

Scientific Article

Clinical Experience With Machine Learning-Based Automated Treatment Planning for Whole Breast Radiation Therapy



Sua Yoo, PhD,^{*,1} Yang Sheng, PhD,¹ Rachel Blitzblau, MD, PhD, Susan McDuff, MD, PhD, Colin Champ, MD, Jay Morrison, BS, CMD, Leigh O'Neill, BS, CMD, Suzanne Catalano, BS, CMD, Fang-Fang Yin, PhD and Q. Jackie Wu, PhD

Duke University Medical Center, Durham, North Carolina

Received 19 October 2020; revised 15 December 2020; accepted 23 December 2020

Abstract

Purpose: The machine learning–based automated treatment planning (MLAP) tool has been developed and evaluated for breast radiation therapy planning at our institution. We implemented MLAP for patient treatment and assessed our clinical experience for its performance.

Methods and Materials: A total of 102 patients of breast or chest wall treatment plans were prospectively evaluated with institutional review board approval. A human planner executed MLAP to create an auto-plan via automation of fluence maps generation. If judged necessary, a planner further fine-tuned the fluence maps to reach a final plan. Planners recorded the time required for auto-planning and manual modification. Target (ie, breast or chest wall and nodes) coverage and dose homogeneity were compared between the auto-plan and final plan.

Results: Cases without nodes ($n = 71$) showed negligible ($<1\%$) differences for target coverage and dose homogeneity between the auto-plan and final plan. Cases with nodes ($n = 31$) also showed negligible difference for target coverage. However, mean \pm standard deviation of volume receiving 105% of the prescribed dose and maximum dose were reduced from $43.0\% \pm 26.3\%$ to $39.4\% \pm 23.7\%$ and $119.7\% \pm 9.5\%$ to $114.4\% \pm 8.8\%$ from auto-plan to final plan, respectively, all with $P \leq .01$ for cases with nodes ($n = 31$). Mean \pm standard deviation time spent for auto-plans and additional fluence modification for final plans were 12.1 ± 9.3 and 13.1 ± 12.9 minutes, respectively, for cases without nodes, and 16.4 ± 9.7 and 26.4 ± 16.4 minutes, respectively, for cases with nodes.

Conclusions: The MLAP tool has been successfully implemented for routine clinical practice and has significantly improved planning efficiency. Clinical experience indicates that auto-plans are sufficient for target coverage, but improvement is warranted to reduce high dose volume for cases with nodal irradiation. This study demonstrates the clinical implementation of auto-planning for patient treatment and the significant importance of integrating human experience and feedback to improve MLAP for better clinical translation.

© 2021 The Authors. Published by Elsevier Inc. on behalf of American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Sources of support: This work was partially supported by an NIH grant (#R01CA201212) and a master research grant from Varian Medical Systems.

Disclosures: Q.J.W. and Y.S. report grants from NIH (R01CA201212) and Varian Medical Systems during the conduct of this study. F.Y. reports grants from Varian Medical Systems during the conduct of this study. R.B. reports grants from Gateway for Cancer Research during the conduct of this study. Other authors have nothing to disclose.

Data sharing statement: Research data are stored in an institutional repository and will be shared upon request to the corresponding author.

* Corresponding author: Sua Yoo, PhD; E-mail: sua.yoo@duke.edu

¹ S.Y. and Y.S. contributed equally to this study.

<https://doi.org/10.1016/j.adro.2021.100656>

2452-1094/© 2021 The Authors. Published by Elsevier Inc. on behalf of American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Radiation therapy is often a critical component of standard treatments for patients with breast cancer. Conventional 3-dimensional conformal radiation therapy (3D CRT) is the most common radiation therapy technique, traditionally using 2 opposing tangential beams with static multileaf collimators (MLCs) and wedges for breast or chest wall irradiation.^{1,2} Due to anatomic variation, it can be challenging to achieve desired target coverage and homogeneous dose distribution using 3D CRT. Various different techniques have been studied to improve target coverage and dose homogeneity while sparing the heart, lungs, and contralateral breast.³⁻⁹

At our institution, an irregular surface compensator technique with static 2 tangential beams and dynamic MLC has been used for breast treatment planning. The irregular surface compensator technique provides good target coverage and homogeneous dose distribution because it uses intensity modulation based on fluence map.^{4,10} A limitation of this technique is that a planner has to manually and iteratively modify the fluence maps until satisfactory dose distribution is achieved.^{4,10} This process takes 2 to 3 hours or more for iterative manual fluence modification and dose calculation, and the plan quality is highly dependent on the planner's experience.⁶

An in-house machine learning–based automated treatment planning (MLAP) tool was developed to improve planning efficiency and plan quality for the irregular surface compensatory technique.¹¹ In late 2018, this tool was successfully implemented in the clinical setting through Eclipse (Varian Medical System, Palo Alto, CA) Scripting Application Programming Interface (ESAPI), and a retrospective study was performed to evaluate breast radiation treatment planning.¹² After the evaluation period, our institution launched the MLAP tool for actual patient treatment in May 2019. Our study aimed to explore the potential and limitations of this auto-planning tool beyond technology development, which is of more interest in a real clinical setting. In addition, human experience of an auto-planning tool should feed back to the technology development phase to enhance further clinical translation. We aimed to collect the experience from the human–automation interaction in this first attempt of implementing an auto-planning tool clinically and learn from these data as we plan for future improvement. This study reports the MLAP performance in terms of efficiency and plan quality in the clinical setting, as well as the results of human modification of the automated plans.

Methods and Materials

The MLAP tool has been used for patient treatment since May 2019 at our institution. This study included all

patients with breast cancer treated using the irregular surface compensator technique between May and December 2019 under institutional review board approval (Pro00102095) at our institution. The number of patients was 102, either breast (ie, lumpectomy cases) or chest wall (CW) (ie, mastectomy cases), and the distribution of the 2 subgroups represents the case spectrum at our institution. All patients underwent computed tomography simulation scans for planning. Prone, supine in free breathing, or supine in deep inspiration breath-hold position was used at the physician's discretion. Planners imported computed tomography images, contoured normal structures and critical organs, and set temporary beams. Physicians contoured targets including nodes following the Radiation Therapy Oncology Group breast atlas.¹³ For breast and CW cases, planning target volume for evaluation was also generated. Physicians also reviewed, modified, and approved the beams to encompass necessary targets with adequate margin before planners proceeded with dose optimization.

In this study, patients were categorized into 2 groups: group 1 included 71 patients without nodal irradiation, and group 2 included 31 patients with nodal irradiation. Group 1 cases had plans with 2 opposing tangential beams set to irradiate the whole breast/CW. Three cases in group 1 required only partial breast irradiation with 2 mini-tangential beams. For group 2, a composite plan of 2 to 3 separate plans was generated. A tangential plan included partial-wide tangential beams to irradiate the entire breast/CW and internal mammary nodes (IMNs).¹⁴ Depending on patient anatomy, a separate plan with an electron IMN beam and a tangential plan with shallow tangential beams were considered in place of a tangential plan with partial-wide tangential beams to improve IMN coverage or to spare the heart and lungs further.^{15,16}

Another separate plan included a half-blocked anterior oblique beam to irradiate supraclavicular nodes (SCLs) at a different isocenter. For patients with deep SCL, anterior and posterior oblique SCL beams were used.¹⁴ The superior border of the tangential beams was matched with the inferior border of the SCL beam(s) using collimator and couch rotation. All beams except separate electron IMN beams used MLCs to encompass targets and spare the heart and lungs. For cases with axillary node (Ax) irradiation, MLC in tangential beams and SCL beam(s) were adjusted to encompass Ax properly. The prescriptions ranged from 1.8 to 2.67 Gy per fraction to a total dose of 40.05 to 50 Gy. Boost plans were not included in this study.

After beams and MLC shapes were approved by physicians, planners executed the MLAP script through ESAPI for tangential beams. MLAP generated optimal fluence map for each tangential beam within the approved MLC shape, and planners manually imported the optimal fluence map. If mixed low- and high-energy beams were used, the MLAP generated corresponding fluence maps

for each beam and energy. Dose distributions were calculated with the anisotropic analytical algorithm (AAA version 15.6.03) in Eclipse. For CW cases, a plan with 0.5 cm bolus was generated by repeating this process, and a composite plan combined the plan with bolus and the plan without bolus with the fractionation ratio of 2:1 or 5:3 following the prescription. A planner adjusted normalization to achieve adequate target coverage. Either the tangential plan using MLAP or the composite plan including the tangential plan(s) using MLAP was labeled as the auto-plan in this study.

Planners reviewed the isodose distributions of the auto-plan and, if judged necessary, modified fluence maps to reach to the final plan. Planners recorded time spent for auto-planning and additional time spent for manual fluence modification for the final plan. Therefore, the total planning time was the sum of the auto-planning time and manual modification time. If no modification was made, the auto-plan was used as the final plan with no manual modification time. Physicians reviewed and approved the final plan for patient treatment. If improvement was necessary based on physician’s judgment, physicians discussed with planners and planners further modified fluence maps to update the final plan. The time spent for this second-round modification was also added to the time spent for manual fluence modification in the final plan.

To evaluate the target coverage, the percentage volume receiving 90% or 95% of the prescription (V90 or V95) was compared between the auto-plan and final plan for breast/CW, breast/CW planning target volume for evaluation, SCL, IMN, and Ax. To evaluate the dose homogeneity, maximum dose (D_{max}) and the percentage volumes of 105% and 110% of the prescription (V105 and V110) were compared. Mean dose (D_{mean}) and percentage dose to 5% of the volume (D5) to the heart and ipsilateral lung were also compared. The statistical significance was determined with the 2-tailed paired *t* test (*P* value) with a significance threshold of 0.01 ($P < .01$).

Results

Table 1 lists the summary of patient characteristics. Group 1 (cases without nodes) consisted of 71 cases, and group 2 (cases with nodal irradiation) had 31 cases. All cases in group 1 had tangential beams alone to irradiate whole or partial breast/CW. All cases in group 2 had separate 3D beam(s) to irradiate SCL except 1 CW case, which included IMN in partial-wide tangential beams without SCL treatment. Planners selected low-energy beams (eg, 6 MV) or mixed energy beams (eg, 6 MV, 10 MV, and/or 15 MV) based on the patient anatomy.

Table 2 shows the dosimetric parameters and planning time for comparison. For both treatment groups, target coverage was very similar between the auto-plan and final plan except IMN V90, which showed about 3% better

Table 1 Characteristics of patients (n = 102)

Categories	Variables	n	Group 1 without nodes	Group 2 with nodes
Side	Left	57	43	14
	Right	43	27	16
RT position	Bilateral	2	1	1
	Supine	95	64	31
	Prone	7	7	0
Motion management	Deep-inspiration breath hold	57	40	17
	Free breathing	45	31	14
	Postoperative status	79	70	9
Treatment area	Mastectomy	23	1	22
	Breast/CW without nodes (group 1)	71		
	Mini-tangential beams for tumor bed only	3		
	Tangential beams for whole breast/CW	68		
	Breast/CW with nodes (group 2)			31
	With IMN included in PWT			27
	With IMN included in separate electron beam			4
	With SCL included in separate 3D beam(s)			30
	With Ax included in PWT/SCL 3D beam(s)			17
Beam energy used	6 MV only	50	39	11
	6 MV and 15 MV	51	32	19
	10 MV and 15 MV	1	0	1

Abbreviations: 3D = 3-dimensional; Ax = axillary node; CW = chest wall; IMN = internal mammary nodes; PWT = partial-wide tangential beams; RT = radiation therapy; SCL = supraclavicular nodes.

coverage on average in final plans than in auto-plans. However, the difference was not statistically significant ($P = .08$), suggesting that the auto-plan was satisfactory for target coverage, and manual modification performed by a planner brought marginal improvement. The mean V105 for group 2 was very large for both auto-plan and final plan because the absolute V105 value was obtained from the composite plan including dose contribution from SCL and IMN plans but normalized to breast/CW volume. For group 1, differences in dose homogeneity parameters (eg, V105, V110, and D_{max}) were small and statistically insignificant, whereas they were 3% to 5% and statistically significant ($P \leq .01$) for group 2. Thus,

Table 2 Summary of the dosimetric parameters and planning time

	Group 1			Group 2		
	Auto-plan	Final plan	<i>P</i>	Auto-plan	Final plan	<i>P</i>
Breast/CW V95 (%)	97.5 ± 2.6	97.4 ± 2.6	.78	92.4 ± 6.6	92.3 ± 6.8	.45
Breast/CW PTVeval V95 (%)	93.2 ± 4.0	93.1 ± 4.0	.34	89.9 ± 7.9	89.6 ± 8.2	.37
SCL V90 (%)				97.0 ± 4.6	97.0 ± 4.6	.01
IMN V90 (%)				90.4 ± 17.9	93.4 ± 10.3	.08
Ax V90 (%)				92.1 ± 5.9	91.5 ± 6.2	.17
V105 (%)	11.5 ± 11.1	11.2 ± 9.9	.46	43.0 ± 26.3	39.4 ± 23.7	.01
V110 (%)	0.6 ± 4.6	0.4 ± 3.2	.31	6.5 ± 10	3.3 ± 9.7	<.01
D _{max} (%)	109.2 ± 1.4	108.8 ± 1.5	.01	119.7 ± 9.5	114.4 ± 8.8	<.01
Heart D _{mean} (%)	1.7 ± 0.7	1.7 ± 0.7	.82	2.9 ± 1.7	2.9 ± 1.7	.07
Heart D5% (%)	5.0 ± 2.2	5.0 ± 2.2	.86	7.4 ± 4.5	7.3 ± 4.5	.76
Ipsilateral lung D _{mean} (%)	13.7 ± 6.0	13.7 ± 5.9	.99	31.8 ± 5.3	31.8 ± 5.3	.76
Ipsilateral lung D5% (%)	73.2 ± 27.4	73.3 ± 27.3	.55	88.4 ± 9.1	88.8 ± 9.2	.45
Planning time (min)	12.1 ± 9.3	13.1 ± 12.9		16.4 ± 9.7	26.4 ± 16.4	
Total planning time (min)		25.7 ± 16.8			42.9 ± 20.3	

Abbreviations: Ax = axillary node; CW = chest wall; IMN = internal mammary nodes; PTVeval = planning target volume for evaluation; RT = radiation therapy; SCL = supraclavicular nodes.

auto-plan was sufficient to achieve homogeneous dose distribution for cases without nodal irradiation, whereas manual modification was necessary for cases with nodal irradiation to improve dose homogeneity.

On average, auto-planning time was 12.1 ± 9.3 minutes for cases without nodes and 16.4 ± 9.7 minutes for cases with nodes. Auto-planning time included all aspects of treatment planning, such as copying and pasting beams and plans, changing energy, adding bolus, importing fluence maps, running dose calculation, reviewing isodose distribution and dose-volume histograms, and so forth. On average, planners spent an additional 13.1 ± 12.9 minutes to modify fluence for cases without nodes and 26.4 ± 16.4 minutes for cases with nodes, attempting improvements. Therefore, the mean total planning time for the final plans was 25.7 ± 16.8 minutes for cases without nodes and 42.9 ± 20.3 minutes for cases with nodes.

Figure 1 shows the binned boxplot of dose difference (auto-plan minus final plan) in absolute volume. A negative difference for bins in dose range of 95% to 100% indicates improved coverage in final plans, whereas a positive difference for bins in dose range of 100% and beyond indicates a reduced high-dose volume and thus improved dose homogeneity in final plans. For group 1 (Fig 1a), the median value for each box was around zero, which indicates similar overall coverage as well as high dose volume control. For group 2 (Fig 1b), the boxes in the target coverage range (95%-100% bins) were centered around the zero line whereas the boxes in the high-dose volume range (105%-109% bins noted with the red bracket in Fig 1b) were trending in the positive direction, indicating that manual modification improved dose homogeneity. Both groups showed a larger interquartile range within 101% to 103% bins, indicating differences between auto-plans and final plans; however, the median

differences were stable around zero, indicating that auto-plans and final plans were overall similar in this dose range.

Figure 2 displays the dosimetric results (V95 and V105) for individual cases. Differences of V95 between auto-plan and final plan (Diff V95 = V95 of auto-plan – V95 of final plan) and differences of V105 (Diff V105 = V105 of auto-plan – V105 of final plan) are plotted for group 1 using an “X” mark and for group 2 using an “O” mark. The vertical axis indicates Diff V95, and positive Diff V95 implies that the manual modification did not improve the target coverage. The horizontal axis indicates Diff V105, and positive Diff V105 implies that the manual modification improved the dose homogeneity. Therefore, the blue shaded area indicates overall better plan quality for auto-plan and the red shaded area favors final plan with manual modification. Two cases from group 2 are in the red shaded area, suggesting that the manual modification improved both target coverage and dose homogeneity. A majority of cases showed small Diff V95 (<1%) in groups 1 and 2 and small Diff V105 (<5%) in group 1. Group 2 had 13 cases (“O” in the red box in Fig 2) with Diff V105 greater than +5%, which implies final plan with manual modification improved the dose homogeneity while maintaining the target coverage within 1% difference compared with auto-plan.

Figures 3 and 4 show example cases from each of the 2 groups for isodose comparison (a) versus (b) between auto-plans and final plans and for fluence comparison (c) versus (d). Target coverage and high dose volume were very similar between the auto-plan and the final plan for group 1 (Fig 3a vs Fig 3b) and fluence maps are slightly different (Fig 3c vs Fig 3d). In contrast, the group 2 example shows a 110% isodose volume (purple line in Fig 4a) at the junction between the tangential beams and

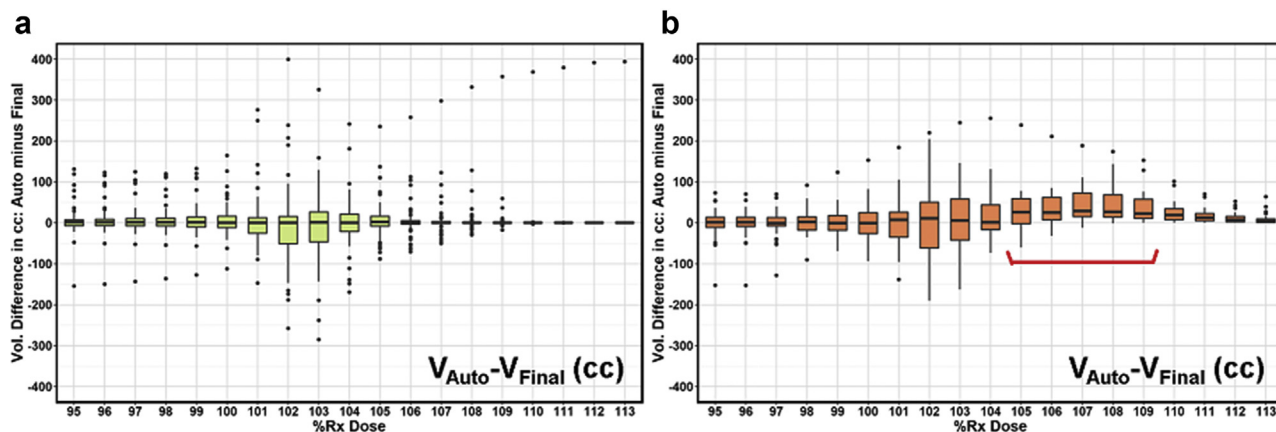


Figure 1 Dose-volume histogram (DVH) difference (auto-plan minus final plan) in target coverage and dose homogeneity for bins in dose range of 95% to 113% for group 1 (a) and group 2 (b). The red bracket notes the range where final plans show improvement in dose homogeneity.

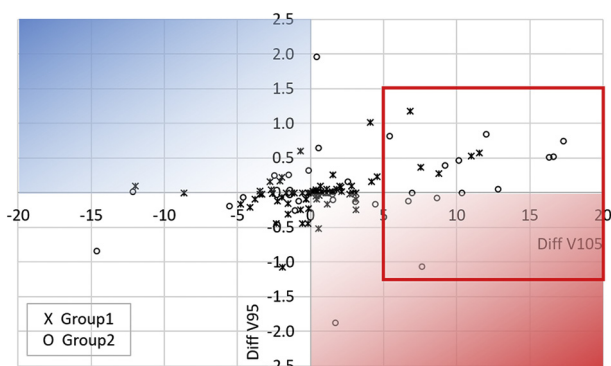


Figure 2 Scatter plot of individual cases for Diff V105 (horizontal axis) and Diff V95 (vertical axis). Diff V105 = V105 of auto-plan – V105 of final plan. Diff V95 = V95 of auto-plan – V95 of final plan. Group 1 cases are marked as “X” and group 2 cases are marked as “O.”

SCL beam in the auto-plan whereas the final plan no longer has a 110% isodose volume in Figure 4b. The fluence map comparison shows that some modification was made from auto-plan to final plan.

Discussion

We developed MLAP for breast/CW radiation treatment planning using a machine learning algorithm that learned the correlation between anatomic features of breast/CW patients and optimal fluence maps from tangential beams using a random forest (RF) model.¹¹ The RF model summarizes the relationship between input features (eg, shape-based features, gray level intensity, penetration depth in breast target, penetration depth in lung) and output variables (pixel-wise fluence map). RF is a nonlinear model that initializes decision trees from randomly sampled data in a training data set and generates a prediction by averaging the output from all trees. The

RF model was trained using 20 plans with 150 trees. For query cases, the RF model predicted fluence intensity at the pixel level, and the entire fluence map served as the fluence estimation for the corresponding beam. A deterministic process that estimated fluence map is reproducible so long as beam’s preset parameters, such as gantry, jaw setting, aperture shape, beam weight, energy, and plan normalization, are consistent. Validation using actual patient cases demonstrated a comparable plan.¹¹

We implemented the MLAP in the clinical setting using Eclipse ESAPI and performed a retrospective study for clinical feasibility in 2018.¹² Thirty whole breast/CW cases without nodal irradiation were planned by using the MLAP and compared with the manual plans used for patient treatment. As estimated from the previous validation study in the research setting, auto-plans were comparable with manual plans for plan quality and the planning time was reduced dramatically. The mean breast/CW V95 was 96.7% (standard deviation [SD], 5.0%) for the manual plans and 96.7% (SD, 4.8%) for the auto-plans ($P = .89$). The mean V105 was 21.6% (SD, 29.8%) for the manual plans and 20.4% (SD, 30.5%) for the auto-plans ($P = .22$). The mean planning time was reduced from 110.2 min to 6.4 min by using MLAP.¹²

The MLAP was trained with cases without nodal irradiation, and both prior validation studies (in the research setting and clinical retrospective study in the clinical setting)^{11,12} were also performed with cases without nodal irradiation. The group 1 results in this study follow the consistent trend that was seen in the previous 2 studies, that that auto-plans are satisfactory for target coverage and dose homogeneity. This result implies that auto-plans were acceptable clinically for cases without nodal irradiation. The data for group 2 also show that auto-plans achieve satisfactory target coverage comparable to final plans. However, we noted that the supraclavicular junction area had noticeable elevated V105% volume for group 2, indicated in Figure 2 with the red box

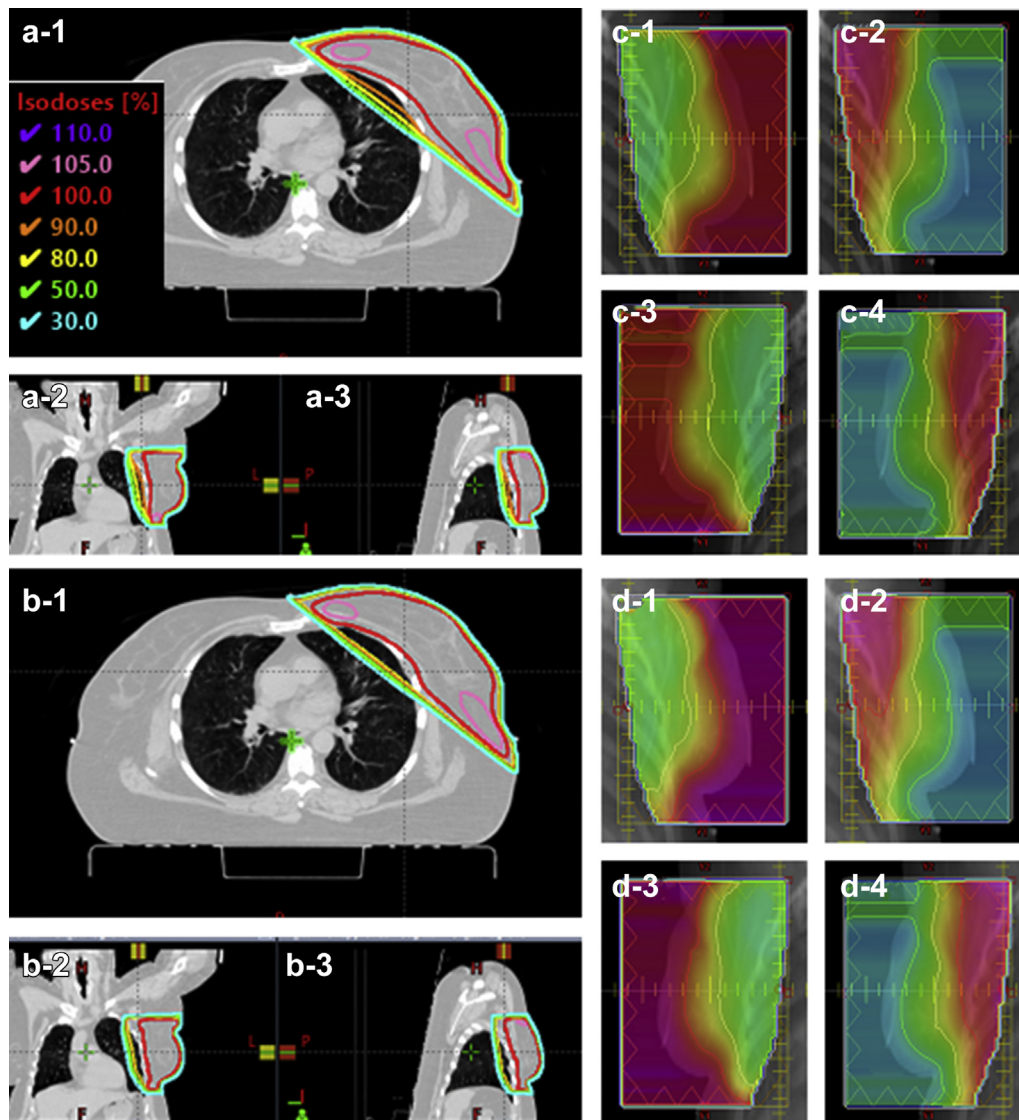


Figure 3 Example case from group 1 for isodose and fluence comparison. (a) and (c) are from the auto-plan and (b) and (d) are from the final plan. (a-1) and (b-1) are axial view, (a-2) and (b-2) are coronal view, and (a-3) and (b-3) are sagittal view. (c-1) and (d-1) are medial tangential beams using 6 MV, (c-2) and (d-2) are medial tangential beams using 15 MV, (c-3) and (d-3) are lateral tangential beams using 6 MV, and (c-4) and (d-4) are lateral tangential beams using 15 MV.

(“O” marks in the red box). Similarly, 4 cases in group 2 with a separate electron IMN plan showed very high dose along the junction between the medial tangential beam and electron IMN beam. A separate IMN plan tends to add dose at the junction with photon beams due to the abutting photon and electron fields.¹⁷ This increased V105 volume in the junction areas because the dose contributions from the SCL and the IMN plans were not considered in the original modeling. These cases are the areas where the MLAP can further improve by incorporating dose contributions from the SCL and IMN plans and providing special fine-tuning processes targeting these types of scenarios.

After the 2 previous validation studies,^{11,12} we took one step closer in automating treatment planning for

clinic. We launched MLAP for real patient treatment planning in clinical settings in May 2019 and performed this prospective study to assess our planning experience with this first in-house auto-planning tool between May and December 2019. To maintain our clinical workflow and to integrate the MLAP tool seamlessly in clinic, blind review was not performed in this study. Our experience indicated that planners had a tendency to modify fluence maps as they used to do in the manual irregular surface compensator technique. The study results indicated negligible differences between auto-plan and final plan for group 1 and for half of group 2. Therefore, physicians would have accepted the auto-plan without manual edits. In addition, all plans (with or without manual modification) in this study passed our quality assurance.

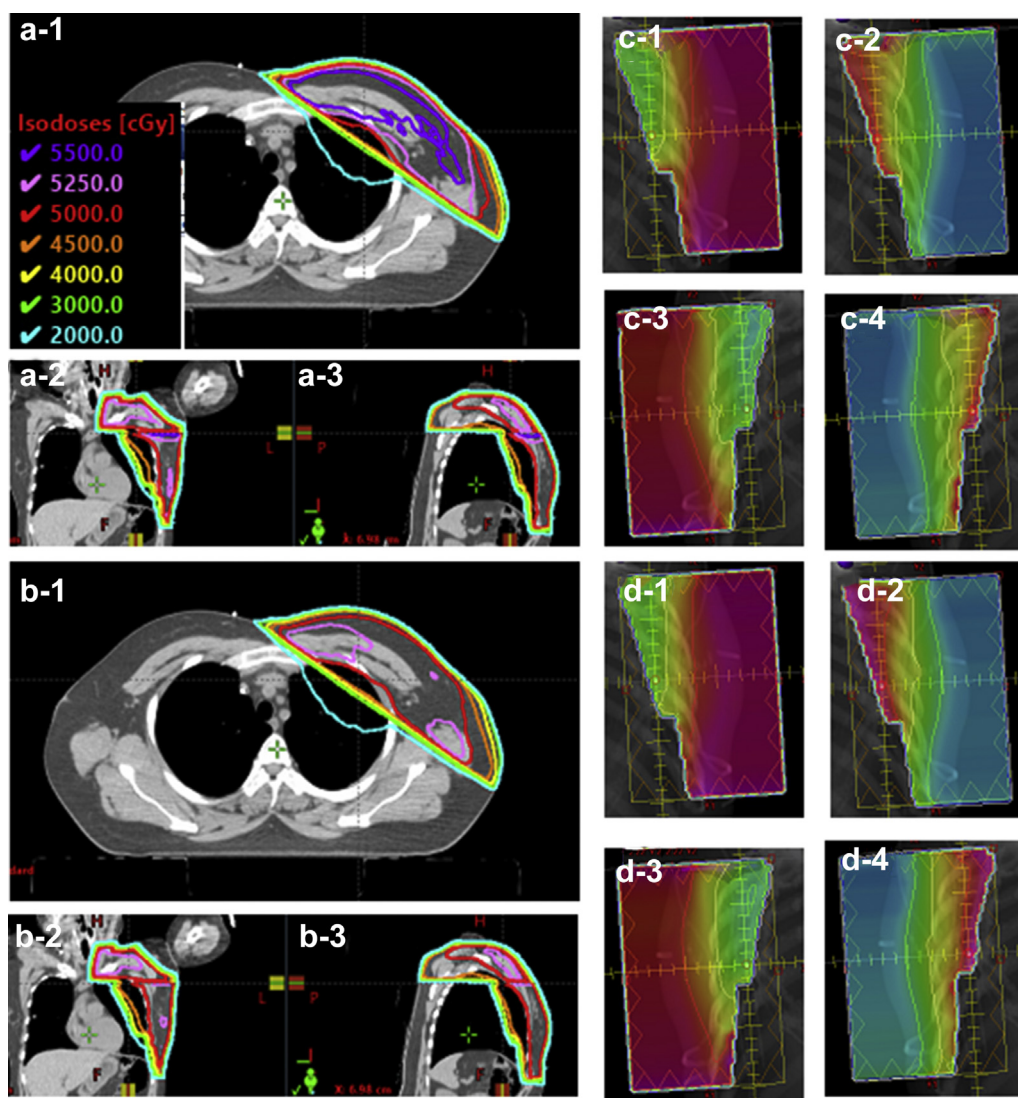


Figure 4 Example case with 50 Gy prescription from group 2 for isodose and fluence comparison. (a) and (c) are from the auto-plan and (b) and (d) are from the final plan. (a-1) and (b-1) are axial view, (a-2) and (b-2) are coronal view, and (a-3) and (b-3) are sagittal view. (c-1) and (d-1) are medial tangential beam using 6 MV, (c-2) and (d-2) are medial tangential beams using 15 MV, (c-3) and (d-3) are lateral tangential beams using 6 MV, and (c-4) and (d-4) are lateral tangential beams using 15 MV.

We identified 4 outlier cases (2 from group 1 and 2 from group 2) with very large separation (eg, 32–36 cm) from the beam entry to the beam exit along the central axis. After reviewing the plans, physicians requested further improvement in target coverage and reduction of high-dose volumes for these patients. Planners generated a hybrid plan as a final plan by adding an open 3D beam with static MLC for each tangential angle and further modifying fluence maps for existing tangential beams. The total modification time was 70 and 90 minutes for 2 cases and not recorded for the other 2 cases. One case had IMN coverage V90 improved from 4.4% to 54.8% and 3 cases had V105 reduced by 6.8% to 17.3%. Our experience with the initial version of MLAP advises more customization modules and features to handle such unique outliers for future improvement.

There have been excellent studies to develop automated tools for breast treatment planning processes with the goals of improving planning efficiency and standardizing plan quality.^{18–23} Those tools were developed to automate 3D CRT,¹⁸ field-in-field,^{19,23} or intensity modulated radiation therapy techniques.^{21,22} Our auto-planning tool was developed based on the irregular surface compensator technique because our institution has been using it for over 15 years. The transition from manually planning to auto-planning required minimal training for planners and no changes in plan evaluation for physicians or quality assurance processes for physicists because MLAP maintained the same planning and delivery techniques.

The MLAP tool incorporated advanced machine learning techniques and designed an algorithm that can

generate fluence maps robust to various breast shapes and sizes. Hence, we believe the technology itself is novel for radiation therapy treatment planning. However, as is the case for many ML techniques designed in the laboratory, its utility and performance in the real clinical setting could be very different from the analysis performed on a well-controlled, retrospective data set. The current study is one of the first prospective studies of in-house auto-planning tools in clinical practice and aims to explore the potential and limitations of such a tool beyond technology development, which is of more interest in clinical practice. Therefore, we believe the study also has value in its robustness and consistent performance in the clinic environment and case spectrum. We hope the results of this study will encourage the research community to explore the translation of automation technologies to clinics and offer confidence to the clinics on the adoption of in-house-developed, dedicated auto-planning tools.

Over the study period and during our current clinical practice, planners have been continuously using the MLAP tool for breast/CW radiation therapy planning and have provided constructive feedback. The initial MLAP currently implemented in the clinic setting has some manual steps such as energy selection and importing optimal fluence maps. Planners' feedback included integration of such manual steps in the auto-planning process. An assessment of our experience provided a critical insight about the current MLAP: The dose contributions for SCL and separate IMN beam need to be considered in the tangential plan. In addition, we learned that more customization modules and features can be added for future development to handle outliers and unique scenarios specific to patients.

Conclusions

The MLAP tool has been successfully implemented for routine clinical practice at our institution and has significantly improved planning efficiency. Clinical experience with MLAP indicates that auto-plans are sufficient for target coverage, but improvement is warranted to reduce high-dose volume for cases with nodal irradiation. This study demonstrates clinical implementation of auto-planning for patient treatment and the importance of integrating human experience and feedback to improve MLAP for better clinical translation.

References

1. Solin LJ, Chu JC, Sontag MR, et al. Three-dimensional photon treatment planning of the intact breast. *Int J Radiat Oncol Biol Phys.* 1991;21:193-203.
2. Sonnik D, Selvaraj RN, Faul C, Gerszten K, Heron DE, King GC. Treatment techniques for 3D conformal radiation to breast and chest

wall including the internal mammary chain. *Med Dosimet.* 2007;32:7-12.

3. Evans PM, Donovan EM, Fenton N, et al. Practical implementation of compensators in breast radiotherapy. *Radiother Oncol.* 1998;49:255-265.
4. Caudell JJ, De Los Santos JF, Keene KS, et al. A dosimetric comparison of electronic compensation, conventional intensity modulated radiotherapy, and tomotherapy in patients with early-stage carcinoma of the left breast. *Int J Radiat Oncol Biol Phys.* 2007;68:1505-1511.
5. Beckham W, Popescu C, Patenaude V, Wai E, Olivetto I. Is multibeam IMRT better than standard treatment for patients with left-sided breast cancer? *Int J Radiat Oncol Biol Phys.* 2007;69:918-924.
6. Hideki F, Nao K, Hiroyuki H, Hiroshi K, Haruyuki F. Improvement of dose distribution with irregular surface compensator in whole breast radiotherapy. *J Med Phys.* 2013;38:115-119.
7. Kuwahata N, Fujita H, Yamanishi H, Okazaki E, Fukuda H. Dosimetric comparison of irregular surface compensator and field-in-field for whole breast radiotherapy. *J Med Phys.* 2018;43:79-84.
8. Boman E, Rossi M, Haltamo M, Skyttä T, Kapanen M. A new split arc VMAT technique for lymph node positive breast cancer. *Physica Medica.* 2016;32:1428-1436.
9. Balaji K, Subramanian B, Yadav P, Anu Radha C, Ramasubramanian V. Radiation therapy for breast cancer: Literature review. *Med Dosimet.* 2016;41:253-257.
10. Furuya T, Sugimoto S, Kurokawa C, Ozawa S, Karasawa K, Sasai K. The dosimetric impact of respiratory breast movement and daily setup error on tangential whole breast irradiation using conventional wedge, field-in-field and irregular surface compensator techniques. *J Radiat Res.* 2013;54:157-165.
11. Sheng Y, Li T, Yoo S, et al. Automatic planning of whole breast radiation therapy using machine learning models. *Front Oncol.* 2019;9:750.
12. Yoo S, Sheng Y, Blitza R, et al. Implementation of machine learning-based treatment planning tool for whole breast radiotherapy using irregular surface compensator technique. The ASTRO 61st Annual Meeting; 2019; Chicago, IL.
13. MacDonalds S, Lee C, Fagundes M, et al. RADCOMP Breast Atlas v.3. Available at: <https://www.nrgoncology.org/Portals/0/Scientific%20Program/CIRO/Atlases/RADCOMP/RADCOMP%20Breast%20Atlas%20v.3%20-%20bigreduced.pdf?ver=2020-08-01-140849-360>. Accessed February 16, 2021.
14. Dumane VA, Bakst R, Green S. Dose to organs in the supraclavicular region when covering the internal mammary nodes (IMNs) in breast cancer patients: A comparison of volumetric modulated arc therapy (VMAT) versus 3D and VMAT. *PLoS One.* 2018;13:e0205770.
15. Taylor CW, Wang Z, Macaulay E, Jaggi R, Duane F, Darby SC. 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.* 2015;93:845-853.
16. Sautter-Bühl M-L, Hültenschmidt B, Melcher U, Ulmer HU. Radiotherapy of internal mammary lymph nodes in breast cancer: Principle considerations on the basis of dosimetric data. *Strahlentherapie und Onkologie.* 2002;178:18-24.
17. Sun C, Cheng C-W, Shimm DS, Cassady JR. Dose profiles in the region of abutting photon and electron fields in the irradiation of head and neck tumors. *Med Dosimet.* 1998;23:5-10.
18. Chen G-P, Ahunbay E, Li XA. Automated computer optimization for 3D treatment planning of breast irradiation. *Med Phys.* 2008;35:2253-2258.
19. Kislak K, Zhang L, Shaitelman SF, et al. Automated treatment planning of postmastectomy radiotherapy. *Med Phys.* 2019;46:3767-3775.
20. Lin T-C, Lin C-Y, Li K-C, et al. Automated hypofractionated IMRT treatment planning for early-stage breast Cancer. *Radiat Oncol.* 2020;15:67.

21. Purdie TG, Dinniwell RE, Letourneau D, Hill C, Sharpe MB. Automated planning of tangential breast intensity-modulated radiotherapy using Heuristic optimization. *Int J Radiat Oncol Biol Phys.* 2011;81:575-583.
22. Purdie TG, Dinniwell RE, Fyles A, Sharpe MB. Automation and intensity modulated radiation therapy for individualized high-quality tangent breast treatment plans. *Int J Radiat Oncol Biol Phys.* 2014;90:688-695.
23. Esho TO, Chung CV, Thompson JU, et al. Optimization of auto-generated chest-wall radiation treatment plans developed for post-mastectomy breast cancer patients in underserved clinics. *Med Dosimet.* 2020;45:102-107.