human review. Other examples of tasks that have 2 independent algorithms are segmentation of the body contour and creation of the field apertures. We pushed the treatment plans to Mobius3D (Mobius Medical Systems, Houston, TX) to perform the secondary dose verification. We also automatically verified the patient's orientation and anatomical site (e.g., head vs. pelvis) using both the DICOM header information and a simple rigid registration to a full-body patient CT scan.

The second type of QA was an automatic check of the result against expected values. The expected values may be a range derived from the population of patient plans, such as for field size, or a single value, such as the collimator angle always being equal to 0.

Finally, the third type of QA was a series of specialized manual checks. One such manual check was a set of specific questions in the user interfaces for approving the plan directive and planning CT. For example, we asked the user to verify that the CT scan field of view was appropriate for the patient. Another manual check was plan documentation that was automatically created and designed to guide the appropriate staff, such as a physicist, through checking the treatment planning parameters.

5.2.4 Assessment of the Effect of the QA Program

We next assessed the effect of the QA program on the detectability of each failure mode and cause. First, we determined which failure modes could be detected by any of the 3 types of QA described above. For each type of QA, we estimated how effectively that type of QA could improve the detectability of that failure mode and cause using a 3-tier scale: very effectively, moderately effectively, or somewhat effectively. We used these ratings to adjust the D value (the likelihood of going undetected) for each failure mode by reducing the percentage of undetected failures (from the corresponding score values in Table II from TG-100). To facilitate the reduction in the D value, we assigned residual percent undetected values of 20%, 50%, and 80% to the very, moderately, and somewhat effective types of QA, respectively, based on the likely effectiveness as determined by group consensus. These reduced values were then converted back to a D score on the 10-point scale from TG-100. We then calculated the differences in the RPN of all potential failure modes with and without the QA program.

5.3 Results

The RPA process, with its 4 major subprocesses (CT simulation, plan directive, RPA plan creation, and plan approval) and 30 steps, is shown in Figure 11. Using FMEA, we identified 68 failure modes with 113 potential causes. The full results of the FMEA can be found in Appendix C. Of the 113 potential causes, 79 (70%) were subject to at least 1 type of QA as part of our QA program (not including typical clinical QA practices). Without the QA program, the average RPN was 91, and the maximum RPN was

504. With the QA program, these values were reduced to 68 and 315, respectively. The distribution of the RPNs with and without the QA program for all potential failure modes and causes is shown in Figure 12, where the overall shift to lower RPN values is apparent. Since the QA program only affected the detectability of failures, we compared the change in the distribution of the D scores with and without the QA program (Figure 13). The median D score without the QA program was 5.0 and was reduced to 3.0 with the QA program.

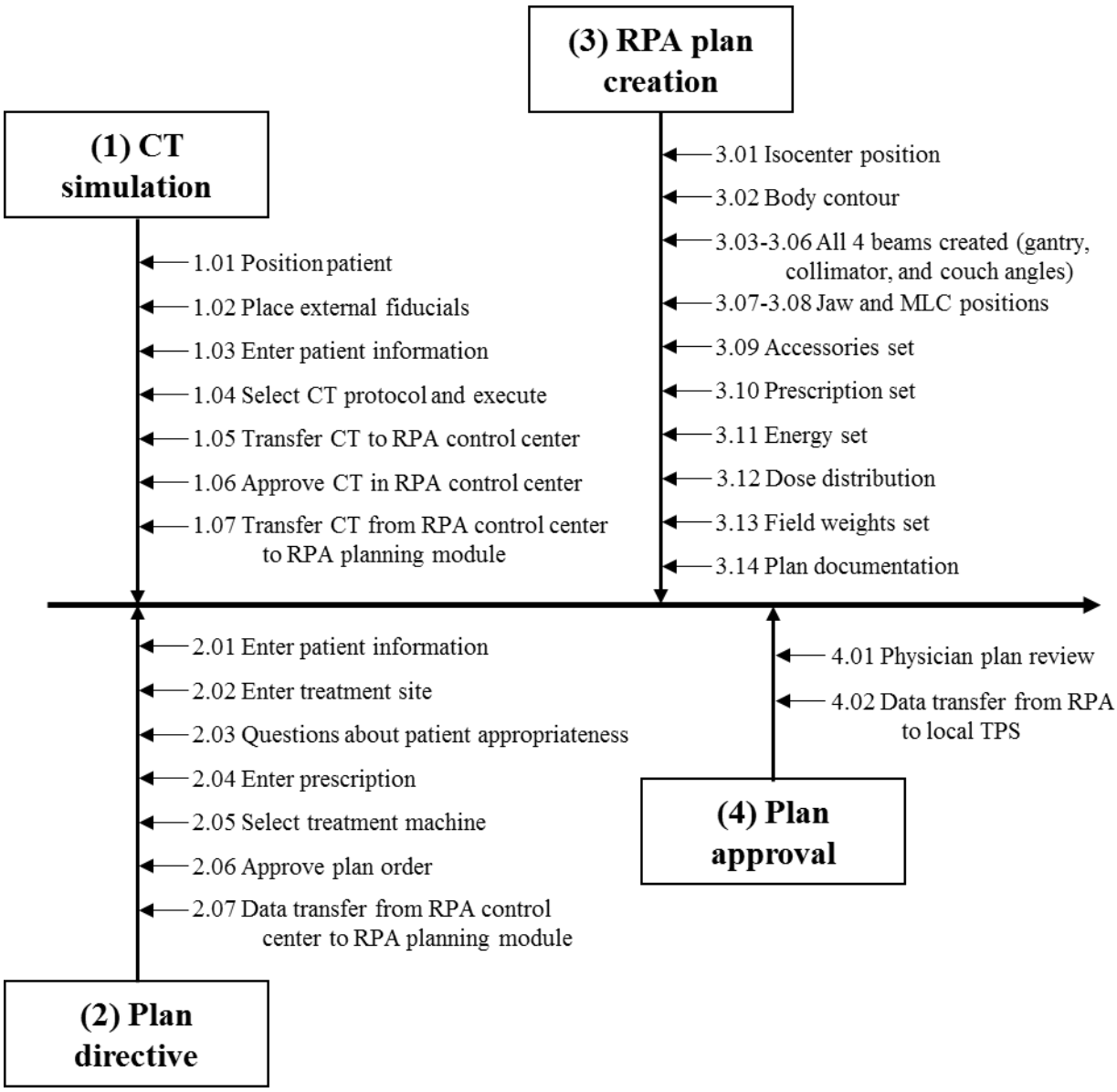

Figure 11. Depiction of the subprocesses and steps involved in automatically planning a 4-field box radiotherapy treatment for cervical cancer with the Radiation Planning Assistant (RPA). Subprocesses 1 and 2 (CT simulation and plan directive) involve many manual steps from which errors could propagate downstream. Subprocess 3 (RPA plan creation) is entirely automatic. Abbreviations: MLC, multileaf collimator; TPS, treatment planning system.

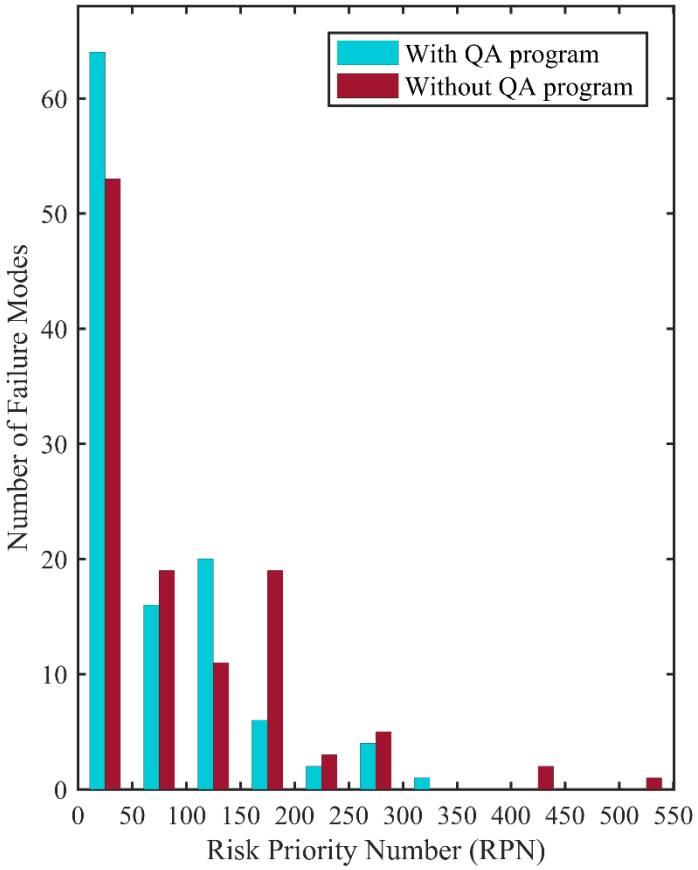

Figure 12. Histogram of the risk priority numbers (RPN) for all potential failure modes identified for automatic planning of a cervical cancer treatment using the Radiation Planning Assistant (RPA) with (blue) and without (red) the QA program.

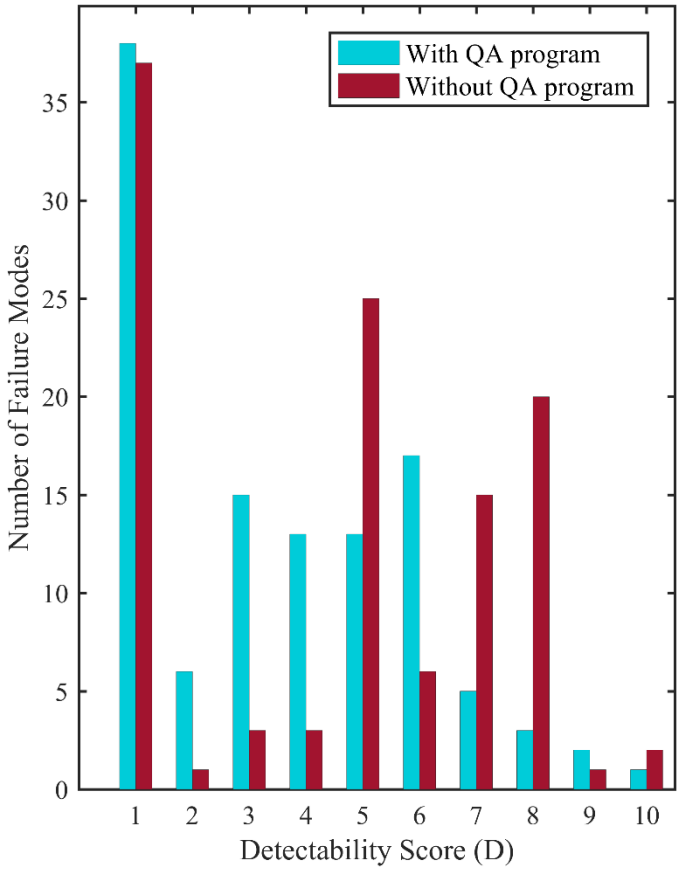

Figure 13. Histogram of the detectability score (D) for all potential failure modes identified for automatic planning of a cervical cancer treatment using the Radiation Planning Assistant (RPA) with (blue) and without (red) the QA program.

Scores for the top 10 potential failure modes and associated causes are shown in Table 4 and Table 5 for the RPA without the QA program and with the QA program, respectively. Without the QA program, 3 of the top 10 failure modes were related to a failure to correctly identify the marked isocenter based on the external fiducials. A failure in this step could have severe consequences if not detected and is not unique to automated treatment planning. Incorrect definition of the isocenter was also identified as a relatively high-risk failure mode in an FMEA of pretreatment steps for TomoTherapy by Broggi et al [55]. The QA program of the RPA includes automatic verification of the isocenter and a guided manual check of the isocenter identification in the plan documentation. With the QA program,

only 1 of the top 10 failure modes involved identification of the isocenter. This failure mode was caused by the presence of other external fiducials on the planning CT scan (e.g., wires), which could reduce the effectiveness of the automatic verification QA process.

Table 4. The top 10 potential failure modes and their causes in automated planning with the Radiation Planning Assistant (RPA) without the quality assurance (QA) program.

#	Major process	Step	Potential failure mode	Potential causes of failure	O	S	D	RPN
1	RPA plan creation	Isocenter position	Incorrectly identified	Other external fiducials	7	9	8	504
2	RPA plan creation	Jaw positions	Inappropriate position	Algorithm error	10	7	6	420
3	RPA plan creation	MLC positions	Inappropriate position	Algorithm error	10	7	6	420
4	Plan approval	Physician plan review	No comprehensive review	Human error	3	10	10	300
5	RPA plan creation	Isocenter position	Incorrectly identified	Algorithm error	4	9	8	288
6	CT simulation	Select CT protocol and execute	Field-of-view is too small	Human error	5	8	7	280
7	CT simulation	Select CT protocol and execute	Field-of-view is too small	Patient is too large	5	8	7	280
8	Plan directive	Enter prescription	Incorrect (not changed from default)	Human error	4	9	7	252
9	RPA plan creation	Dose distribution	Calculation point is inappropriate	Located in high or low CT number	10	4	6	240
10	RPA plan creation	Isocenter position	Incorrectly identified	External fiducials out of range of scan	3	9	8	216

Abbreviations: O, occurrence score; S, severity score; D, detectability score; RPN, risk priority number; MLC, multileaf collimator; CT, computed tomography.

Table 5. The top 10 potential failure modes and their causes in automated planning with the Radiation Planning Assistant (RPA) with the quality assurance (QA) program.

#	Major process	Step	Potential failure mode	Potential causes of failure	O	S	D	RPN
1	RPA plan creation	Isocenter position	Incorrectly identified	Other external fiducials	7	9	5	315
2	Plan approval	Physician plan review	No comprehensive review	Human error	3	10	10	300
3	RPA plan creation	Jaw positions	Inappropriate position	Algorithm error	10	7	4	280
4	RPA plan creation	MLC positions	Inappropriate position	Algorithm error	10	7	4	280
5	Plan directive	Enter prescription	Incorrect (not changed from default)	Human error	4	9	7	252
6	CT simulation	Select CT protocol and execute	Field-of-view is too small	Human error	5	8	6	240
7	CT simulation	Select CT protocol and execute	Field-of-view is too small	Patient is too large	5	8	6	240
8	Plan directive	Questions about patient appropriateness	Completed incorrectly	Human error	4	9	5	180
9	Plan directive	Approve plan directive	Approved by staff without correct rights	Shared login/Incorrect rights	4	9	5	180
10	CT simulation	Position patient	Inappropriate positioning	Human error	6	4	7	168

Abbreviations: O, occurrence score; S, severity score; D, detectability score; RPN, risk priority number; MLC, multileaf collimator; CT, computed tomography.

Other important potential failure modes, with and without the QA program, were the definition of the beam apertures (jaws and multileaf collimator). We implemented automatic QA verification to detect gross errors in beam aperture definition. However, it remains vital that the physician review the appropriateness of the beam apertures, regardless of whether the beam apertures are determined

automatically by computer algorithms or drawn manually by a dosimetrist or resident. In fact, a potential failure mode identified by both our FMEA and that of other groups was failure of the physician review of plan quality [53], [56]. While physician review is standard clinical practice, and such a failure is unlikely, it would be very difficult to detect and could have severe consequences if the plan quality is inadequate. We have previously shown that nonspecialists are unlikely to reliably identify these errors, even when presented with examples of correct beam apertures as a reference [37].

We specifically investigated the 15 higher-risk failure modes—those for which the RPN was greater than or equal to 200. Of these, 13 (87%) were subject to at least 1 type of QA technique. The number of these higher-risk failure modes was reduced from 15 to 7 when the QA program was implemented, with an average reduction in RPN score of 93 points (range, 40-189). Physician review of the plan was 1 of the 2 higher-risk failure modes that could not be mitigated by a specialized QA technique (beyond standard clinical practice).

The other top failure mode that was not subject to a specialized QA technique was an incorrectly entered prescription in the plan directive workspace, specifically cases in which the physician intended to change the prescription from the default or typical value. While such a scenario is unlikely, this failure would be more difficult to catch later in a manual check, since the intended prescription would not be the expected, typical value for that patient type. Automation does mitigate many potential transcription errors (such as those that occur between the plan directive and the treatment planning system). However, the initial entry into the system must be as the planning physician intends, regardless of whether manual or automated treatment planning is used. In fact, incorrect prescription has been identified as a potentially severe failure mode by other groups studying manual treatment planning [55], [56].

We also investigated the 22 failure modes and associated causes with high severity scores ($S \geq 9$). Table 6 lists all failure modes and causes with S scores of 9 or higher, including their overall scores with the specialized QA program. Here, we found that 14 (64%) failure modes were subject to at least 1 type

of QA. On average, the RPN of the high-severity failure modes was reduced from 146 to 113 with the QA program implemented. The maximum RPN for these was reduced from 504 to 315. Most of these failure modes were unlikely to occur; 20 (91%) had O scores of 4 or less. With 1 exception, the high-severity potential failure modes that were not subject to any specialized QA test were caused by human error, such as incorrectly entering the prescription in the plan directive.

Table 6. Potential automated treatment planning failure modes and associated causes with severity (S) scores of 9 or higher. Scores shown are for the Radiation Planning Assistant (RPA) with the quality assurance (QA) program implemented.

Major process	Step	Potential failure mode	Potential causes of failure	O	S	D	RPN
CT simulation	Enter patient information	Incorrect name or ID entered	Human error	1	10	9	90
Plan directive	Enter patient information	Incorrect name or ID entered	Human error	2	10	1	20
RPA plan creation	Prescription set	Does not match the plan directive	Algorithm error	1	10	5	50
RPA plan creation	Prescription set	Incorrect normalization	Algorithm error	1	10	5	50
RPA plan creation	Dose distribution	Calculation point not at isocenter	Algorithm error	1	10	5	50
RPA plan creation	Plan documentation	Data corrupted	Algorithm error	3	10	4	120
Plan approval	Physician plan review	No comprehensive review	Human error	3	10	10	300
Plan approval	Data transfer from RPA to local TPS	Data corrupted	Network error	2	10	3	60
CT simulation	Position patient	Incorrect orientation	Human error	3	9	2	54
CT simulation	Position patient	Incorrect orientation	Standard technique varies from RPA protocol	6	9	1	54
CT simulation	Position patient	Incorrect orientation	Intentional non-standard technique	4	9	2	72

Plan directive	Questions about patient appropriateness	Completed incorrectly	Human error	2	9	7	126
Plan directive	Questions about patient appropriateness	Completed incorrectly	Human error	4	9	5	180
Plan directive	Enter prescription	Incorrect (not changed from default)	Human error	4	9	7	252
Plan directive	Enter prescription	Incorrect (changed from default)	Human error	3	9	6	162
Plan directive	Approve plan order	Approved by person without correct rights	Shared login/incorrect rights	4	9	5	180
RPA plan creation	Isocenter position	Incorrectly identified	Other external fiducials	7	9	5	315
RPA plan creation	Isocenter position	Incorrectly identified	Fiducials out of range of CT	3	9	4	108
RPA plan creation	Isocenter position	Incorrectly identified	Algorithm error	4	9	4	144
RPA plan creation	All 4 beams created	Not created at the isocenter	Algorithm error before aperture generation	1	9	6	54
RPA plan creation	All 4 beams created	Not created at the isocenter	Algorithm error after aperture generation	1	9	2	18
RPA plan creation	MLC positions	MLC missing from plan	Algorithm error	1	9	2	18

Note: Potential failures in the "Plan directive" step "Questions about patient appropriateness" were scored by considering 2 separate scenarios: (1) when the result does not affect the results of automated planning, but still poses a risk (such as prior irradiation); and (2) when the result would technically affect the result of automated planning (such as the presence of an artificial hip, which may cause errors in contouring the bony anatomy).
Abbreviations: O, occurrence score; D, detectability score; RPN, risk priority number; CT, computed tomography; TPS, treatment planning system; MLC, multileaf collimator.

5.4 Discussion

Following the recommendation of TG-100 to perform a risk analysis of new technologies to be implemented in clinical practice, we assessed the risk of our recently developed automated treatment planning tool, the RPA. To our knowledge, this is the first published work assessing the risk of automated treatment planning using FMEA. Additionally, we determined the effect of the specialized QA program on the RPA's failure risk. This work has enabled us to systematically and prospectively

identify the highest-risk steps involved in our automated treatment planning workflow and aided us in developing a QA program specific to the RPA.

While the FMEA presented here is specific to the workflow of the RPA, the lessons gained can be applied broadly to implementations of automated planning. The current analysis has shown us that many of the highest-risk steps, both with and without the QA program (as shown in Table 4 and Table 5), are similar to what might be expected in a manual treatment planning process, including correct identification of the marked isocenter, use of appropriate beam apertures, and use of the correct prescription. However, while the resulting failures may be the same, the causes of these failures may be different in automated planning, in which an algorithm, rather than a human planner, may fail to perform adequately. We have found that many of the highest risk errors in our automated planning workflow were caused by human error. Such errors in an automated plan can be readily detected by the physicist or physician who reviews the plan, as in standard clinical practice. The results of this study emphasize the importance of these plan reviews prior to patient treatment, regardless of whether the plan was generated automatically or using standard manual techniques. In fact, with the advent of automated treatment planning, we need to ensure that we do not develop an overreliance on automation and forego the usual attention to detail in the manual review of treatment plans. In general, automation can improve the safety and consistency of many steps of treatment planning, especially for more objective tasks such as setting prescriptions and creating the beams at the correct isocenter. However, for more subjective tasks, such as designing treatment beams or contouring, the automation techniques and algorithms, like a human planner, have a certain level of skill, so their results should always be scrutinized by qualified staff.

Using automated QA techniques for plan review has been proposed by several groups [60]–[63]. Here, we incorporated such checks for the consistency and reasonability of the treatment planning parameters as part of our QA program. Additionally, we included automatic verification of more subjective tasks, such as designing beam apertures. These automatic checks add additional risk

mitigation without additional workload on staff, which is especially important for resource-constrained settings. However, our QA program only addresses potential failure modes in treatment planning and does not address failures that may occur as plans are transferred to the local treatment planning system or record-and-verify system, although we have developed software to check the integrity of this data transfer.

In addition, our study revealed that automation of treatment planning and QA does not completely remove the risk of human error. For example, the effect of an incorrectly entered prescription is still a high-risk point that can only be detected by a diligent human review. To address this, the third component of our QA program, the specialized manual checks, was designed to draw the reviewer's attention to these important components of the treatment plan. However, there are still potential failures that are not covered by any of our QA program, including the specialized manual checks. Many of these potential failures could be caught by standard QA steps, such as the typical physics and physician plan reviews which could detect if there was a discrepancy of the prescription or an abnormality in the plan. To reinforce the importance of the manual checks as we clinically implement automated planning, training of the users of the RPA will be vital to overcoming the residual higher-risk failure modes. The results of this FMEA will inform the training by educating the users on what are the highest-risk potential failure modes that should be included in their manual checks.

Our analysis of approximately 500 test cases was able to quantify some of the values for the likelihood of occurrence of potential failure modes caused by algorithm error. However, most of the failure modes we identified prospectively through the FMEA did not materialize in this testing. Therefore, the O scores were estimated for these potential failure modes on the basis of the FMEA team's experience and knowledge of the RPA algorithms. In order to further quantify and improve the reliability of the FMEA results, more extensive testing will be necessary in the course of pre- and early implementation. Therefore, the results of this FMEA will be a living document that will be updated as more testing occurs. This is similar to techniques that incorporate data from incident-learning systems

to validate and improve upon FMEA findings. We intend to conduct regular reviews and, if necessary, update the FMEA throughout the testing, clinical implementation, and use of the RPA. By continuing to monitor the RPA's performance, the occurrence of failures, and the ability of the QA program to detect failures, we will collect quantitative data that will improve the validity of the FMEA.

Additionally, before implementing automated planning for other treatment sites or techniques using the RPA, we will reassess the process map and failure modes. For example, in planning volumetric modulated arc therapy for head and neck cancers, contouring is a vital step that has been identified by other groups as a high-risk potential failure mode [43], [53], [55]. Therefore, this FMEA will be updated to consider the risk of failures in automatic contouring, and QA techniques will be added to reduce the risks of these failures.

FMEA is by nature subjective; it incorporates the experience and bias of the team performing it. Thus, the results of this analysis may not necessarily represent the true risk of the system. Still, FMEA is a valuable tool for prospectively identifying potential failure modes and risks, which can help to design QA programs to mitigate the identified risks. Given this limitation, continuous monitoring of the performance of automated planning tools is vital to the safe use of automated treatment planning. Moreover, because the RPA is intended for global use in clinics that may have various levels of resources and follow different practice guidelines, it is important to assess how any deviations from the procedures and workflows assumed to be in place for this FMEA will affect its evaluations of risk. As part of initial clinical implementation, the results of this FMEA will be reassessed with multi-disciplinary teams from each clinic to ensure the results are reflective of each clinics particular risk profile.

Finally, on the basis of the results of this FMEA and the experience we gained through this process, we identified the following 3 key components of the safe deployment of automated treatment planning:

1. **Training.** Training should educate the end users of automated planning systems about the potential failure modes, the impact of these failures on patients, and the need for careful manual review of the plans to prevent these failures.
2. **Manual plan checks.** Physician review of the plans (and contours, where necessary) and physics checks are essential components of automated treatment planning.
3. **Automated QA.** It is important to not only automate the planning, but also to include automated QA steps, as these can substantially mitigate the risks of automated planning.

5.5 Conclusion

We carried out an FMEA to assess the risk involved in the clinical deployment of automated treatment planning. We determined that while automated QA reduces the risks of automated planning, effective training and manual plan checks by radiation oncologists and physicists remain extremely important parts of the deployment process.

Chapter 6 : Automated Plan Quality Verification for Cervical Cancer Radiotherapy

6.1 Introduction

The treatment planning process is a primary source of radiation therapy incidents [52], [64]. For this reason, treatment planning is the focus of many quality assurance (QA) tasks, including physicist and physician reviews of treatment plans. Reviewing the treatment plan is a time-consuming task, and although it is considered one of the most effective forms of QA, it still does not reliably catch all errors that occur in treatment planning [38], [65], [66]. Automating the plan review process would improve the effectiveness of error detection by drawing reviewers' attention to aspects of the treatment plan that are incorrect or are suboptimal and sparing them from reviewing every minute plan detail.

Many radiation therapy groups have already automated various components of the plan review process. Some automation techniques verify the technical accuracy of the plans (e.g., the consistency of the prescription, correct settings for the dose calculation, and completeness of physician contouring) [60], [61], [63], whereas other techniques verify the quality of the treatment plans by comparing them to physician-specified plan quality objectives, estimated achievable dose metrics, or other similar plans [62], [67], [68]. The development of techniques to verify plan quality is more subjective and difficult than the development of techniques to verify the technical accuracy of plans, and most work to date has focused on the quantitative dose metrics involved in assessing plan quality for advanced treatments, such as intensity-modulated radiation therapy. However, many clinics rely heavily on simpler treatment techniques. For example, radiotherapy treatment for patients with cervical cancer can be delivered through the use of a four-field box with beam apertures based on the bony pelvic anatomy. This treatment technique is common in resource-constrained settings in which soft-tissue contouring is not available. In fact, the International Atomic Energy Agency and American Society of Clinical Oncology recommend that resource-constrained clinics use it to deliver radiotherapy to patients with cervical cancer [10], [11].

We have developed a technique to automatically verify the plan quality of three-dimensional radiation treatments for cervical cancer through detection of clinically unacceptable beam apertures for the four-field box treatment. The technique involves comparing the planned beam apertures to a set of independent beam apertures automatically created for the same patient using a secondary technique. This concept is similar to the concept of verifying dose calculation using a secondary, independent algorithm to calculate dose for a given patient's treatment plan.

In this paper, we present our new automated technique for automatically verifying the quality of beam apertures and describe its application in plans created by a recently developed, fully automated treatment planning tool intended for use in resource-constrained settings [37], [44]. Because staff is one of the most limited resources in such clinics [24], any improvement in the efficiency and effectiveness of QA tasks has a substantial clinical impact.

6.2 Methods and Materials

To detect clinically unacceptable beam apertures, this technique utilizes two independent, fully automated techniques to generate beam apertures for a single patient's treatment planning computed tomography (CT) scan. The beam apertures created using the primary technique are intended to be used for treatment, and those created using the secondary technique are to be used to verify the primary treatment apertures. Poor agreement between the two sets of beam apertures indicates that the treatment beam apertures are potentially clinically unacceptable and can be used to alert the physician that the beam aperture may be in need of editing. The primary and secondary techniques, as well as the comparison technique used to detect when the two sets of beam apertures were not in agreement, are described in the following sections. All patient data used in this study were handled in accordance with an approved institutional review board protocol.

6.2.1 Primary Automation Technique for Treatment

The automated planning tool designs four-field box treatments for locally advanced cervical cancer with beam apertures determined based on a patient's bony anatomy. The automation technique and its validation have been previously described elsewhere [37], [44]. In brief, the input to the automated planning tool is a treatment planning CT scan and a plan directive. First, the marked isocenter and body contour are automatically defined, and the patient's bony anatomy is automatically segmented using a technique that employs deformable registration of multiple atlases [39], [46]. Then, the bony anatomy is projected into the beam's-eye-view for each of the four fields, and the beam apertures are determined based on these projections. Next, the beam apertures are converted into jaws and multi-leaf collimator (MLC) positions specific to the type of linear accelerator and MLCs to be used for treatment. Finally, the dose is calculated, and the beam weights are automatically optimized to deliver a homogenous dose to the treated volume.

In a previous study, two radiation oncologists with expertise in radiotherapy for gynecological cancers retrospectively reviewed the beam apertures for 150 treatment planning CT scans. These beam apertures had been generated automatically using the primary technique described above and each was rated as either "clinically acceptable" or "clinically unacceptable" for treatment. Eighty-nine percent of the beam apertures were found to be clinically acceptable [44].

6.2.2 Secondary Automation Technique for Verification

For the secondary verification technique, we used the DeepLabv3+ deep learning architecture [69] to predict the shape of the beam apertures from the digitally reconstructed radiographs (DRRs) for each beam angle. Essentially, this was a two-dimensional segmentation task in which the deep learning model was trained to learn which parts of the DRRs should be included in the beam apertures on the basis of the visible anatomy [70]. The deep learning model was trained and tested on the beam apertures and DRRs from 310 clinically acceptable four-field box treatment plans (255 for training, 55

for testing). Physicians had previously reviewed these beam apertures for clinical acceptability as part of the development and testing of the primary automation technique [44].

During training, 3-fold cross-validation was used to fine-tune the model parameters (learning rate, decay, etc.). After the optimal parameters were identified during cross-validation, the model was trained on the complete training set (255 CT scans) for 35 epochs. A single model was trained to segment all four beams (anteroposterior [AP], posteroanterior [PA], right lateral [RT], and left lateral [LT]) independently; specifically, the model was trained with an equal distribution of each beam's DRRs. To augment the input data, we applied random vertical and horizontal shifts of +/- 5 mm to the input images. Rotational shifts were investigated but were found to be inappropriate due to the nature of the beam aperture segmentations (flat edges defined by jaws with no collimator rotations).

After training, the performance of the deep learning model was assessed by comparing the predicted beam apertures with the ground truth for the 55 test CT scans. The predicted beam apertures were post-processed prior to comparison in order to impose more typical beam aperture qualities, such as straight edges at the field borders to replicate the jaw edges. Averaged across all four post-processed beams for all patients, the mean surface distance (MSD) was 1.6 mm (standard deviation [SD] = 1.0 mm) and the Hausdorff distance (HD) was 6.6 mm (SD = 4.8 mm) (data not shown). The beam apertures used in the secondary verification technique were not post-processed, but were the raw predictions.

6.2.3 Comparison Technique

To detect whether the beam apertures resulting from the primary automation technique were clinically acceptable, we compared them quantitatively to the beam apertures from the secondary technique using the following spatial agreement and overlap metrics [39]:

- MSD: the average Euclidean distance between the points that define the edges of the beam apertures. A larger MSD indicates worse agreement.

- HD: the maximum Euclidean distance between the points that define the edges of the beam apertures. A larger HD indicates worse agreement.
- Dice similarity coefficient (DSC): the overlap of the beam apertures as measured by the ratio of their intersection to their mean area. A DSC value of 1 indicates perfect agreement, whereas 0 indicates no agreement.

We tested this technique on 320 beam apertures (four apertures from each of 80 treatment plans. In a previous study, these were rated by physicians with 228 of the beam apertures rated as clinically acceptable and 92 rated as clinically unacceptable. These plans were part of the cohort used to validate the automated planning tool for cervical cancer [44]. These CT scans were originally acquired for treatment planning with the patient in a supine position.

To determine if the beam apertures for the clinically acceptable plans had better agreement with the secondary technique than those for the clinically unacceptable plans, we compared the values of the agreement metrics using a one-sided t-test. A p-value of 0.05 was used to determine statistical significance. We then used a receiver operating characteristic (ROC) analysis to quantify the sensitivity and specificity with which we could detect unacceptable beam apertures. In this analysis, a true positive result corresponded to beam apertures that were correctly flagged as unacceptable for treatment.

6.3 Results

Figure 1, in which each panel corresponds to a different patient, shows examples of the beam apertures resulting from both automation techniques. Panel A shows a true negative case in which the two automation techniques agreed and the beam apertures were clinically acceptable. Panels B and C, show true positive cases in which the two techniques disagreed and the primary beam apertures were clinically unacceptable. In Panel C, both techniques failed to create clinically acceptable beam

apertures. Because the two automation techniques use independent methodologies, the disparate failures successfully flagged the unacceptable beam apertures.

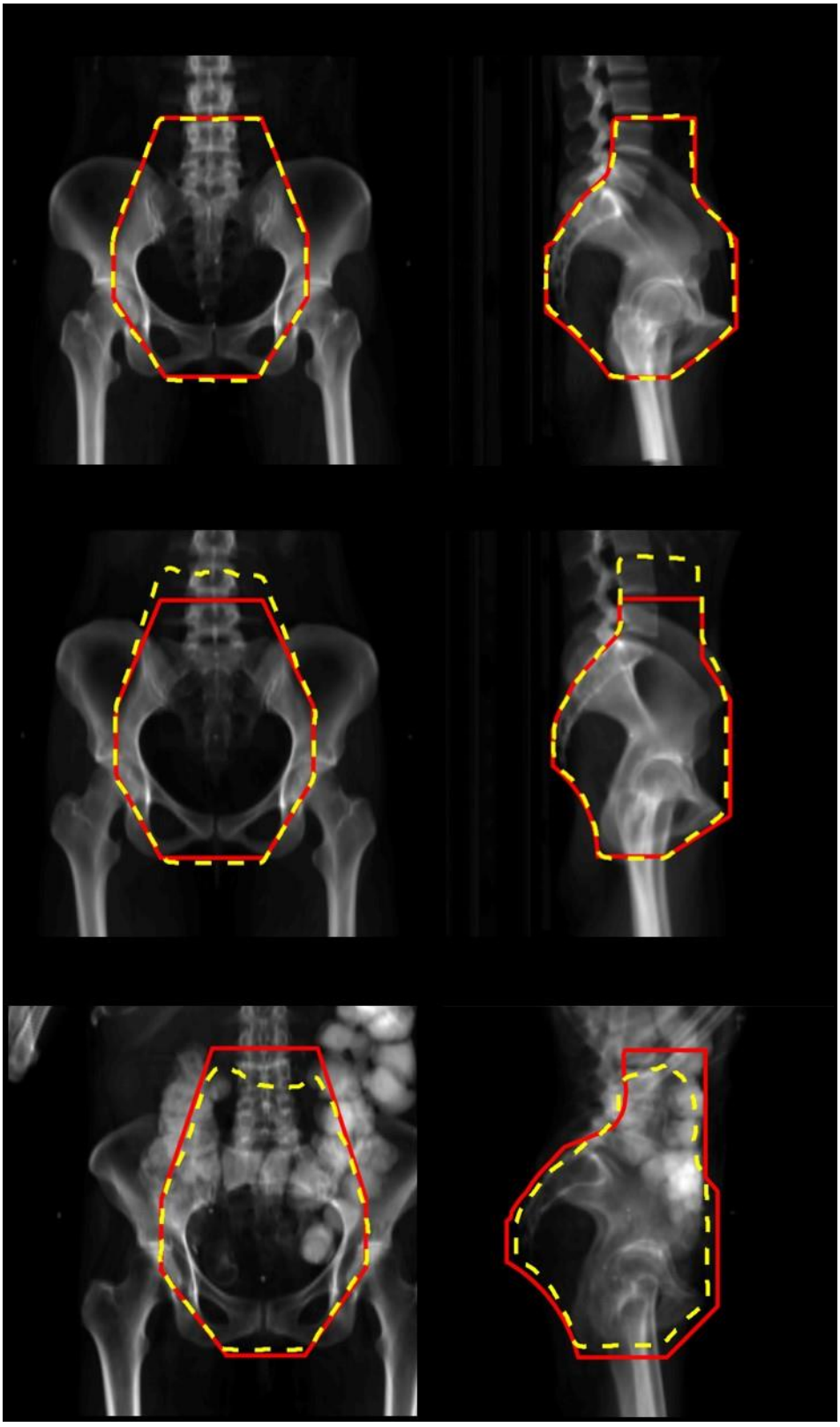

Figure 14. Examples of Beam Apertures Resulting from Both Automation Techniques. The beam apertures are shown for the anteroposterior and right lateral beams (left and right images,

respectively) with results from the primary (red solid line) and secondary (yellow dotted line) techniques. Panel A shows examples of beam apertures that were correctly classified as clinically acceptable by the QA technique (a true negative result). Panels B and C show examples of beam apertures that were correctly classified as clinically unacceptable by the QA technique (true positive results).

6.3.1 Comparison Metrics

The histogram distributions of the MSD, HD, and DSC values for both clinically acceptable and unacceptable beam apertures are shown in Figure 15, Figure 16, and Figure 17, respectively. The means and standard deviations of the comparison metrics tested are embedded in the corresponding figures. For all comparison metrics, the average agreement was significantly better (lower for MSD and HD, higher for DSC) for clinically acceptable plans than for clinically unacceptable plans (all $p < 0.001$). The histograms show only a small overlap between the two sets of beam apertures.

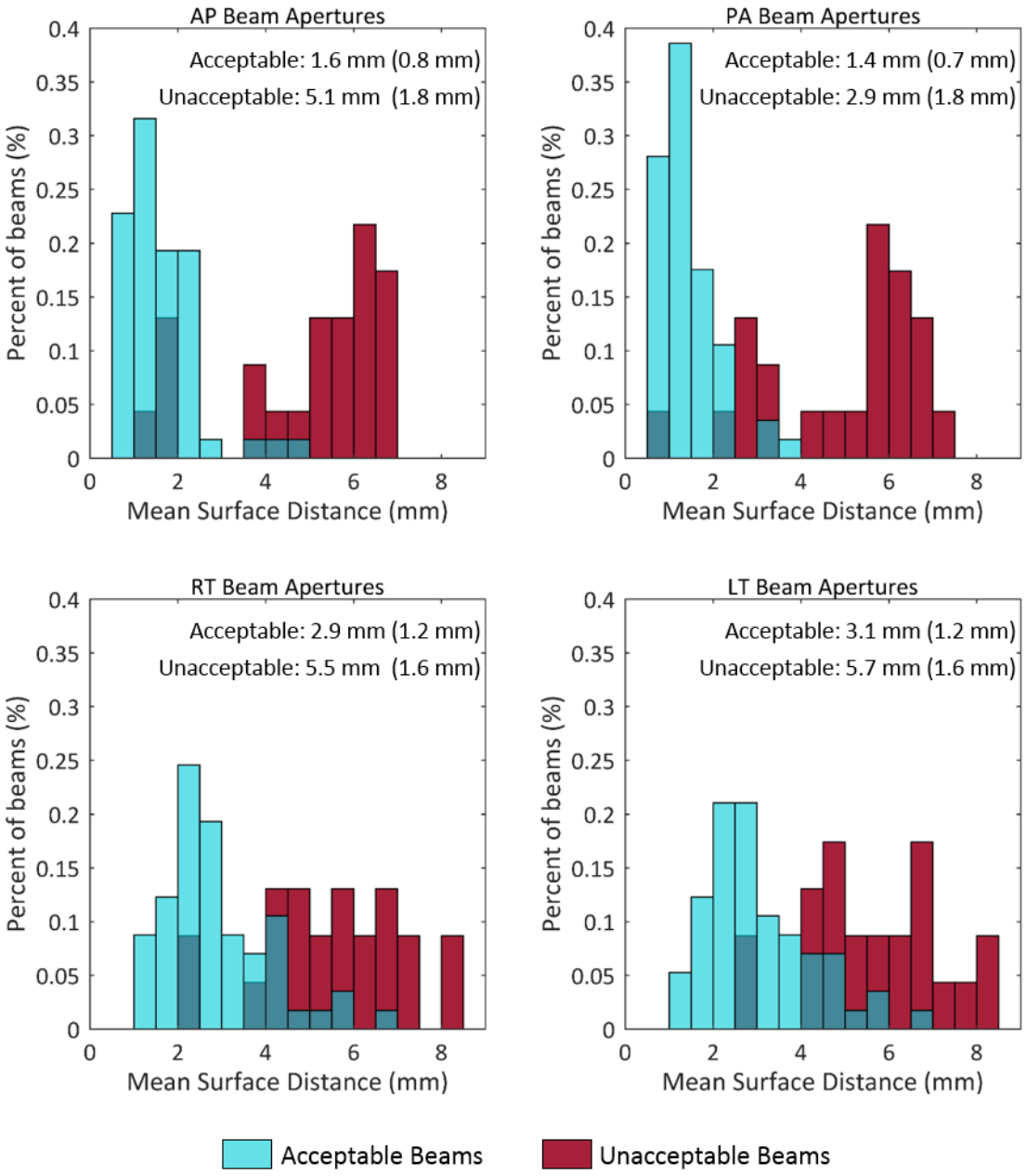

Figure 15. Histogram of the Mean Surface Distance (MSD) Values. Comparison of the beam apertures created by the primary and secondary automation techniques, shown for apertures rated clinically acceptable (blue) or unacceptable (red) by physicians. In each subfigure, the mean is reported and the standard deviation is in parentheses for both the acceptable and unacceptable beams. Lower MSD values indicate better agreement. Abbreviations: AP, anteroposterior; PA, posteroanterior; RT, right lateral; LT, left lateral.

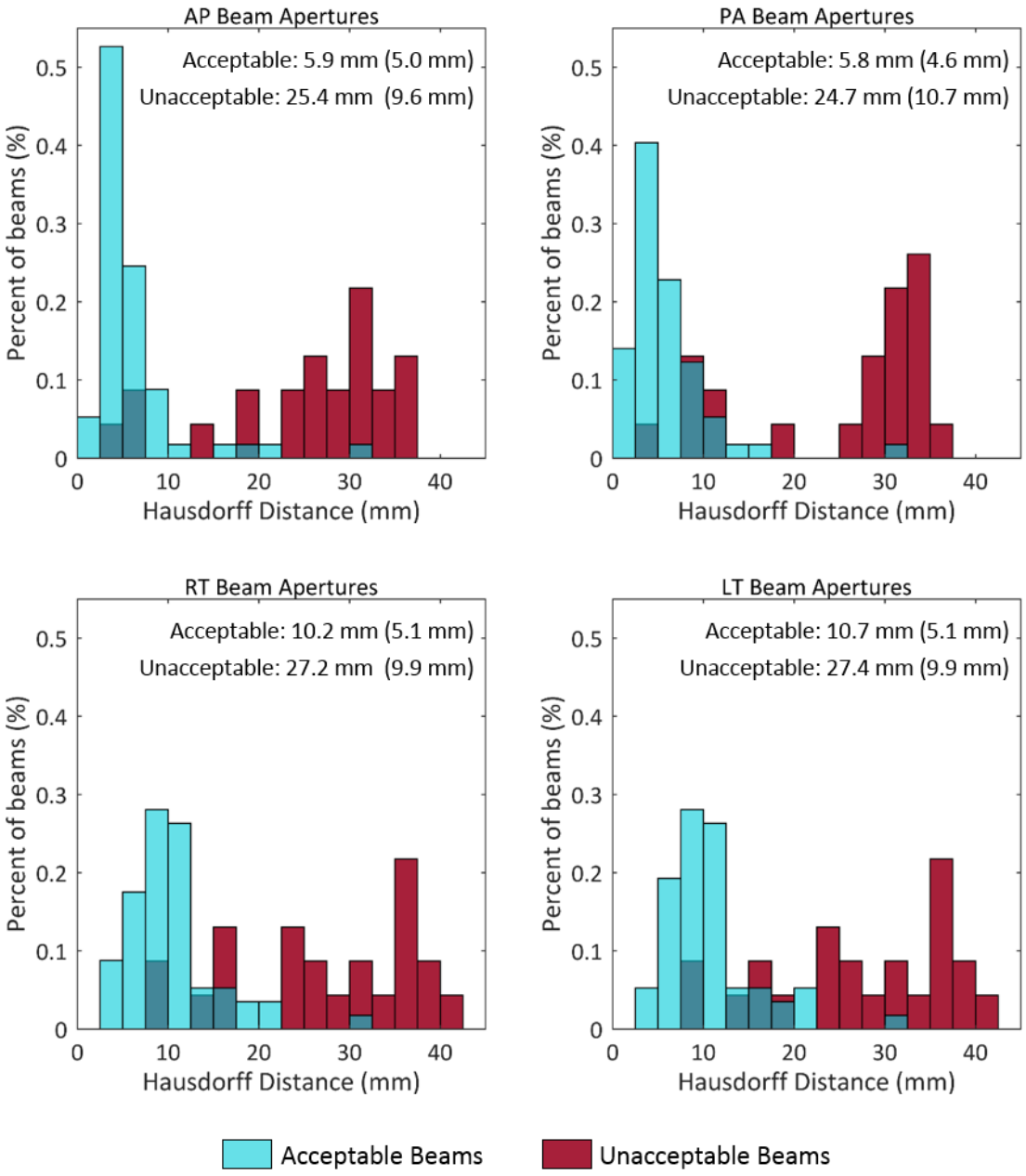

Figure 16. Histogram of the Hausdorff Distance (HD) Values. Comparison of the beam apertures created by the primary and secondary automation techniques, shown for apertures rated clinically acceptable (blue) or unacceptable (red) by physicians. In each subfigure, the mean is reported and the standard deviation is in parentheses for both the acceptable and unacceptable beams. Lower HD values indicate better agreement. Abbreviations: AP, anteroposterior; PA, posteroanterior; RT, right lateral; LT, left lateral.

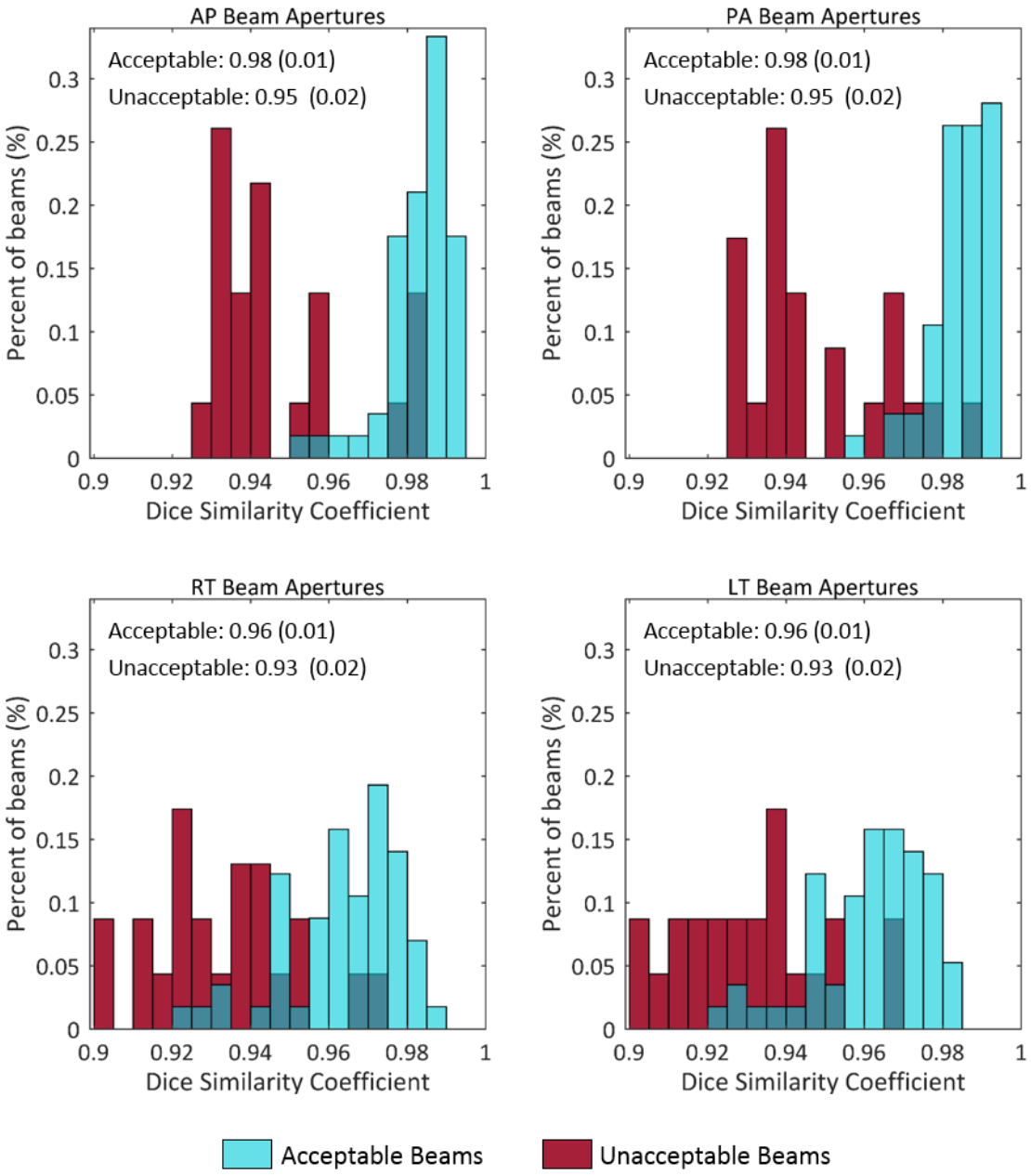

Figure 17. Histogram of the Dice Similarity Coefficient (DSC) Values. Comparison of the beam apertures created by the primary and secondary automation techniques, shown for apertures rated clinically acceptable (blue) or unacceptable (red) by physicians. In each subfigure, the mean is reported and the standard deviation is in parentheses for both the acceptable and unacceptable beams. Higher DSC values indicate better agreement. Abbreviations: AP, anteroposterior; PA, posteroanterior; RT, right lateral; LT, left lateral.

6.3.2 ROC Analysis

Using ROC analysis, we quantified how well each of the comparison metrics could determine if a beam aperture was unacceptable. The resulting ROC curves for all three metrics are shown in Figure 18. The area under the curve (AUC) for the AP, PA, RT, and LT beams are also shown (minimum AUC = 0.89). The high AUC values indicate that these comparison metrics could detect unacceptable beam apertures with good sensitivity and specificity. Overall, the HD metric resulted in the highest AUC for three of the four beam apertures.

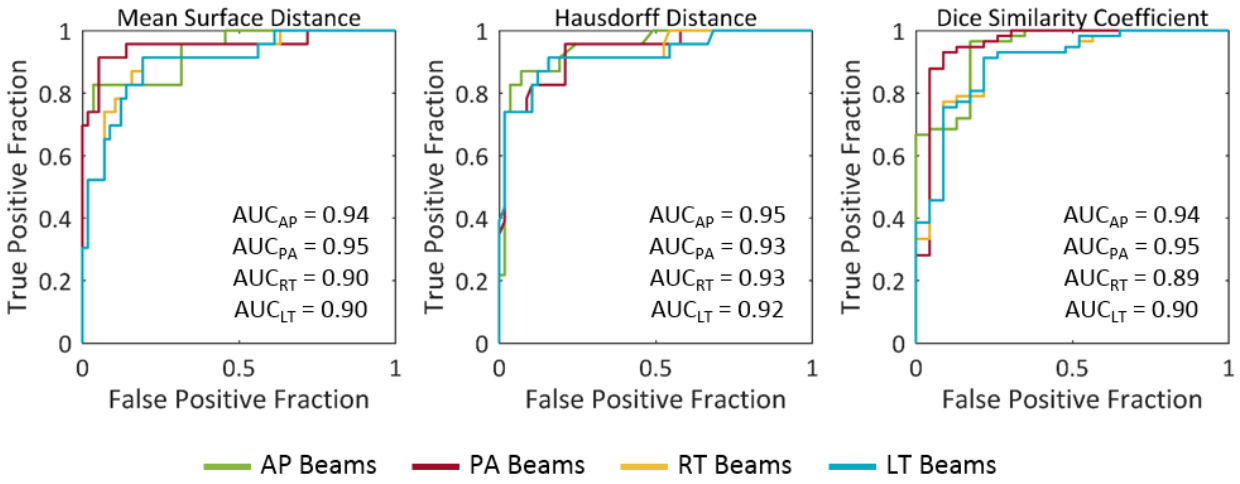

Figure 18. Receiver Operating Characteristic (ROC) Curves for Each Comparison Metric. The area under the curve (AUC) for each metric and beam angle are shown in the corresponding subfigure. Abbreviations: AP, anteroposterior; PA, posteroanterior; RT, right lateral; LT, left lateral.

In order to determine a threshold at which to flag plans that may be deemed unacceptable, we considered two scenarios: one with high sensitivity, giving a true-positive fraction (TPF) of 0.90, and another with high specificity, giving a low false-positive fraction (FPF) of 0.10. Using the HD metric, we determined the threshold values and the corresponding FPFs and TPFs for the beams in the high-specificity and high-sensitivity scenarios, respectively (Table 7). For the high-sensitivity scenario, the corresponding FPF was relatively low, ranging from 0.16 to 0.21. For the high specificity scenario, the corresponding TPF was relatively high, ranging from 0.74 to 0.87.

Table 7. HD Thresholds for Two Verification Scenarios.

		AP Beam	PA Beam	RT Beam	LT Beam
High sensitivity	Threshold	7.0 mm	8.3 mm	14.1 mm	14.1 mm
(TPF = 0.90)	Corresponding FPF	0.19	0.21	0.16	0.16
High specificity	Threshold	11.0 mm	10.4 mm	19.3 mm	19.9 mm
(FPF = 0.10)	Corresponding TPF	0.87	0.78	0.74	0.74

Abbreviations: HD, Hausdorff Distance; TPF, true-positive fraction; FPF, false-positive fraction; AP, anteroposterior; PA, posteroanterior; RT, right lateral; LT, left lateral.

6.4 Discussion

In a previous risk assessment of automated treatment planning using failure modes and effects analysis, it was found that beam aperture creation was one of the high-risk areas subject to failure [71]. Currently, standard practice relies solely on one physician to determine the clinical acceptability of the beam apertures; there is no secondary check by an independent expert. To our knowledge, the QA technique presented in this work is the first technique to automatically verify the clinical acceptability of beam apertures. Because this technique was effective at detecting the clinically unacceptable beam apertures in need of editing prior to patient treatment with high sensitivity and specificity, it could make plan reviews more effective without requiring additional effort from staff. The increases in efficiency and safety wrought by automatic verification could be especially impactful in resource-constrained settings, where there are immense staff shortages.

We implemented this novel automatic beam aperture verification technique as part of an automatic QA program that accompanies a fully automated treatment planning tool for cervical cancer radiotherapy. The goal of the technique is to alert the physician (or other reviewer of plan quality) to potentially unacceptable beam apertures that, when compared with a secondary set of beam

apertures, exceed a threshold for a comparison metric. Based on the results of this study, we will initially use the HD as our comparison metric in our clinically deployed beam aperture verification and will select a threshold that results in high sensitivity to unacceptable beam apertures. However, this may result in an excessive number of false positives (i.e., it may create alerts for plans that a physician would deem acceptable for treatment); in turn, this could cause the alerts to eventually be disregarded. Accordingly, in the initial clinical deployment of the automated planning tool and QA program, we will quantify the effectiveness of the automatic verification of beam apertures and solicit feedback from the physician users regarding the practicality of the thresholds employed for the automatic verification.

By inspecting the plans that were misclassified by the QA technique, we found that most could be attributed to DRRs that were difficult to interpret. Causes of poor quality DRRs included obscured anatomy, poor image quality, or high contrast material in the CT scan. An example of a DRR that was difficult to interpret is shown in Panel C of Figure 1. Poor DRR quality presents a challenge for both the secondary automation technique, which predicts the beam aperture using the DRR, and for physician review of the beam apertures via the DRR.

In addition to having utility within the automated planning tool for cervical cancer, the technique we have presented in this work could be translated to other treatment sites for which beam apertures are designed automatically or manually using the beam's eye view. Our group is currently creating such a tool for use in whole-brain treatments with lateral beams. The methodology presented in this work can be implemented relatively easily in other clinics wanting to verify the quality of their beam apertures and to use their existing, clinically approved treatment plans to train the deep learning model.

With the development and clinical deployment of automated treatment planning, it is natural to include automated processes that aid the physicist and physician in the review of treatment plans prior to treatment. However, although automated plan QAs are a useful means of enhancing the QA process and aiding staff in planning safe and effective radiation treatments, they should only be used for those

purposes and never as a substitute for final manual reviews by staff. Physicians always have the responsibility of ensuring that only high-quality plans are approved for treatment. In this study, even with a threshold set for high sensitivity, 10% of beam apertures rated as clinically unacceptable would not have been flagged by the QA technique.

6.5 Conclusions

In this work, we tested the ability of a novel beam aperture quality verification technique to detect clinically unacceptable beam apertures. We found the technique to be very effective, with AUC values of 0.89-0.95. By comparing the beam apertures with a secondary set that were created using an independent automated technique, we were able to detect clinically unacceptable beam apertures with high sensitivity and specificity. This technique will be deployed as part of a fully automated treatment planning tool for cervical cancer and could be translated to other treatment sites.

Chapter 7 : Discussion and Conclusion

7.1 Project Summary

The primary objective of this work was to automate radiotherapy treatment planning for patients with locally advanced cervical cancer and node-positive, postmastectomy breast cancer. Additionally, we aimed to assess the risk of automated planning and to develop techniques to automatically verify the quality of automatically-generated treatment plans. To achieve these objectives we first developed techniques to automate treatment planning for both cervical and breast cancer radiotherapy in collaboration with hospitals in South Africa and the United States. These automation algorithms were integrated with a commercial treatment planning system and can generate patient-specific treatment plans with no further human input once automated planning is initiated. We then validated our techniques retrospectively by testing these techniques on patient CT scans. Upon physician review of plans, we found they overwhelmingly accepted the cervical cancer plans as-is (90%) and the remaining with minor edits (10%). They also accepted the breast cancer plans as-is (50%) or with only minor edits (50%). Next, we used failure modes and effects analysis (FMEA) to assess the risk involved in the workflow of using this automated planning tool and demonstrated a shift to lower relative risk when a quality assurance (QA) program specific to the automated planning tool was included in the analysis. We then created techniques for automated plan quality verification that were able to successfully detect potential deviations in plan quality in the automatically planned treatments generated for both cervical and breast cancer treatments.

7.2 Specific Aim One

In Specific Aim One, we developed and validated a fully automatic treatment planning tool for four-field box treatments with beam apertures based on bony anatomy, which is the recommended technique for the delivery of external beam radiotherapy for locally advanced cervical cancer in low-

resource clinics [10], [11]. We developed automation algorithms for the design of the treatment beam apertures and optimization of the beam weights to minimize dose heterogeneity.

We have validated the clinical acceptability of the treatments planned by our automated planning tool for cervical cancer in three parts. First, we performed an extensive retrospective validation of the algorithms developed by testing them using 150 planning CTs. The automatically generated body contours and marked isocenters agreed well with their manually defined counterparts. In validating the automatically designed beam apertures, two physicians, one from our institution and one from a South African partner institution, rated 90.7% and 87.5% of plans acceptable for treatment, respectively. The use of automatically optimized beam weights significantly reduced the maximum dose while maintaining coverage within the treated volume. Second, we conducted a successful test of our fully automated planning tool on-site at two South African hospitals using 14 planning CTs from patients previously treated at those institutions. Upon physician review of these plans, 100.0% were rated clinically acceptable. Third, automatically planned treatment beams were clinically deployed at our institution at have been used for 24 patients with cervical cancer by physicians at our institution, with edits as needed, and its use is ongoing.

Overall, we have found that the automation techniques developed in this work are effective for planning patient-specific treatments for locally advanced cervical cancer and may provide a reliable option for resource-constrained clinics. To our knowledge, this is the first work toward fully automated treatment planning for radiotherapy of cervical cancer. As cervical cancer is one of the most prevalent forms of cancer in LMICs, this automated planning tool could help alleviate staff shortages in resource-constrained clinics. In addition to reducing the workload on staff, another benefit of this automated planning tool is that patient treatments could be prepared more quickly since the number of handoffs between physicians and planners would be reduced. This could also improve safety since staff handoffs is a known weakness in the safety of radiotherapy [42], [43].

7.3 Specific Aim Two

In Specific Aim Two, we developed and validated a fully automated treatment planning tool for radiotherapy of the chest wall and ipsilateral lymph nodes using tangential fields matched to an en-face supraclavicular (SCV) field, a standard treatment for node-positive, post mastectomy breast cancer patients. Overall, physicians accepted 100% of patient plans with no or minor changes, with 40% of plans using an alternative match line, which was placed more superior than the clavicular head to reduce the amount of lung exposed in the SCV field. The major automated steps include (1) segmentation of relevant structures (targets, normal tissues, and other planning structures); (2) setup of the beams (tangent fields matched to a supraclavicular field); and (3) optimization of the dose distribution using a mix of high- and low-energy photon beams and field-in-field modulation for the tangent fields.

This automated postmastectomy radiation therapy (PMRT) planning tool was tested on ten CT scans of patients who had previously received radiation to the left chest wall. First we assessed the accuracy of the automated segmentation of the heart and ipsilateral lung. These contours agreed well when compared with manually edited contours. Any differences in contouring did not have a substantial impact on the dose calculated to the organs-at-risk. We then assessed the automatically created PMRT treatment plans quantitatively and found that they met the acceptable dose objectives, including target coverage, maximum plan dose, and dose to organs-at-risk, for all patient plans except for one in which the heart objectives were exceeded. The plans were then reviewed by two physicians who rated the plans on a three-tier scale: use as-is, minor changes, or major changes. Physicians accepted the treatment plans for 50% of the plans as-is and required only minor changes for the remaining 50%.

Given the nuances involved in preparing PMRT and the proximity of the sensitive normal tissues to the targets, there are many trade-offs that must be balanced in order to achieve the desired final plan quality. These tradeoffs include covering the target, reducing dose to organs-at-risk, and minimizing the

maximum dose. As a result, for half of the test patients, physicians would accept the automatically planned treatments, but would attempt minor changes. This is still a positive outcome given that having a complete plan that only requires minor changes would likely still result in a substantial reduction in planning time for these challenging treatments.

To our knowledge, we have developed the first automated planning tool for PMRT and, through retrospective testing, demonstrated its viability for clinical implementation. This tool has the potential to improve the efficiency with which we plan these complex treatments. And given the prevalence of breast cancer in LMICs and the advanced stage at which it is normally diagnosed, this automated PMRT planning tool can potentially reduce the workload on radiotherapy staff.

7.4 Specific Aim Three

In Specific Aim Three, we first assessed the risk of failure of the automated treatment planning tool that we have developed using FMEA and demonstrated that a specialized QA program, which included automatic QA techniques, reduced this risk by improving the detectability of failures. Next, we created and tested plan quality verification techniques for automatically planned treatments generated in this work for both cervical and breast cancer treatments which were both more than 90% effective at detecting potential deviations in plan quality, while having a relatively low false positive rate.

Treatment planning is one of the most error-prone processes in radiation therapy [52], [64]. While the introduction of automated treatment planning techniques has the potential to reduce this risk, no formalized assessment of automated planning existed before now. This work has enabled us to systematically and prospectively identify the highest-risk steps involved in our automated treatment planning workflow and aided us in developing a specialized QA program. We have found that, while automated planning is likely a safe process overall, some residual risks persisted, which were similar to those found in manual treatment planning. Also human error remained a major cause of potential errors.

The risk assessment presented in this work is not only important for the safe clinical implementation of this automated planning tool, it also has generated broader reaching lessons that can inform the safe deployment of other automated planning tools. As the prevalence of automated treatment planning grows, such recommendations for safe deployment will be vital to maintaining safety for patients treated with radiation therapy. Through this risk analysis process, we identified three key aspects of safe deployment of automated planning: (1) user training on potential failure modes; (2) comprehensive manual plan review by physicians and physicists; and (3) automated QA of the treatment plan.

In the FMEA, beam aperture creation for automated planning for cervical cancer was identified as one of the higher-risk potential failure modes. This is because the shape of the beam in four-field box treatments is the main driver of the quality of the treatment plan. As a result, the beam aperture quality was the focus of the plan quality verification technique developed for automated planning for locally advanced cervical cancer. We developed a technique to automatically verify the clinical acceptability of the beam apertures by comparing them with a secondary set of beam apertures developed using an independent technique. If the two sets of beam apertures disagree, this technique alerts the user that the beam apertures may need to be edited prior to treatment. We found that our technique was an effective tool for flagging potentially unacceptable beam apertures during the treatment plan review process.

This plan quality verification technique is effective because of the total independence of the two methodologies used for creating the beam apertures. Both methodologies had instances where they failed to produce acceptable beam apertures, and sometimes this occurred for the same patient. However, the reason for each methodology failing was different, resulting in different appearances for the beam apertures, allowing this QA tool to correct flag the unacceptable beam apertures in these cases.

The plan quality verification technique presented in this work is the first to automatically verify beam apertures. This technique was extremely effective at detecting clinically unacceptable beam apertures that were in need of editing. Additionally, this technique has applications for radiotherapy of other treatment sites, including whole-brain treatments. This automated technique could improve the effectiveness of plan reviews without additional effort from staff, which could improve the efficiency and safety of the treatment planning QA process.

We also have tested a plan quality verification tool for automatically created PMRT plans that alerts the user to any possible deviations in the plan quality. Most of these quality verification tests were based on the plan's dose distribution. In the testing presented in this work, this tool was able to detect 92% of the changes requested by physicians. This methodology will be integrated with the clinical automated planning tool for PMRT, where it could improve the safety of treatment planning by alerting the physician to potential discrepancies, such as high organ-at-risk dose. These alerts could also enable quicker adjustment of the treatment plans, if needed, by automatically drawing attention specifically to the aspect of the treatment plan that may be in need of edit.

7.5 General Discussion

The resulting automated planning tools presented in this work are a consensus based on the expertise of several physicians from the United States and South Africa and was honed through several rounds of testing. For cervical cancer, we conducted preliminary testing on more than 250 patient CT scans. For breast cancer, we conducted preliminary testing on 29 additional patient CT scans prior to the final test reported in this work.

Through this process, we have become aware of variations in radiotherapy treatment techniques for similar patients across institutions and physicians. Many of these variations in clinical practice can likely be accommodated by slight adjustments to the automation algorithms, which have flexibility in the parameters used to generate the resulting treatment plans (e.g. setting the jaws at 1.5 cm versus

2.0 cm beyond the widest extent of the pelvis for cervical cancer plans, or alternative superior border locations for the SCV field of PMRT plans). Such common variations could be accommodated through the use of institution-specific configuration files containing the preferred techniques. However, if the variations in clinical practice are too numerous, there is a risk that it becomes impractical to adapt the algorithms for every new clinic in which the automated planning tool is implemented.

Other variations in treatment plan technique that could easily be accommodated are for patients with different characteristics than those for which these tools were designed, described in detail in Appendix A and Appendix B. For example, the automated planning for cervical cancer can be easily adapted for patients with more extensive nodal disease requiring adjustment of the inferior and/or superior border locations. For automated planning of breast cancer, the same methodologies can be used to automatically plan right-sided PMRT, and even for the treatment of intact breast and treatments without the inclusion of the ipsilateral lymph nodes (SCV and level III axillary). There are certain variations in radiotherapy technique that would require more extensive changes to the automation algorithms. One such example would be the treatment of the internal mammary chain lymph nodes with an en-face electron field, which would require more complex adaptation of the current algorithms and likely the development of new automation algorithms. Prior to the clinical implementation of any adaptation of the current automated planning tools, there should be extensive testing to ensure there are no unexpected outcomes.

The radiotherapy techniques that we selected to automate in this work were based on several factors including, resource-stratified recommendations from international organizations, the current clinical practice of our international partner institutions and other institutions in low- and middle-income countries, and the currently available computational techniques for automating the treatment planning process. As computational techniques progress, it is likely that our ability to automate more advanced techniques will improve. For example, improved automated segmentation of soft tissues in

the pelvis would facilitate automated planning of conformal techniques or even inversely-planned treatment techniques, such as volumetric-modulated arc therapy (VMAT).

The automation of these advanced techniques has the potential to help resource-constrained clinics transition to more state-of-the-art radiotherapy treatment techniques. In order to make the transition to more advanced treatments safely, other developments in clinical practice would need to coincide, including upgraded technology for treatment, new QA processes and equipment, and additional staff expertise. For example, in order to implement VMAT for cervical cancer, clinics will need to have external beam radiotherapy machines capable of delivering such treatments, as well as the equipment necessary to perform the necessary QA tasks. Furthermore, staff will need the expertise to perform these QA tasks and to be able to sufficiently judge the quality of these treatment plans.

7.6 Study Limitations

A limitation of this study was lack of access to patient data from our partner institution. Due to the limitations of sharing data across international institutions, it was difficult to have access to large numbers of patient CT datasets, and especially to the corresponding treatment plans for those patients. As a result, the large cohort of patient CTs used for testing the automated planning tool for cervical cancer were from our own institution, with only a small cohort of 14 CT scans used for testing from two hospitals in South Africa. For automated PMRT planning, we found early in our development that differences in the position of the arms of patients treated at our institution and our South African partner institution was hindering the development process. As a result we were only able to conduct development and testing on data collected from other institutions where the setup was more similar to our partner institution. This severely limited the number of datasets we were able to use for testing.

Finally, retrospective reviews are a vital step prior to the implementation of any clinical tool developed. They are useful in gauging the performance of the tool while minimizing interference in clinical processes. However, they do have several limitations as they do not fully replicate the actual

clinical conditions in which a physician would normally evaluate the quality of a plan. First, the physician does not have the extensive patient-specific information that could alter their opinion of the plan otherwise. Instead, they are instructed to assume the patient is a standard patient (based on the criteria outline in Appendix A and Appendix B, rather than a special case. Retrospective reviews also tend to be conducted in a manner in which the physician is reviewing multiple plans consecutively, which is not typical of the clinic. Prospective reviews in which the physician is comparing the automated treatment plan to a manually planned treatment, or reviewing and editing a manually planned treatment prior to patient treatment would give a better indication of the true performance of the automated planning tools.

7.7 Future Direction

Moving forward, the next step that needs to occur is on-site training and testing with the prospective end users of the tools for automated planning and plan quality verification. These tools are in the process of being converted from a standalone research system to a multi-server clinical implementation that will function as a cloud-based solution. Our partner institutions in southern Africa and the Philippines can access these tools remotely by submitting approved patient CT scans and corresponding approved plan directives through specialized user interfaces developed by our research group [37]. Using these interfaces, they can retrospectively evaluate the performance of these automated planning tools on larger numbers of their own patient cohorts and even compare the results of automated planning to the manually planned treatments. All end-users will need to be trained on the use of the automated planning user interfaces, the appropriate patient characteristics and setup procedures, and any potential failure modes in automated planning.

Once the prospective end-users have completed retrospective validation of the automated planning tools, the next step would be for prospective testing with two main goals: evaluating the

performance under actual clinical conditions and evaluating the reduction in workload and time savings afforded by automated planning.

Prospective testing can be performed in various ways. One such way is a side-by-side test in which the treatment plan is generated simultaneously using both automatic and manual planning. Then the physician would review both plans, ideally blinded to the source of the plan, select their preferred plan, and give feedback on any discrepancies in the two plans [72]. A benefit of this method is that we could also collect data on how long each planning method takes and quantify any potential time-savings and reduction in workload on staff with automated planning. A drawback is that this would require the physician to evaluate two plans, which would involve more of their limited time during the testing period. Another way to prospectively test the automated planning is to present the physician with only the automated plan and then ask them to accept or reject the plan. If accepted, we could then capture any edits made prior to final approval for patient treatment. The benefit of this method is that it does not require the physician to review more than one plan. However, it would not allow us to conduct a head-to-head comparison of the time required for automated and manual planning.

Beyond clinical implementation at our current partner institutions, there are a few other steps that need to occur to facilitate widespread adoption of these automated planning tools. One such step is the adaptation of the current versions of the automated planning tools to accommodate more patients treated for cervical and breast cancer at these institutions. In their current form, the automated tools only apply to a subset of these patients, although we did select the largest subset based on our partner's clinical experience. However, we are still limited by laterality for PMRT and extent of disease for cervical cancer. With some adaptation, these tools can be adapted to have utility for a majority of the patients receiving radiotherapy for cervical and breast cancer.

Another step to enable widespread adoption is adapting these automated tools for other clinical scenarios that are common in resource-constrained clinics. One such example is the lack of access to high energy photons. These automated tools can be adapted and tested for treatments using only low-

energy photon beams, or even for photon beams generated by Co-60 teletherapy machines. Another scenarios could be the use of photon blocks rather than MLCs.

Additionally, prior to clinical deployment, the FMEA results should be revisited and updated based on the final workflow, automation methodologies, and the clinical practice of our partner hospitals where the tool will be implemented. Any findings that result in higher relative risk should be addressed by changes in the workflow, training, or other risk mitigation strategies. Through continued use of our automated planning tools, we can collect quantitative data on the frequency and detectability of any planning failures, and update the values used in our FMEA. This is similar to the use of an incident-learning system to inform the results of an FMEA.

The automation tools and techniques developed in this work are part of a larger project with the long term goal of automating treatment planning for all (or nearly all) radiation treatments. Initially our work (including that presented here and those of others), have spanned the extent of types of radiation treatments, from conventional (four-field boxes for cervical cancer) to inversely planned (VMAT for head-and-neck caner). These techniques build the foundation on which future automated planning techniques can be developed. For example, the techniques developed for four-field box beam aperture creation and beam weight optimization can be applied to other sites using this “2.5-D” technique, such as whole brain treatments using lateral beams or rectal cancer treatments using a three-field belly board technique. The techniques developed to determine optimal gantry angles for left-sided, postmastectomy radiotherapy can be used for determine treatment field borders for all types of breast tangent plans, any treatment using oblique gantry angles, such as the en-face SCV field, and even for setting up electron beams to treat the internal mammary chain lymph nodes. Additionally, the techniques developed for automating the optimization of the dose distribution using mixed energy photon beams and field-in-field segments could be applied to any radiation treatment using field-in-field style modulation to achieve more homogenous dose distributions.

7.8 Conclusion

For automated planning for locally advanced cervical cancer, radiation oncologists accepted as-is 89.6% of treatments plans reviewed in the two test cohorts, with three expert physicians evaluating different subsets of the plans. The remaining 10% required only minor changes in the beam apertures. For node-positive, postmastectomy breast cancer, 100% of the plans were accepted with no or only minor changes and are therefore considered clinically viable. Therefore, our hypothesis that 90% of the treatment plans created with our automated planning techniques would be rated as clinically acceptable by radiation oncologists was proven in this work. Furthermore, the techniques we developed for the automated detection of poor quality beam apertures for cervical cancer radiotherapy was able to detect 90% of the unacceptable beam apertures with good specificity. The automated detection of deviations in plan quality for PMRT treatments detected 92% of the changes requested by radiation oncologists. Therefore, our hypothesis that our automated quality verification techniques would detect 90% of unacceptable plans was proven in this work.

Appendix A

Statement of appropriate use of the automated planning tool for cervical cancer

The automated planning tool developed in this work for radiotherapy of locally advanced cervical cancer has been designed and tested for a certain disease extent, treatment setup, treatment technique. Below are a list of criteria relating to the appropriate use of the current version of the automated planning tool presented in this work. Applicability of this tool to other similar conditions are feasible but not yet vetted.

Patient Disease Extent

- Locally advanced cervical cancer
- Disease limited to the upper two-thirds of the vagina (ie, not involving the distal vagina)
- Disease only involving the pelvic lymph node (eg, not involving the paraaortic lymph nodes).

Patient Treatment Setup

- Head first, supine
- Marked and treatment isocenter indicated with a three-point external fiducial setup
- No additional high density objects (eg, wires or fiducials) should be placed on the patient's body surface
- No intravenous or oral contrast should be used in the planning CT scan

Treatment Technique

- Four-field box treatment technique with beam apertures based on bony anatomy
- Delivered using a linear accelerator with high energy photons
- Treatment fields defined by asymmetrical jaws and MLCs (ie, no photon blocks)

- Dose normalized 100% to the marked/treatment isocenter

AP and PA Beam Aperture Description

Table 8. The AP and PA beams are not mirror images of each other. The location of each jaw or block described below is indicated by the Roman numeral and corresponds to the location shown in panel A of Figure 19.

Superior-right block (i)	From transverse process, straight line to level of SI (sacroiliac) joint (where sacrum touches the pelvis, ie, the top of the pelvic inlet)
Superior jaw (ii)	L3/L4 (ie, top of L4)
Superior-left block (iii)	From transverse process, straight line to level of SI (sacroiliac) joint (where sacrum touches the pelvis, ie, the top of the pelvic inlet)
Left jaw (iv)	2.0 cm from widest part of pelvic inlet
Inferior-left block (v)	From where the top of the femoral head crosses the jaw, a straight line to skirt past the obturator foramen. Make sure to not block the medial/superior aspect of the femoral head
Inferior jaw (vi)	Bottom of the pelvis
Inferior-right block (vii)	From where the top of the femoral head crosses the jaw, a straight line to skirt past the obturator foramen. Make sure to not block the medial/superior aspect of the femoral head
Right jaw (viii)	2.0 cm from widest part of pelvic inlet

Lateral Beam Aperture Description

Table 9. The right and left lateral beams are mirror images of each other. The location of each jaw or block described below is indicated by the Roman numeral and corresponds to the location shown in panel B of Figure 19.

Superior-posterior block (i)	Follow curve of vertebral bodies
Superior jaw (ii)	L3/L4 (ie, top of L4)
Superior-anterior block (iii)	Start 3.5 cm out anterior to the vertebral bodies then straight inferior till bottom of L5. From there, diagonal out till anterior jaw, the slope of which is parallel to the slope from S1/L5 to top of pubic symphysis
Anterior jaw (iv)	Anterior to the pubic symphysis by 0.5 cm. However, if this puts the jaw <3.5 cm anterior to the vertebral bodies, then set at 3.5 cm anterior to the vertebral bodies.
Inferior-anterior block (v)	Following just inferior to the pubic symphysis
Inferior jaw (vi)	Bottom of the pelvis
Inferior-posterior block (vii)	Start at posterior jaw at level of coccyx. Follow the curve of the coccyx and draw a curved line toward where the pelvis crosses the inferior jaw
Posterior jaw (viii)	Posterior to the sacrum

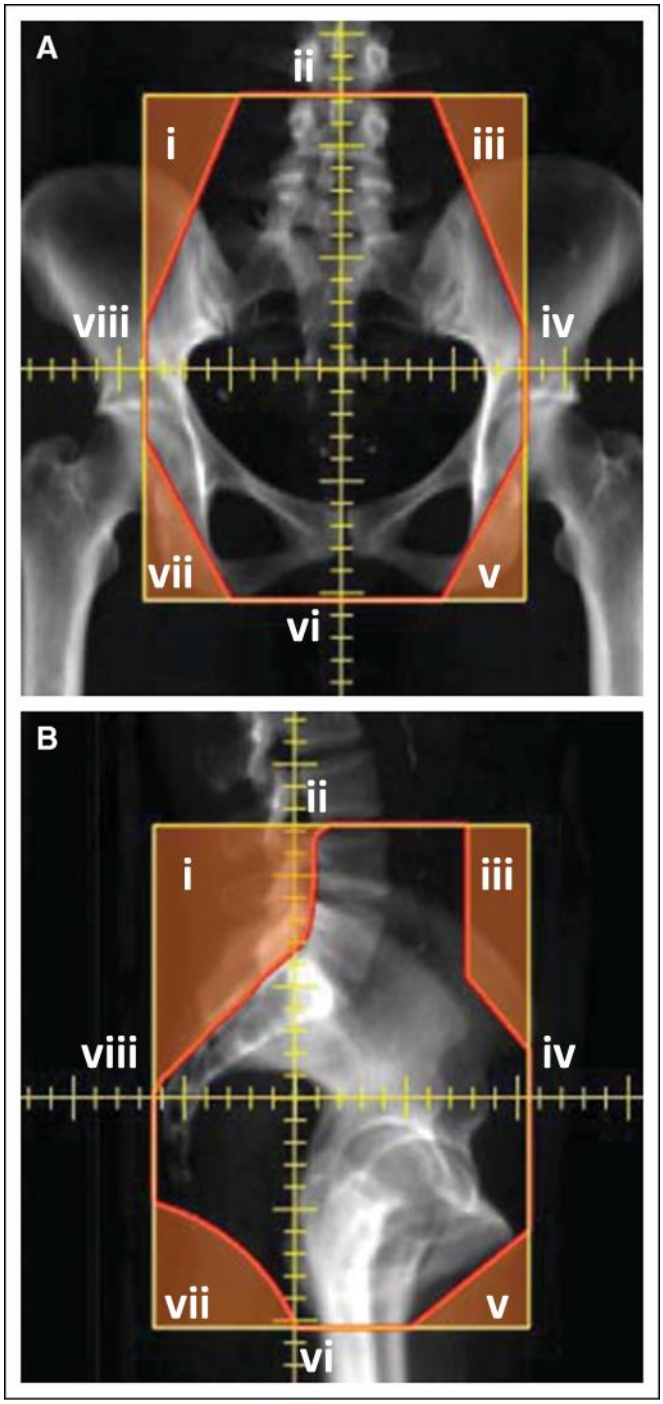

Figure 19. Figure adapted from Kisling et al [44]. Beam's eye view of the (A) anteroposterior (AP) and (B) right lateral beam angles. The Roman numerals indicate the location of a jaw or block which corresponds to the descriptions above in Table 8 and Table 9.

Appendix B

Statement of appropriate use of the automated planning tool for breast cancer

The automated planning tool developed in this work for radiotherapy of node-positive, postmastectomy breast cancer has been designed and tested for a certain disease extent, treatment setup, treatment technique. Below are a list of criteria relating to the appropriate use of the current version of the automated planning tool presented in this work. Applicability of this tool to other similar conditions are feasible but not yet vetted.

Patient Disease Extent

- Node-positive, postmastectomy breast cancer requiring treatment of the chest wall, supraclavicular (SCV) lymph nodes, and level III axillary lymph nodes
- For treatment of left-sided disease only
- Not for treatment of the internal mammary chain lymph nodes
- Not for patients with skin involvement

Patient Treatment Setup

- Head first, supine
- Both arms raised over the patient's head
- Patient scanned while freely breathing (ie, no deep inspiration breathhold)
- Patient positioned on an angled breast board
- Marked isocenter indicated with a three-point external fiducial setup
- No intravenous or oral contrast should be used in the planning CT scan

Treatment Technique

- Monoisocentric medial and lateral tangential fields matched to an oblique, en-face supraclavicular field
- Designed for a hypofractionated treatment (40.05 Gy in 15 fractions)
- Delivered using a linear accelerator with both low and high energy photons (specifically 6 MV and 18 MV)
- Treatment fields defined by asymmetrical jaws and MLCs (ie, no photon blocks)
- Mixed energies allowed for all fields
- Up to two field-in-field segments per medial and lateral tangential beam
- Match line between tangential and SCV fields is placed at the inferior extent of the clavicle
    - If the tangent fields are too long (>20 cm maximum field size for a half-beam blocked field for Varian C-arm machines, such as 2100 series or TrueBeam), the match line if moved more inferior
    - If there is insufficient open field inferior to the humeral head (<2 cm), the match line if moved more inferior
    - If there is too much lung in the SCV field (>4 cm from the inferior border of field to the apex of the lung in the beam's eye view), the match line is moved more superior (as long as the previous two points allow this move)

Tangential Field Description

- A beam's eye view of an example of the medial tangential field is shown in Figure 20. The location of each border described below is indicated by a Roman numeral that corresponds to the location shown in the figure.
- The collimator angle is 0 degrees
- The superior border (i) is at the match line location

- The inferior border (ii) is 1 cm below the most inferior extent of the breast tissue (which is manually identified)
- The anterior border (iii) allows for 2 cm of flash
- The posterior border (iv) is defined by the MLCs

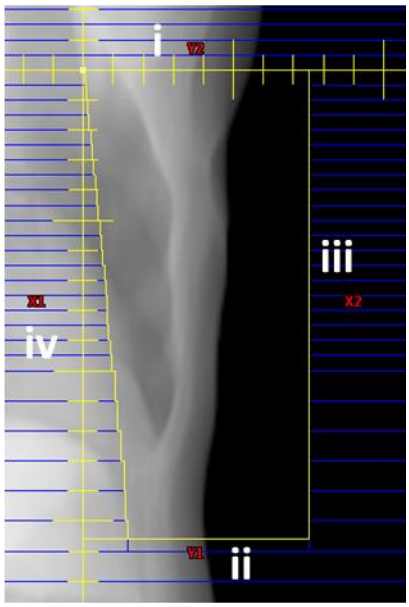

Figure 20. Beam's eye view of the medial tangential field. The Roman numerals indicate the location of a border that corresponds to the descriptions of the field.

SCV Field Description

- A beam's eye view of an example of the SCV field is shown in Figure 21. The location of each border described below is indicated by a Roman numeral that corresponds to the location shown in the figure.
- The superior border (i) is at the cricoid
- The inferior border (ii) is at the match line location
- The medial border (iii) shields the esophagus and spinal cord
- At the lateral border of the field (iv), the humeral head is shielded

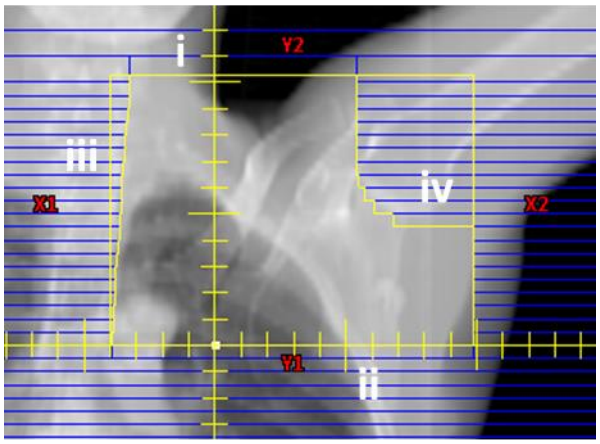

Figure 21. Beam's eye view of the SCV field. The Roman numerals indicate the location of a border that corresponds to the descriptions of the field.

Dosimetric Goals

Table 10. These objectives were determined using several sources, including objectives from a hypofractionated PMRT clinical trial (unpublished protocol) [49], recommendations from The Royal College of Radiologists [45], and clinical objectives from our partner hospitals.

Region of Interest	Dose Metric	Preferred Objective	Acceptable Limit
Heart	Mean Dose	< 4 Gy	< 6 Gy
	Volume > 25 Gy	< 7%	< 10%
Ipsilateral (left) lung	Volume > 17 Gy	< 35%	< 40%

Appendix C

Full results of failure modes and effects (FMEA) analysis

Table 11. The following table contains the potential failure mode and potential cause of failure for each process step assessed in the Radiation Planning Assistant (RPA) workflow (depicted in Figure 11 in Chapter 5). The FMEA scores are shown for the likelihood of occurrence (O), severity (S), likelihood of going undetected (D), and the overall risk priority number (RPN) both without and with the specialized quality assurance (QA) program.

				Without QA program				With QA program			
Process Step	Process Step Description	Potential Failure Mode	Potential Cause of Failure	O	S	D	RPN	O	S	D	RPN
1	**CT simulation**										
1.01	Position patient	Incorrect orientation	Human error	3	9	2	54	3	9	2	54
1.01	Position patient	Incorrect orientation	Standard technique varies from RPA protocol	6	9	3	162	6	9	1	54
1.01	Position patient	Incorrect orientation	Intentional non-standard technique	4	9	3	108	4	9	2	72
1.01	Position patient	Unexpected immobilization (that affects CT data)	Human error	3	5	8	120	3	5	6	90
1.01	Position patient	Unexpected immobilization (that affects CT data)	New immobilization device	2	5	7	70	2	5	6	60
1.01	Position patient	Unexpected immobilization (that affects CT data)	Standard technique varies from RPA protocol	3	5	8	120	3	5	7	105
1.01	Position patient	Unexpected immobilization (that affects CT data)	Intentional non-standard technique	4	5	8	160	4	5	6	120
1.01	Position patient	Inappropriate positioning (eg arm in-field)	Human error	6	4	7	168	6	4	7	168
1.01	Position patient	Inappropriate positioning (eg arm in-field)	Intentional non-standard technique	4	4	8	128	4	4	8	128
1.02	Place external fiducials	Inappropriate type of external fiducials	Human error	2	4	8	64	2	4	4	32

1.02	Place external fiducials	Inappropriate type of external fiducials	Standard technique varies from RPA protocol	2	4	8	64	2	4	4	32
1.02	Place external fiducials	External fiducials exceed planar limitations	Human error	3	3	1	9	3	3	1	9
1.02	Place external fiducials	External fiducials exceed planar limitations	Patient moves	4	3	1	12	4	3	1	12
1.02	Place external fiducials	<3 external fiducials placed	Human error	1	3	1	3	1	3	1	3
1.02	Place external fiducials	<3 external fiducials placed	Standard technique varies from RPA protocol	8	3	1	24	8	3	1	24
1.02	Place external fiducials	<3 external fiducials placed	External fiducials fall off	2	3	1	6	2	3	1	6
1.02	Place external fiducials	Out of range of treatment/wrong location	Human error	3	6	4	72	3	6	2	36
1.03	Enter patient information	Incorrect ID or name	Human error	2	1	1	2	2	1	1	2
1.03	Enter patient information	Incorrect ID or name	Human error	1	10	9	90	1	10	9	90
1.04	Select CT protocol and execute	Orientation entered incorrectly	Human error	5	2	1	10	5	2	1	10
1.04	Select CT protocol and execute	Orientation entered incorrectly	Standard technique varies from RPA protocol	5	2	1	10	5	2	1	10
1.04	Select CT protocol and execute	Orientation entered incorrectly	Incorrect CT protocol selected (eg HN protocol)	5	2	1	10	5	2	1	10
1.04	Select CT protocol and execute	Orientation entered incorrectly	Standard CT protocol is inappropriate	5	2	1	10	5	2	1	10
1.04	Select CT protocol and execute	Too large slice spacing	Human error	2	8	5	80	2	8	3	48
1.04	Select CT protocol and execute	Too large slice spacing	Standard technique varies from RPA protocol	2	8	5	80	2	8	3	48
1.04	Select CT protocol and execute	Too large slice spacing	Incorrect CT protocol selected (eg HN protocol)	2	8	5	80	2	8	3	48
1.04	Select CT protocol and execute	Too large slice spacing	Standard CT protocol is inappropriate	2	8	5	80	2	8	3	48

1.04	Select CT protocol and execute	FOV is too small	Human error	5	8	7	280	5	8	6	240
1.04	Select CT protocol and execute	FOV is too small	Patient is too large	5	8	7	280	5	8	6	240
1.04	Select CT protocol and execute	FOV is too small	Incorrect CT protocol selected (eg HN protocol)	3	8	7	168	3	8	6	144
1.04	Select CT protocol and execute	FOV is too small	Standard CT protocol is inappropriate	2	8	7	112	2	8	6	96
1.04	Select CT protocol and execute	Insufficient scan length	Human error	4	3	1	12	4	3	1	12
1.04	Select CT protocol and execute	kV/mA inappropriate (affects image quality)	Human error	1	5	5	25	1	5	5	25
1.04	Select CT protocol and execute	kV/mA inappropriate (affects image quality)	Incorrect CT protocol selected (eg HN protocol)	1	5	5	25	1	5	5	25
1.04	Select CT protocol and execute	kV/mA inappropriate (affects image quality)	Standard CT protocol is inappropriate	1	5	5	25	1	5	5	25
1.04	Select CT protocol and execute	Used contrast (IV or oral)	Human error	3	5	5	75	3	5	3	45
1.04	Select CT protocol and execute	Used contrast (IV or oral)	Standard technique varies from RPA protocol	8	5	5	200	8	5	3	120
1.04	Select CT protocol and execute	Used contrast (IV or oral)	Intentional non-standard technique	4	5	5	100	4	5	3	60
1.04	Select CT protocol and execute	Scan incomplete/failure	Power failure	1	3	1	3	1	3	1	3
1.04	Select CT protocol and execute	Scan incomplete/failure	Patient issues	1	3	1	3	1	3	1	3
1.05	Transfer CT to RPA control center	Incorrect CT sent	Human error	4	4	8	128	4	4	8	128
1.05	Transfer CT to RPA control center	Data corrupted/incomplete	Network error	1	7	5	35	1	7	3	21
1.05	Transfer CT to RPA control center	Data corrupted/incomplete	Human error (too quick)	6	7	5	210	6	7	3	126

1.05	Transfer CT to RPA control center	Not completed	Network error	1	2	1	2	1	2	1	2
1.05	Transfer CT to RPA control center	Not completed	Human error	3	2	1	6	3	2	1	6
1.06	Approve CT in RPA control center	Not approved	Human error	8	2	1	16	8	2	1	16
1.06	Approve CT in RPA control center	Approve incorrect CT (same patient, if >1 CT)	Human error	3	4	8	96	3	4	8	96
1.07	Transfer CT from RPA control center to RPA planning module	Data corrupted	Network error	1	7	5	35	1	7	3	21
1.07	Transfer CT from RPA control center to RPA planning module	Not completed	Network error	1	2	1	2	1	2	1	2
2	**Plan directive**										
2.01	Enter patient information	Incorrect ID or name	Human error	2	10	1	20	2	10	1	20
2.01	Enter patient information	Incorrect ID or name	Human error	2	1	1	2	2	1	1	2
2.01	Enter patient information	Incomplete ID or name	Human error	2	1	1	2	2	1	1	2
2.01	Enter patient information	Incompatible digits	Algorithm error	2	1	1	2	2	1	1	2
2.02	Enter treatment site	Incorrect treatment site	Human error	3	1	1	3	3	1	1	3
2.02	Enter treatment site	Not completed	Human error	2	1	1	2	2	1	1	2
2.03	Questions about patient appropriateness	Not completed	Human error	2	1	1	2	2	1	1	2
2.03	Questions about patient appropriateness	Completed incorrectly	Human error	2	9	7	126	2	9	7	126
2.03	Questions about patient appropriateness	Completed incorrectly	Human error	4	9	5	180	4	9	5	180
2.04	Enter prescription	Incorrect dose (not changed from default)	Human error	4	9	7	252	4	9	7	252
2.04	Enter prescription	Incorrect dose (changed from default incorrectly)	Human error	3	9	6	162	3	9	6	162

2.05	Select treatment machine	Incorrect	Human error	6	3	6	108	6	3	6	108
2.05	Select treatment machine	Not selected	Human error	2	1	1	2	2	1	1	2
2.06	Approve plan directive	Not completed	Human error	2	1	1	2	2	1	1	2
2.06	Approve plan directive	Approved by staff without correct rights	Shared login/incorrect rights	4	9	5	180	4	9	5	180
2.07	Data transfer from RPA control center to RPA planning module	Data corrupted	Network error	2	1	1	2	2	1	1	2
2.07	Data transfer from RPA control center to RPA planning module	Not completed	Network error	2	1	1	2	2	1	1	2
3	**RPA plan creation**										
3.01	Isocenter position	Incorrectly identified	Other external fiducials	7	9	8	504	7	9	5	315
3.01	Isocenter position	Incorrectly identified	Out of range of CT scan	3	9	8	216	3	9	4	108
3.01	Isocenter position	Incorrectly identified	Algorithm error	4	9	8	288	4	9	4	144
3.01	Isocenter position	Not identified	Out of range of CT scan	3	2	1	6	3	2	1	6
3.01	Isocenter position	Not identified	Algorithm error	2	2	1	4	2	2	1	4
3.02	Body contour	Inadequately contoured	Inappropriate immobilization	5	5	8	200	5	5	6	150
3.02	Body contour	Inadequately contoured	Poor image quality	5	4	8	160	5	4	5	100
3.02	Body contour	Inadequately contoured	Unexpected item in CT (eg arm, cell phone, etc)	6	4	7	168	6	4	5	120
3.02	Body contour	Inadequately contoured	Unexpected couch (eg low density material)	3	4	8	96	3	4	6	72
3.02	Body contour	Inadequately contoured	Algorithm error	4	5	8	160	4	5	4	80
3.02	Body contour	Not created	Algorithm error	2	2	1	4	2	2	1	4
3.03	All 4 beams created	Not at isocenter	Algorithm error before aperture creation	1	9	8	72	1	9	6	54

3.03	All 4 beams created	Not at isocenter	Algorithm error after aperture creation	1	9	5	45	1	9	2	18
3.03	All 4 beams created	<4 beams created	Algorithm error	1	2	1	2	1	2	1	2
3.03	All 4 beams created	Wrong machine	Algorithm error	1	4	5	20	1	4	5	20
3.04	Gantry angles	Incorrect	Algorithm error/corrupted template	1	5	8	40	1	5	6	30
3.05	Collimator angle = 0	Incorrect	Algorithm error/corrupted template	1	5	5	25	1	5	3	15
3.06	Couch angle = 0	Incorrect	Algorithm error/corrupted template	1	5	7	35	1	5	5	25
3.07	Jaw positions	Inappropriate position	Poor image quality	5	5	5	125	5	5	4	100
3.07	Jaw positions	Inappropriate position	Algorithm error	10	7	6	420	10	7	4	280
3.07	Jaw positions	Inappropriate position	Presence of high contrast material	8	5	5	200	8	5	4	160
3.07	Jaw positions	Inappropriate position	Unexpected item in CT (eg arm, cell phone, etc)	6	4	7	168	6	4	6	144
3.07	Jaw positions	Inappropriate position	Isocenter at extreme location	6	6	5	180	6	6	3	108
3.07	Jaw positions	Exceeds delivery limits	Algorithm error	7	2	1	14	7	2	1	14
3.08	MLC positions	Inappropriate position	Poor image quality	5	5	5	125	5	5	4	100
3.08	MLC positions	Inappropriate position	Algorithm error	10	7	6	420	10	7	4	280
3.08	MLC positions	Inappropriate position	Presence of high contrast material	8	5	5	200	8	5	4	160
3.08	MLC positions	Inappropriate position	Unexpected item in CT (eg arm, cell phone, etc)	6	4	7	168	6	4	6	144
3.08	MLC positions	Inappropriate position	Isocenter at extreme location	6	6	5	180	6	6	3	108
3.08	MLC positions	Not present	Algorithm error	1	9	3	27	1	9	2	18
3.08	MLC positions	Exceeds delivery limits	Algorithm error	7	2	1	14	7	2	1	14
3.09	Accessories set	Incorrect	Algorithm error/corrupted template	1	7	4	28	1	7	2	14

3.10	Prescription set	Incorrect (doesn't match plan directive	Algorithm error	1	10	7	70	1	10	5	50
3.10	Prescription set	Incorrect normalization	Algorithm error/corrupted template	1	10	7	70	1	10	5	50
3.11	Energy set	Incorrect energy	Algorithm error/corrupted template	1	4	8	32	1	4	6	24
3.12	Dose distribution	Calculation point is inappopriate	Located in high or low CT number	10	4	6	240	10	4	3	120
3.12	Dose distribution	Incorrect but approved calculation algorithm	Algorithm error	1	1	10	10	1	1	9	9
3.12	Dose distribution	No heterogeneity correction	Algorithm error	1	4	8	32	1	4	6	24
3.12	Dose distribution	Dose grid inappropriate (size/location)	Algorithm error	1	4	6	24	1	4	4	16
3.12	Dose distribution	Calculation point is not at isocenter	Algorithm error	1	10	7	70	1	10	5	50
3.13	Field weights set	Inappropriate	Algorithm error	5	4	8	160	5	4	7	140
3.14	Plan documentation	Data corrupted	Algorithm error	3	10	4	120	3	10	4	120
3.14	Plan documentation	Not created	Algorithm error	1	2	1	2	1	2	1	2
4	**Plan approval**										
4.01	Physician plan review	Not completed	Human error	2	3	1	6	2	3	1	6
4.01	Physician plan review	No comprehensive review	Human error	3	10	10	300	3	10	10	300
4.02	Data transfer from RPA to local TPS	Not completed	Network error	3	3	1	9	3	3	1	9
4.02	Data transfer from RPA to local TPS	Data corrupted	Network error	2	10	5	100	2	10	3	60

Appendix D

Quality assurance (QA) for the location of the calculation point in four-field box treatment plans for cervical cancer radiotherapy

In the automated planning tool for cervical cancer radiotherapy that was developed in this work, the calculation point is set at the location of the marked isocenter, per the local practice of our partner hospitals located in low- and middle-income countries. This calculation point is used as the prescription point, where the dose is set to be normalized to 100% of the prescription, and as the reference point for the relative field weights.

In our testing of automated planning for cervical cancer radiotherapy, we discovered that sometimes the location of the calculation point was not suitable to be used for dose calculation. There were two main causes: (1) the calculation point was very close to the edge of the treatment field and (2) the calculation point was located in high density material, such as bone, or low density material, such as gas within the bowels.

In order to avoid creating treatment plans with an inadequate location for the calculation point, we developed an automatic QA test of the location of the calculation point. The first QA test was based on the proximity of the calculation point to the edge of the treatment field and creates a flag if the proximity is less than 2 cm. In a test of 366 patient plans, 1.1% were found to have the calculation point located within 2 cm of the edge of at least one beam aperture.

The second QA test is based on the distribution of the CT numbers in the region around the calculation point. This QA test evaluates a 1-cm radius sphere around the calculation point and searches for high-density by assessing the 90^{th} percentile of the CT numbers within this sphere (ie, the most dense values) and searches for low-density by assessing the 10th percentile of the CT numbers within this sphere (ie, the least dense values). In a test of 366 patient plans, using an upper threshold of 200 and a lower threshold of -200 created flags for 5.7% and 9.6% of patient plans, respectively. The distributions of the CT number at the 90^{th} and 10^{th} percentiles for these patients are shown in Figure 22

below. The red line indicates the threshold value used in this test. Examples of patient plans that were flagged for high and low density are shown in Figure 23.

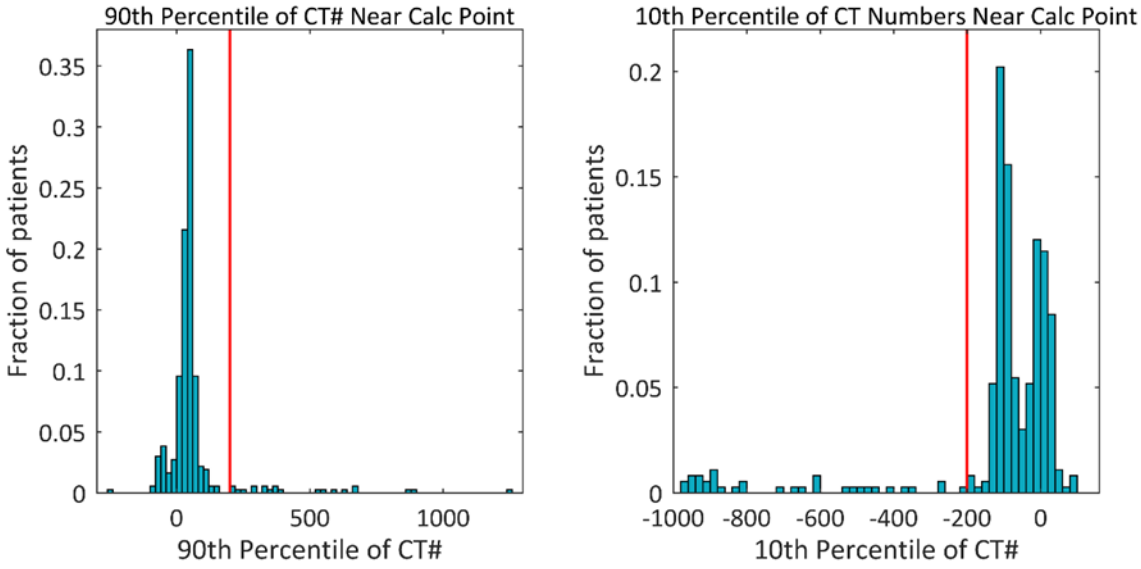

Figure 22. Histogram of the 90th percentile (left) and 10th percentile (right) of the CT numbers within a 1-cm radius sphere around the calculation point for 366 patient plans. The red line indicates a threshold that was used in a QA test of the location of the calculation point.

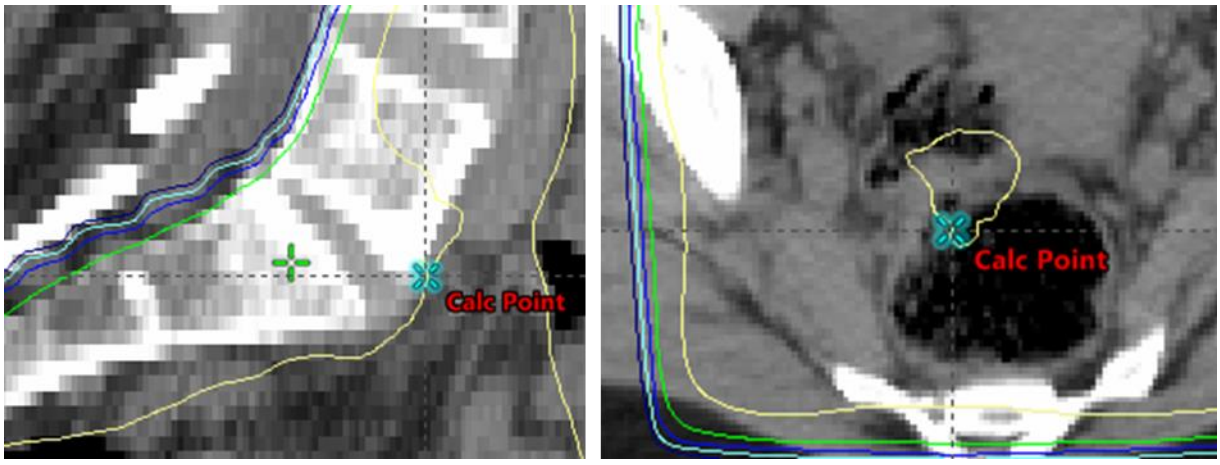

Figure 23. Examples of patient plans that were flagged for high and low density (left and right, respectively). The plan flagged for high density had a CT number of 548 at the 90th percentile, and the plan flagged for low density had a CT number of -375 at the 10th percentile.

Bibliography

[1] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.

[2] F. Bray, "Transitions in human development and the global cancer burden," in *World Cancer Report 2014*, B. W. Stewart and C. P. Wild, Eds. Lyon, France: International Agency for Research on Cancer, 2014, pp. 54–68.

[3] L. N. Shulman, W. Willett, A. Sievers, and F. M. Knaul, "Breast Cancer in Developing Countries : Opportunities for Improved Survival," *J. Oncol.*, vol. 2010, pp. 1–6, 2010.

[4] H. M. Keys, B. N. Bundy, F. B. Stehman, L. I. Muderspach, W. E. Chafe, C. L. Suggs, J. L. Walker, and D. Gersell, "Cisplatin, Radiation, and Adjuvant Hysterectomy Compared with Radiation and Adjuvant Hysterectomy for Bulky Stage IB Cervical Carcinoma," *N. Engl. J. Med.*, vol. 340, no. 15, pp. 1154–1161, Apr. 1999.

[5] M. Morris, P. J. Eifel, J. Lu, P. W. Grigsby, C. Levenback, R. E. Stevens, M. Rotman, D. M. Gershenson, and D. G. Mutch, "Pelvic Radiation with Concurrent Chemotherapy Compared with Pelvic and Para-Aortic Radiation for High-Risk Cervical Cancer," *N. Engl. J. Med.*, vol. 340, no. 15, pp. 1137–1143, Apr. 1999.

[6] P. G. Rose, B. N. Bundy, E. B. Watkins, J. T. Thigpen, G. Deppe, M. A. Maiman, D. L. Clarke-Pearson, and S. Insalaco, "Concurrent Cisplatin-Based Radiotherapy and Chemotherapy for Locally Advanced Cervical Cancer," *N. Engl. J. Med.*, vol. 340, no. 15, pp. 1144–1153, Apr. 1999.

[7] C. W. Whitney, W. Sause, B. N. Bundy, J. H. Malfetano, E. V Hannigan, W. C. Fowler, Jr, D. L. Clarke-Pearson, and S. Liao, "Randomized Comparison of Fluorouracil Plus Cisplatin Versus Hydroxyurea as an Adjunct to Radiation Therapy in Stage IIB-IVA Carcinoma of the Cervix With Negative Para-Aortic Lymph Nodes: A Gynecologic Oncology Group and Southwest Oncology Group Study," *J. Clin. Oncol.*, vol. 17, no. 5, pp. 1339–1339, May 1999.

[8] W. A. Peters, P. Y. Liu, R. J. Barrett, R. J. Stock, B. J. Monk, J. S. Berek, L. Souhami, P. Grigsby, W. Gordon, and D. S. Alberts, "Concurrent Chemotherapy and Pelvic Radiation Therapy Compared With Pelvic Radiation Therapy Alone as Adjuvant Therapy After Radical Surgery in High-Risk Early-Stage Cancer of the Cervix," *J. Clin. Oncol.*, vol. 18, no. 8, pp. 1606–1613, Apr. 2000.

[9] P. J. Eifel, K. Winter, M. Morris, C. Levenback, P. W. Grigsby, J. Cooper, M. Rotman, D. Gershenson, and D. G. Mutch, "Pelvic irradiation with concurrent chemotherapy versus pelvic and para-aortic irradiation for high-risk cervical cancer: An update of Radiation Therapy Oncology Group Trial (RTOG) 90-01," *J. Clin. Oncol.*, vol. 22, no. 5, pp. 872–880, 2004.

[10] L. T. Chuang, S. Temin, R. Camacho, S. Feldman, M. Gultekin, V. Gupta, S. Horton, G. Jacob, E. A. Kidd, K. Lishimpi, C. Nakisige, J.-H. Nam, N. Y. Sheung, W. Small, G. Thomas, and J. S. Berek, "Management and Care of Women With Invasive Cervical Cancer : American Society of Clinical Oncology Resource-Stratified Clinical Practice Guideline," *J. Glob. Oncol.*, vol. 2, no. 5, pp. 1–30, 2016.

[11] International Atomic Energy Agency, "IAEA Human Health Reports No.6 Management of Cervical Cancer: Strategies for Limited-resource Centres — A Guide for Radiation Oncologists," 2013. [Online]. Available: http://www.rho.org/files/IAEA_Radiation_oncology_2013.pdf. [Accessed: 09-Mar-2017].

[12] M. Overgaard, P. S. Hansen, J. Overgaard, C. Rose, M. Andersson, F. Bach, M. Kjaer, C. C. Gadeberg, H. T. Mouridsen, M.-B. Jensen, and K. Zedeler, "Postoperative Radiotherapy in High-Risk Premenopausal Women with Breast Cancer Who Receive Adjuvant Chemotherapy," *N. Engl. J. Med.*, vol. 337, no. 14, pp. 949–955, Oct. 1997.

[13] M. Overgaard, M.-B. Jensen, J. Overgaard, P. S. Hansen, C. Rose, M. Andersson, C. Kamby, M. Kjaer, C. C. Gadeberg, B. B. Rasmussen, M. Blichert-Toft, and H. T. Mouridsen, "Postoperative radiotherapy in high-risk postmenopausal breast-cancer patients given adjuvant tamoxifen: Danish Breast Cancer Cooperative Group DBCG 82c randomised trial," *Lancet*, vol. 353, no. 9165,

pp. 1641–1648, May 1999.

[14] J. Ragaz, I. A. Olivotto, J. J. Spinelli, N. Phillips, S. M. Jackson, K. S. Wilson, M. A. Knowling, C. M. L. Coppin, L. Weir, K. Gelmon, N. Le, R. Durand, A. J. Coldman, and M. Manji, "Locoregional Radiation Therapy in Patients With High-Risk Breast Cancer Receiving Adjuvant Chemotherapy: 20-Year Results of the British Columbia Randomized Trial," *J. Natl. Cancer Inst.*, vol. 97, no. 2, pp. 116–126, Jan. 2005.

[15] P. M. Poortmans, S. Collette, C. Kirkove, E. Van Limbergen, V. Budach, H. Struikmans, L. Collette, A. Fourquet, P. Maingon, M. Valli, K. De Winter, S. Marnitz, I. Barillot, L. Scandolaro, E. Vonk, C. Rodenhuis, H. Marsiglia, N. Weidner, G. van Tienhoven, C. Glanzmann, A. Kuten, R. Arriagada, H. Bartelink, and W. Van den Bogaert, "Internal Mammary and Medial Supraclavicular Irradiation in Breast Cancer," *N. Engl. J. Med.*, vol. 373, no. 4, pp. 317–327, Jul. 2015.

[16] C. W. Taylor, Z. Wang, E. Macaulay, R. Jagsi, F. Duane, and S. C. Darby, "Exposure of the Heart in Breast Cancer Radiation Therapy: A Systematic Review of Heart Doses Published During 2003 to 2013," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 93, no. 4, pp. 845–853, Nov. 2015.

[17] A. Recht, E. A. Comen, R. E. Fine, G. F. Fleming, P. H. Hardenbergh, A. Y. Ho, C. A. Hudis, E. S. Hwang, J. J. Kirshner, M. Morrow, K. E. Salerno, G. W. Sledge, L. J. Solin, P. A. Spears, T. J. Whelan, M. R. Somerfield, and S. B. Edge, "Postmastectomy Radiotherapy: An American Society of Clinical Oncology, American Society for Radiation Oncology, and Society of Surgical Oncology Focused Guideline Update," *J. Clin. Oncol.*, vol. 34, no. 36, pp. 4431–4442, Dec. 2016.

[18] C. W. Taylor and A. M. Kirby, "Cardiac Side-effects From Breast Cancer Radiotherapy," *Clin. Oncol.*, vol. 27, no. 11, pp. 621–629, Nov. 2015.

[19] S. C. Darby, M. Ewertz, P. McGale, A. M. Bennet, U. Blom-Goldman, D. Brønnum, C. Correa, D. Cutter, G. Gagliardi, B. Gigante, M.-B. Jensen, A. Nisbet, R. Peto, K. Rahimi, C. Taylor, and P. Hall, "Risk of Ischemic Heart Disease in Women after Radiotherapy for Breast Cancer," *N. Engl. J. Med.*, vol. 368, no. 11, pp. 987–998, Mar. 2013.

[20] K. E. Henson, P. McGale, C. Taylor, and S. C. Darby, “Radiation-related mortality from heart disease and lung cancer more than 20 years after radiotherapy for breast cancer,” *Br. J. Cancer*, vol. 108, no. 1, pp. 179–182, Jan. 2013.

[21] M. C. Aznar, F. K. Duane, S. C. Darby, Z. Wang, and C. W. Taylor, “Exposure of the lungs in breast cancer radiotherapy: A systematic review of lung doses published 2010–2015,” *Radiother. Oncol.*, vol. 126, no. 1, pp. 148–154, Jan. 2018.

[22] J. Pignol, I. Olivotto, and E. Rakovitch, “A Multicenter Randomized Trial of Breast Intensity-Modulated Radiation Therapy to Reduce Acute Radiation Dermatitis,” *J. Clin. Oncol.*, vol. 26, no. 13, pp. 2085–2092, 2008.

[23] E. H. Zubizarreta, E. Fidarova, B. Healy, and E. Rosenblatt, “Need for Radiotherapy in Low and Middle Income Countries – The Silent Crisis Continues,” *Clin. Oncol.*, vol. 27, no. 2, pp. 107–114, Feb. 2015.

[24] N. R. Datta, M. Samiei, and S. Bodis, “Radiation therapy infrastructure and human resources in low- and middle-income countries: Present status and projections for 2020,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 89, no. 3, pp. 448–457, 2014.

[25] R. Atun, D. A. Jaffray, M. B. Barton, F. Bray, M. Baumann, B. Vikram, T. P. Hanna, F. M. Knaul, Y. Lievens, T. Y. M. Lui, M. Milosevic, B. O’Sullivan, D. L. Rodin, E. Rosenblatt, J. Van Dyk, M. L. Yap, E. Zubizarreta, and M. Gospodarowicz, “Expanding global access to radiotherapy,” *The Lancet Oncology*, vol. 16, no. 10. pp. 1153–1186, 2015.

[26] S. F. Petit, B. Wu, M. Kazhdan, A. Dekker, P. Simari, R. Kumar, R. Taylor, J. M. Herman, and T. McNutt, “Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma,” *Radiother. Oncol.*, vol. 102, no. 1, pp. 38–44, Jan. 2012.

[27] B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, “Data-Driven Approach to Generating Achievable Dose–Volume Histogram Objectives in

Intensity-Modulated Radiotherapy Planning," *Int. J. Radiat. Oncol.*, vol. 79, no. 4, pp. 1241–1247, Mar. 2011.

[28] S. Shiraishi, J. Tan, L. A. Olsen, and K. L. Moore, "Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery," *Med. Phys.*, vol. 42, no. 2, pp. 908–917, Jan. 2015.

[29] I. Xhaferllari, E. Wong, K. Bzdusek, M. Lock, and J. Z. Chen, "Automated IMRT planning with regional optimization using planning scripts," *J. Appl. Clin. Med. Phys.*, vol. 14, no. 1, pp. 176–191, 2013.

[30] T. G. Purdie, R. E. Dinniwell, A. Fyles, and M. B. Sharpe, "Automation and intensity modulated radiation therapy for individualized high-quality tangent breast treatment plans," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 90, no. 3, pp. 688–695, 2014.

[31] T. G. Purdie, R. E. Dinniwell, D. Letourneau, C. Hill, and M. B. Sharpe, "Automated Planning of Tangential Breast Intensity-Modulated Radiotherapy Using Heuristic Optimization," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 81, no. 2, pp. 575–583, Oct. 2011.

[32] G.-P. Chen, E. Ahunbay, and X. A. A. Li, "Automated computer optimization for 3D treatment planning of breast irradiation," *Med. Phys.*, vol. 35, no. 6Part1, pp. 2253–2258, May 2008.

[33] X. Zhao, D. Kong, G. Jozsef, J. Chang, E. K. Wong, S. C. Formenti, and Y. Wang, "Automated beam placement for breast radiotherapy using a support vector machine based algorithm," *Med. Phys.*, vol. 39, no. 5, p. 2536, 2012.

[34] J. Fan, J. Wang, Z. Zhang, and W. Hu, "Iterative dataset optimization in automated planning: Implementation for breast and rectal cancer radiotherapy," *Med. Phys.*, vol. 44, no. 6, pp. 2515–2531, Jun. 2017.

[35] M. B. Barton, M. Frommer, and J. Shafiq, "Role of radiotherapy in cancer control in low-income and middle-income countries.," *Lancet. Oncol.*, vol. 7, no. 7, pp. 584–95, 2006.

[36] M. B. Barton, S. Jacob, J. Shafiq, K. Wong, S. R. Thompson, T. P. Hanna, and G. P. Delaney,

"Estimating the demand for radiotherapy from the evidence: A review of changes from 2003 to 2012," *Radiother. Oncol.*, vol. 112, no. 1, pp. 140–144, 2014.

[37] L. E. Court, K. Kisling, R. McCarroll, L. Zhang, J. Yang, H. Simonds, M. du Toit, C. Trauernicht, H. Burger, J. Parkes, M. Mejia, M. Bojador, P. Balter, D. Branco, A. Steinmann, G. Baltz, S. Gay, B. Anderson, C. Cardenas, A. Jhingran, S. Shaitelman, O. Bogler, K. Schmeller, D. Followill, R. Howell, C. Nelson, C. Peterson, and B. Beadle, "Radiation Planning Assistant - A streamlined, fully automated radiotherapy treatment planning system," *J. Vis. Exp.*, vol. 2018, no. 134, pp. e57411–e57411, Apr. 2018.

[38] E. C. Ford, S. Terezakis, A. Souranis, K. Harris, H. Gay, and S. Mutic, "Quality Control Quantification (QCQ): A Tool to Measure the Value of Quality Control Checks in Radiation Oncology," *Int. J. Radiat. Oncol.*, vol. 84, no. 3, pp. e263–e269, Nov. 2012.

[39] J. Yang, A. Amini, R. Williamson, L. Zhang, Y. Zhang, R. Komaki, Z. Liao, J. Cox, J. Welsh, L. Court, and L. Dong, "Automatic contouring of brachial plexus using a multi-atlas approach for lung cancer radiation therapy," *Pract. Radiat. Oncol.*, vol. 3, no. 4, pp. e139–e147, Oct. 2013.

[40] R. Zhou, Z. Liao, T. Pan, S. A. Milgrom, C. C. Pinnix, A. Shi, L. Tang, J. Yang, Y. Liu, D. Gomez, Q. N. Nguyen, B. S. Dabaja, L. Court, and J. Yang, "Cardiac atlas development and validation for automatic segmentation of cardiac substructures," *Radiother. Oncol.*, vol. 122, no. 1, pp. 66–71, 2017.

[41] R. E. McCarroll, B. M. Beadle, P. A. Balter, H. Burger, C. E. Cardenas, S. Dalvie, D. S. Followill, K. D. Kisling, M. Mejia, K. Naidoo, C. L. Nelson, C. B. Peterson, K. Vorster, J. Wetter, L. Zhang, L. E. Court, and J. Yang, "Retrospective Validation and Clinical Implementation of Automated Contouring of Organs at Risk in the Head and Neck: A Step Toward Automated Radiation Treatment Planning for Low- and Middle-Income Countries," *J. Glob. Oncol.*, vol. 4, no. 4, pp. 1–11, Nov. 2018.

[42] S. Leonard and A. O'Donovan, "Measuring safety culture: Application of the Hospital Survey on

Patient Safety Culture to radiation therapy departments worldwide," *Pract. Radiat. Oncol.*, vol. 8, no. 1, pp. e17–e26, Jan. 2018.

[43] E. C. Ford, K. Smith, S. Terezakis, V. Croog, S. Gollamudi, I. Gage, J. Keck, T. DeWeese, and G. Sibley, "A streamlined failure mode and effects analysis," *Med. Phys.*, vol. 41, no. 6, May 2014.

[44] K. Kisling, L. Zhang, H. Simonds, N. Fakie, J. Yang, R. McCarroll, P. Balter, H. Burger, O. Bogler, R. Howell, K. Schmeler, M. Mejia, B. M. Beadle, A. Jhingran, and L. Court, "Fully automatic treatment planning for external beam radiation therapy of locally advanced cervical cancer – a tool for low-resource clinics," *J. Glob. Oncol.*, pp. 1–9, 2019.

[45] "Postoperative Radiotherapy for Breast Cancer: UK Consensus Statements," *The Royal College of Radiologist, London*, 2016. [Online]. Available: https://www.rcr.ac.uk/system/files/publication/field_publication_files/bfco2016_breast-consensus-guidelines.pdf. [Accessed: 26-Feb-2019].

[46] J. Yang, Y. Zhang, L. Zhang, and L. Dong, "Automatic segmentation of parotids from CT scans using multiple atlases," in *Medical Image Analysis for the Clinic: A Grand Challenge*, 2010, pp. 323–330.

[47] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation," *IEEE Trans. Med. Imaging*, vol. 23, no. 7, pp. 903–921, Jul. 2004.

[48] J. White, A. Tai, D. Arthur, T. Buchholz, S. MacDonald, L. Marks, L. Pierce, A. Recht, R. Rabinovitch, A. Taghian, F. Vicini, W. Woodward, and X. A. Li, "Breast Cancer Atlas for Radiation Therapy Planning: Consensus Definitions," *Radiation Therapy Oncology Group*. [Online]. Available: https://www.rtog.org/LinkClick.aspx?fileticket=vzJFhPaBipE%3D&tabid=236. [Accessed: 13-Oct-2017].

[49] M. Poppe, "Hypofractionated radiation therapy after mastectomy in preventing recurrence in patients with stage IIa-IIIa breast cancer." 2017.

[50] K. L. Moore, G. C. Kagadis, T. R. McNutt, V. Moiseenko, and S. Mutic, “Vision 20/20: Automation and advanced computing in clinical radiation oncology,” *Med. Phys.*, vol. 41, no. 1, 2014.

[51] G. Amit and T. G. Purdie, “Automated planning of breast radiotherapy using cone beam CT imaging,” *Med. Phys.*, vol. 42, no. 2, pp. 770–779, 2015.

[52] B. G. Clark, R. J. Brown, J. Ploquin, and P. Dunscombe, “Patient safety improvements in radiation treatment through 5 years of incident learning,” *Pract. Radiat. Oncol.*, vol. 3, no. 3, pp. 157–163, Jul. 2013.

[53] M. S. Huq, B. A. Fraass, P. B. Dunscombe, J. P. Gibbons, G. S. Ibbott, A. J. Mundt, S. Mutic, J. R. Palta, F. Rath, B. R. Thomadsen, J. F. Williamson, and E. D. Yorke, “The report of Task Group 100 of the AAPM: Application of risk analysis methods to radiation therapy quality management,” *Med. Phys.*, vol. 43, no. 7, pp. 4209–4262, Jun. 2016.

[54] E. C. Ford, R. Gaudette, L. Myers, B. Vanderver, L. Engineer, R. Zellars, D. Y. Song, J. Wong, and T. L. DeWeese, “Evaluation of safety in a radiation oncology setting using failure mode and effects analysis,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 74, no. 3, pp. 852–858, 2009.

[55] S. Broggi, M. C. Cantone, A. Chiara, N. Di Muzio, B. Longobardi, P. Mangili, and I. Veronese, “Application of failure mode and effects analysis (FMEA) to pretreatment phases in tomotherapy,” *J. Appl. Clin. Med. Phys.*, vol. 14, no. 5, pp. 265–277, 2013.

[56] M. C. Cantone, M. Ciocca, F. Dionisi, P. Fossati, S. Lorentini, M. Krengli, S. Molinelli, R. Orecchia, M. Schwarz, I. Veronese, and V. Vitolo, “Application of failure mode and effects analysis to treatment planning in scanned proton beam radiotherapy.,” *Radiat. Oncol.*, vol. 8, no. 1, p. 127, 2013.

[57] A. Wexler, B. Gu, S. Goddu, M. Mutic, S. Yaddanapudi, L. Olsen, T. Harry, C. Noel, T. Pawlicki, S. Mutic, and B. Cai, “FMEA of manual and automated methods for commissioning a radiotherapy treatment planning system,” *Med. Phys.*, vol. 44, no. 9, pp. 4415–4425, 2017.

[58] “American College of Radiology. Radiation Oncology Practice Accreditation Program

Requirements." [Online]. Available: http://www.acraccreditation.org/~/media/ACRAccreditation/Documents/ROPA/Requirements.pdf?la=en. Published July 2, 2018. [Accessed: 01-Nov-2018].

[59] J. T. Faught, "Quantification of IMRT severity scores for improvement of FMEA results [dissertation]," Houston: The University of Texas Graduate School of Biomedical Sciences, 2015.

[60] R. A. Siochi, E. C. Pennington, T. J. Waldron, and J. E. Bayouth, "Radiation therapy plan checks in a paperless clinic," *J. Appl. Clin. Med. Phys.*, vol. 10, no. 1, pp. 43–62, Dec. 2009.

[61] D. Yang and K. L. Moore, "Automated radiotherapy treatment plan integrity verification," *Med. Phys.*, vol. 39, no. 3, pp. 1542–1551, Feb. 2012.

[62] L. A. Olsen, C. G. Robinson, G. R. He, H. O. Wooten, S. Yaddanapudi, S. Mutic, D. Yang, and K. L. Moore, "Automated radiation therapy treatment plan workflow using a commercial application programming interface," *Pract. Radiat. Oncol.*, vol. 4, no. 6, pp. 358–367, Nov. 2014.

[63] E. E. Furhang, J. Dolan, J. K. Sillanpaa, and L. B. Harrison, "Automating the initial physics chart-checking process," *J. Appl. Clin. Med. Phys.*, vol. 10, no. 1, pp. 129–135, 2009.

[64] A. Novak, M. J. Nyflot, R. P. Ermoian, L. E. Jordan, P. A. Sponseller, G. M. Kane, E. C. Ford, and J. Zeng, "Targeting safety improvements through identification of incident origination and detection in a near-miss incident learning system," *Med. Phys.*, vol. 43, no. 5, pp. 2053–2062, Apr. 2016.

[65] O. Gopan, W. P. Smith, A. Chvetsov, K. Hendrickson, A. Kalet, M. Kim, M. Nyflot, M. Phillips, L. Young, A. Novak, J. Zeng, and E. Ford, "Utilizing simulated errors in radiotherapy plans to quantify the effectiveness of the physics plan review," *Med. Phys.*, vol. 45, no. 12, pp. 5359–5365, Dec. 2018.

[66] O. Gopan, J. Zeng, A. Novak, M. Nyflot, and E. Ford, "The effectiveness of pretreatment physics plan review for detecting errors in radiation therapy," *Med. Phys.*, vol. 43, no. 9, pp. 5181–5187, Aug. 2016.

[67] K. L. Moore, R. S. Brame, D. A. Low, and S. Mutic, “Experience-Based Quality Control of Clinical Intensity-Modulated Radiotherapy Planning,” *Int. J. Radiat. Oncol.*, vol. 81, no. 2, pp. 545–551, Oct. 2011.

[68] B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, M. Chuang, R. Taylor, R. Jacques, and T. McNutt, “Patient geometry-driven information retrieval for IMRT treatment plan quality control,” *Med. Phys.*, vol. 36, no. 12, pp. 5497–5505, Nov. 2009.

[69] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 801–818.

[70] C. Cardenas, B. Anderson, L. Zhang, A. Jhingran, H. Simonds, J. Yang, K. Brock, A. Klopp, B. Beadle, L. Court, and K. Kisling, “A Comparison of Two Deep Learning Architectures to Automatically Define Patient-Specific Beam Apertures,” *Med. Phys.*, vol. 45, no. 6, p. e132, 2018.

[71] K. Kisling, J. Johnson, H. Simonds, L. Zhang, A. Jhingran, B. M. Beadle, H. Burger, M. du Toit, N. Joubert, W. Shaw, C. Trauernicht, P. Balter, R. Howell, K. Schmeler, and L. Court, “A Risk Assessment of Automated Treatment Planning and Recommendations for Clinical Deployment (submitted),” *Med. Phys.*, 2019.

[72] B. P. Ziemer, S. Shiraishi, J. A. Hattangadi-Gluth, P. Sanghvi, and K. L. Moore, “Fully automated, comprehensive knowledge-based planning for stereotactic radiosurgery: Preclinical validation through blinded physician review,” *Pract. Radiat. Oncol.*, vol. 7, no. 6, pp. e569–e578, Nov. 2017.

Vita

Kelly Diane Kisling was born in Palm Beach Gardens, Florida, the daughter of Karen Daniels Kisling and Douglas Leroy Kisling. After completing her work at Tolland High School, Tolland, Connecticut in 2004, she entered the Georgia Institute of Technology (Georgia Tech) in Atlanta, Georgia. She received the degree of Bachelor of Science with a major in Nuclear and Radiological Engineering from Georgia Tech in May 2008. She then entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences in Houston, Texas and received the degree of Master of Science with a major in Medical Physics in August 2010. For the next four years, she worked as a junior medical physicist at the MD Anderson Regional Care Centers. She earned her board certification in Therapeutic Medical Physics from the American Board of Radiology in 2014. In January of 2015, she entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences.

Permanent address:

134 W 10th St Apt 3N

New York, NY 77008