

Article

Evaluation of Patient-Specific Quality Assurance Results to Optimise Automatic Flash Margin in Monte Carlo Algorithm for VMAT of the Breast and Chest Wall Irradiation

Cemile Ceylan ^{1,2}  and Sevim Sahin ^{3,*} 

¹ Department of Radiation Oncology, Medicalpark Istanbul Oncology Hospital, Istanbul 34758, Turkey; cemile.ceylan@yeditepe.edu.tr

² Department of Medical Physics, Faculty of Health Sciences, Yeditepe University, Istanbul 34758, Turkey

³ Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences, Fenerbahce University, Istanbul 34758, Turkey

* Correspondence: sevim.sahin@fbu.edu.tr; Tel.: +90-216-910-1907

Abstract

Background/Objectives: To evaluate the impact of automatic flash (AF) margin thickness on plan complexity and patient-specific quality assurance (PSQA) outcomes in 4-partial-arc VMAT (4pVMAT) for breast and chest wall irradiation using the Monaco Monte Carlo treatment planning system. **Methods:** Twenty patients (11 left-sided, 9 right-sided) previously treated with 4pVMAT were retrospectively replanned with five AF thicknesses (AF0, AF3, AF5, AF10, AF15), yielding 100 plans on Elekta Versa HD/Agility MLC. Plan complexity was assessed via modulation factor (MF), total monitor units (MU), segment number, MU efficiency, and normalisation ratio. PSQA was performed using the IBA Matrixx Resolution detector, with gamma passing rates (GPRs) evaluated under 3%/3 mm, 3%/2 mm, and 2%/2 mm criteria. Friedman and Wilcoxon signed-rank tests were used with Bonferroni correction. **Results:** All five complexity metrics were statistically invariant across AF0–AF15 (Friedman $p > 0.05$; no pairwise comparison against AF10 reached significance). Mean GPR increased monotonically up to AF10— $92.28 \pm 4.83\%$ (AF0) to $93.73 \pm 4.55\%$ (AF10) at 3%/3 mm—and plateaued at AF15. After Bonferroni correction, AF0 vs. AF10 ($p = 0.0003$) and AF3 vs. AF10 ($p = 0.005$) remained significant under the 3%/3 mm criterion. The proportion of plans meeting $GPR \geq 95\%$ at 3%/3 mm rose from 25% (AF0) to 45% (AF10). **Conclusions:** AF thickness has no clinically meaningful effect on plan complexity but significantly affects delivery accuracy in 4pVMAT breast and chest wall irradiation. AF10 emerges as the optimal margin—maximising delivery accuracy at a plateau without increasing plan complexity—with AF5 as the practical lower bound and AF15 as the upper bound.



Academic Editor: Thomas Lindner

Received: 2 June 2026

Revised: 25 June 2026

Accepted: 29 June 2026

Published: 2 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: breast cancer; VMAT; autoflash; plan complexity; patient-specific quality assurance; Monte Carlo

1. Introduction

Radiation therapy techniques such as volumetric-modulated arc therapy (VMAT), intensity-modulated radiation therapy (IMRT), and three-dimensional conformal radiation therapy (3D-CRT) have been effectively employed in the treatment of breast cancer following breast-conserving surgery and in node-positive post-mastectomy cases [1]. The use of VMAT, a technique that can deliver highly conformal dose distributions to the

target while limiting the dose to the organs at risk (OARs), for breast cancer radiotherapy—including breast irradiation for patients having breast-conserving surgery and chest wall irradiation for node-positive patients following mastectomy—is becoming standard in most radiotherapy centres [2–4].

The application of VMAT in breast cancer is controversial due to the potential risk of secondary cancer caused by low-dose exposure; however, several studies have indicated that comparing VMAT with IMRT and traditional 3D conformal techniques highlights critical trade-offs, including better dose homogeneity and improved OAR sparing [5–9]. Currently, various VMAT techniques, such as multi-partial arc VMAT, are employed to produce plans that improve dose homogeneity within the target and assist in achieving dose constraints for OARs even in cases involving the axillary lymph and internal mammary nodes (IMNs).

However, the VMAT technique has raised concerns regarding achieving adequate Planning Target Volume (PTV) coverage with the prescribed dose and optimal superficial dose because of the deformation of the surface of the breast, while addressing setup uncertainties poses a challenge due to inter- and intrafractional motion. This issue has adversely impacted the precision of treatment delivery in comparison to static fields for breast irradiations due to the highly modulated nature of VMAT [10]. In addition, VMAT involves multiple beam incidences in contrast to conventional tangential beam arrangements, often including more perpendicular beam entry angles relative to the skin surface, which may reduce the delivered skin dose and resulting compromise target coverage [11,12]. While the impact of flash thickness on plan quality remains contentious, several studies [13,14] have recommended implementing auto flash (AF) in VMAT breast irradiation to address potential setup errors and breast deformation beyond the skin surface. Most treatment planning systems enable users to specify the thickness of the flash region to achieve a clinically acceptable plan by automatically extending fluence and shaping Multileaf Collimator (MLC) leaves to account for the thickness of the flash region during the planning optimisation phases of the treatment planning algorithm, thereby improving skin dose uniformity and enhancing treatment robustness against patient-specific variations in VMAT breast irradiation. Hubley et al. [15] demonstrated that plan quality reduces as the AF margin increases; however, a recommendation of a 7–10 mm flash region can yield a robust plan.

While many studies have examined the impact of flash thickness on plan quality and robustness, there is a paucity of research connecting flash thickness to complexity metrics, including modulation factor, segment count, and the passing rates of patient-specific quality assurance (PSQA) outcomes that are directly influenced by these metrics in VMAT breast plans. This study aims to identify the optimal flash thickness that helps generate a high-quality treatment plan by ensuring adequate coverage of the PTV and minimising dose discrepancies near the surface, as well as preventing the compromise of PSQA results for breast and chest wall VMAT irradiation with or without regional lymph nodes. Furthermore, to the best of our knowledge, this study is the first to demonstrate the relationship between AF margin thickness and both plan complexity and PSQA outcomes within a partial VMAT breast radiotherapy planning technique using Monte Carlo-based dose calculation with the AF feature.

2. Materials and Methods

2.1. Patient Selection

This retrospective study encompassed a total of 20 patients (11 left-sided and 9 right-sided) who were randomly chosen, including early-stage cases and stages requiring irradiation of regional lymph nodes (RN) (excluding the IMN), treated with either hypofractionation or conventional fractionation. Table 1 summarises the patient specifications, including

treatment volumes and prescribed doses. Eight patients received moderate fractionation, with a prescribed dose of 40.05 Gy delivered in 15 fractions to either the breast or chest wall and regional nodes, while six patients received 40.05 Gy to the breast PTV and 48 Gy to the tumour bed simultaneously. The second prescription was conventional fractionation, delivering 50 Gy in 25 fractions for the breast or chest wall and regional lymph nodes, applied to six out of twenty patients.

Table 1. Treated volume, number of patients, prescription dose, and total planning target volume of the 20 enrolled patients.

Treated Volume	Number of Patients	Prescription Dose (Gy)/fx	PTVTotal (cc), Mean $\pm$ SD (Range)
Breast_R	3	40/15	1491 $\pm$ 336 (1254–1875)
Breast_R and TB	3	40–48/15	577 $\pm$ 105 (487–692)
Breast_L	3	40/15	1060 $\pm$ 238 (838–1311)
Breast_L and TB	3	40–48/15	817 $\pm$ 170 (671–1003)
ChestWall_L and RN	1	40/15	656
ChestWall_R and RN	2	50/25	513 $\pm$ 135 (418–608)
Breast_R and RN *	1	50/25	1484
Breast_R, TB and RN	1	40–48/15	776
Breast_L and RN *	3	50/25	1511 $\pm$ 468 (1014–1943)

* Sequencing boost treatment. L = left; R = right; TB = Tumor Bed; RN = regional nodes; PTV = planning target volume.

2.2. VMAT Planning

Computed tomography (CT) simulation with a 2.5 mm slice thickness (Discovery CT590 RT, GE Healthcare, Chicago, IL, USA) was performed for RT treatment planning to delineate the target volume and adjacent OARs according to European Society for Radiotherapy and Oncology (ESTRO) guidelines [16]. The radiation volumes comprised the clinical target volume (CTV), which included the whole breast or chest wall (CW), and for advanced-stage patients with supraclavicular and infraclavicular involvement, any portion of the axillary bed at risk was delineated as CTV_RN. The Planning Target Volumes (PTV) for the breast (PTV_Breast) and for regional lymph nodes (PTV_RN) were established by applying an isotropic margin expansion of 3 mm to CTV_Breast and CTV_RN, followed by a reduction of 2 mm from the skin and 5 mm from the ipsilateral lung, respectively. The total PTV for patients with a PTV_RN volume was generated by merging PTV_Breast and PTV_RNs. The clinically accepted plans for the retrospectively selected patients were optimised with an auto skin flash of a 10 mm margin at the time of their treatment, which is our clinical decision to ensure that at least 95% of the PTV receives 95% of the prescribed dose of either 40 or 50 Gy to the breast or chest wall and lymph nodes. The treatment plans for these patients were replanned by modifying the AF thickness to 0, 3, 5, 10, and 15 mm, utilising the same CT dataset, delineated contours, and prescribed treatment doses. The abbreviations AF0, AF3, AF5, AF10, and AF15 in this study represent the AF thicknesses employed during the optimisations of 0, 3, 5, 10, and 15 mm, respectively. Plans were developed for an Elekta Versa HD accelerator (MLC Agility with 80 leaves and 0.5 cm leaf width) utilising a Monaco v5.11 treatment planning system (Elekta AB, Stockholm, Sweden), employing a 4-partial-arc VMAT (4pVMAT) with 6 MV photons as shown in Figure 1. In this configuration, the isocenter was set inside of the fields' overlapping area, consisting of 4 arcs that each cover 50°, starting and ending at angles of 300°/20° for the central arcs and 90°/180° for the lateral arcs, thereby allowing a minimum overlap of 15° between the arcs as outlined by Poeta et al. [17].

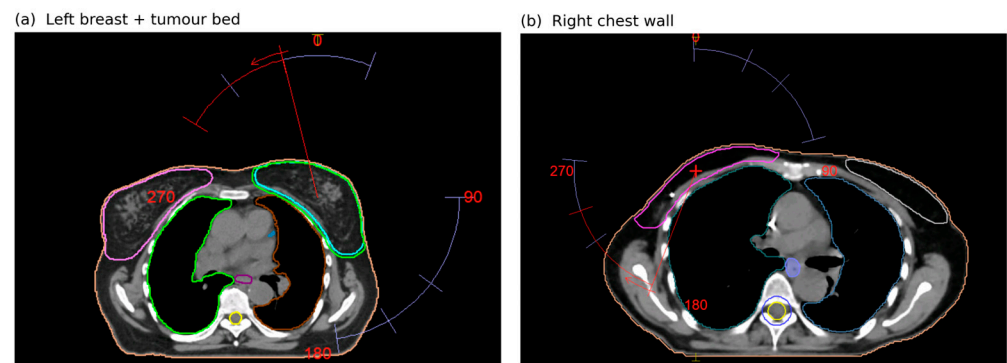


Figure 1. 4pVMAT beam geometry on representative transverse CT slices: (a) left breast + tumour bed and (b) right chest wall. Each panel shows the gantry-angle ring (0° , 90° , 180° , 270°) and the four 50° partial arcs (two medial, two lateral; $\geq 15^\circ$ overlap) per Poeta et al. [17], mirrored between left- and right-sided targets.

All treatment plans were standardised to ensure that 95% of the PTVTotal received at least 95% of the prescribed dose. In cases where the AF plans with varying thicknesses did not achieve this coverage after a second re-optimisation, the plans were normalised to meet the clinical acceptance criterion for PTVTotal. Specifically, the volume receiving more than 110% of the prescribed dose was constrained not to exceed 3 cm^3 , even if the maximum dose surpassed the clinical threshold of $D_{\text{max}} \leq 110\%$. This approach was preferred to mitigate the elevated modulation factor associated with increasing the optimisation weight of the PTVTotal. All plans were required to satisfy the same institutional OAR dose limits and the $V_{95\%} \geq 95\%$ target-coverage criterion prior to comparison; these clinical constraints were therefore held fixed as acceptance conditions across all AF settings, with autoflash thickness as the only variable, rather than treated as per-AF outcome measures. Table 1 summarizes the patient-specific PTV details, including the prescribed dose for each individual patient and the total PTV volume in cubic centimeters (cc). The critical structures used as OARs were delineated, including the heart, contralateral breast, left and right lungs, esophagus, and spinal cord for all involved patients, with the addition of the left anterior descending artery (LAD) for patients with left-sided breast sites as described by Offersen et al. [16] and Nielsen et al. [18].

The number of control points per arc, segment dimensions, and fluence smoothing—parameters that influence complexity metrics such as segment number, modulation factor (MF) and monitor unit (MU) efficiency—remained consistent in the reoptimised plans across all AF thicknesses for all patients. A standard deviation of 1% was used in Monte Carlo dose calculation with a dose grid of 3.0 mm. The replans were named VMAT_AFx, referring to the thickness of the AF (in mm) used in the optimisation. At the beginning of the inverse optimisation for each VMAT_AFx plan, the same optimisation template as VMAT_AF10—the accepted plan at the time of treatment—was used. Dose objectives and dose–volume constraints were defined according to institutional breast VMAT planning practice and established radiotherapy protocols, in line with published guidelines such as ASTRO, DEGRO, and UK practice reports for whole-breast and chest wall irradiation [19–26]. These constraints, summarised in Table 2, include limits for the ipsilateral lung, heart, LAD, and contralateral lung and breast. For the 50 Gy in 25 fractions schedule, conventional fractionation constraints were applied, while for 40 Gy in 15 fractions these were adapted to clinically comparable hypofractionated dose levels (e.g., V17 Gy instead of V20 Gy and V35 Gy instead of V40 Gy). Following optimisation and normalisation for PTV coverage as described above, all plans across the different AF thicknesses were reviewed to ensure compliance with clinical acceptance criteria, with particular emphasis on maintaining heart and contralateral organ doses as low as reasonably achievable (ALARA). When OAR dose

constraints were not satisfied after normalisation, plans underwent further iterative re-optimisation, with adjustments to both target volume objectives and OAR dose constraints to achieve compliance with the criteria defined in Table 2.

Table 2. Dose constraints for OARs for whole-breast/CW with or without regional lymph nodes irradiation using 50 Gy in 25 fx (2 Gy/fx) and 40 Gy in 15 fx (≈ 2.67 Gy/fx).

OAR	Metric	50 Gy/25 fx	40 Gy/15 fx
Ipsilateral Lung	V20 Gy/V17 Gy	$V20 \leq 25\%$	$V17 \leq 25\%$
	V10 Gy/V8 Gy	$V10 \leq 35\text{--}40\%$	$V8 \leq 35\text{--}40\%$
	V5 Gy/V4 Gy	$V5 \leq 55\text{--}65\%$	$V4 \leq 55\text{--}65\%$
	Mean lung dose	≤ 8 Gy (goal); ≤ 10 Gy (upper)	≤ 8 Gy (goal); ≤ 10 Gy (upper)
Heart	Mean dose	< 4 Gy (ideal ≤ 4 ; acceptable ≤ 5)	< 4 Gy (ideal ≤ 3.2 ; acceptable ≤ 4)
	V20 Gy/V17 Gy	$V20 \leq 10\%$	$V17 \leq 10\%$
LAD	V40 Gy/V35 Gy	$V40 \leq 5\%$	$V35 \leq 5\%$
	Dmax (0.035 cc)	≤ 20 Gy	≤ 17 Gy
	Dmean	< 15 Gy	< 15 Gy
Contralateral Breast	Mean dose	ALARA (≤ 1 Gy pref; ≤ 3 Gy accept.)	ALARA (≤ 1 Gy pref; ≤ 3 Gy accept.)
Contralateral Lung	Dmax	≤ 2.40 Gy (≤ 2.64 acceptable)	≤ 2.40 Gy (≤ 2.64 acceptable)
	V5 Gy/V4 Gy	$V5 \leq 10\%$ ($\leq 15\%$ acceptable)	$V4 \leq 10\%$ ($\leq 15\%$ acceptable)

OAR = organ at risk; CW = chest wall; LAD = left anterior descending artery; ALARA = as low as reasonably achievable.

2.3. Plan Complexity and Delivery Accuracy Analysis Patient Selection

Plan complexity and delivery accuracy were evaluated to investigate the impact of AF thickness on VMAT plan deliverability and to derive the correlation between changes in plan complexity metrics and delivery accuracy. For each patient, five treatment plans corresponding to different AF thicknesses (AF0, AF3, AF5, AF10, and AF15) were generated and analysed under identical planning conditions, with AF thickness being the only varying parameter. Plan complexity was quantified using parameters extracted from the treatment planning system, including modulation factor (MF), total MU, segment number, MU efficiency, and normalisation ratio. The modulation factor, defined as the ratio of total MU to the prescribed dose, expressed in MU/cGy, was used as a primary indicator of fluence modulation. Since all plans were normalised to achieve comparable PTV coverage, variations in MF reflect changes in modulation complexity associated with different AF thicknesses. The normalisation ratio, a scaling factor applied after optimisation to satisfy the PTV coverage constraint, was also inspected as a parameter of plan quality and dose uniformity. Together, these parameters provide a comprehensive assessment of plan modulation, delivery efficiency, and optimisation performance.

PSQA was performed for all plans using the IBA Matrixx Resolution 2D detector array system. Measured dose distributions were compared with the corresponding calculated dose distributions from the treatment planning system using gamma index analysis. Gamma passing rates (GPR) were evaluated using 3%/3 mm, 3%/2 mm, and 2%/2 mm criteria, corresponding to standard, intermediate, and strict levels of evaluation, respectively. Clinical acceptance thresholds were defined as $\geq 95\%$ for the 3%/3 mm criterion and $\geq 90\text{--}95\%$ for the 3%/2 mm criterion, while the 2%/2 mm criterion with a 10% dose threshold was used for more sensitive assessment of delivery accuracy.

The relationship between AF thickness, plan complexity metrics, and PSQA results was analysed across all patients to determine whether increasing AF thickness leads to changes in modulation characteristics and delivery performance. Given that each patient was evaluated under all five AF conditions, a paired statistical design was adopted. Data normality was assessed using the Shapiro–Wilk test. To evaluate the overall effect of AF

thickness on each metric across the five paired conditions, the non-parametric Friedman test was applied. Where a significant overall effect was identified, post hoc pairwise comparisons between individual AF levels were performed using the Wilcoxon signed-rank test, with Bonferroni correction applied for ten pairwise comparisons ($\alpha_{\text{corr}} = 0.005$). The Spearman rank correlation was used to assess monotonic associations between AF thickness (treated as a continuous variable, 0–15 mm) and each metric across all 100 plans, as well as inter-metric associations between complexity parameters and GPR. The Pearson product-moment correlation was additionally computed for complexity-vs-PSQA pairs to capture linear relationships. A two-sided p -value < 0.05 was considered statistically significant for individual tests, with Bonferroni-corrected thresholds reported separately for pairwise comparisons. All analyses were performed in Python (version 3.12.3) (SciPy (version 1.17.1) library) on the full dataset of $n = 20$ patients $\times$ 5 AF thicknesses = 100 plans.

To evaluate the consistency of the autoflash effect across clinical scenarios, a linear mixed-effects model was fitted to the primary endpoint (gamma passing rate at 3%/3 mm). The model included autoflash thickness (categorical, five levels: AF0, AF3, AF5, AF10, AF15) and clinical subgroup (categorical, three levels: breast-only, breast plus RNI, chest wall plus RNI) as fixed effects, with their interaction term, and patient as a random intercept to account for the repeated-measures structure of the data. Models were fitted by maximum likelihood; significance of the AF main effect, the subgroup main effect and the AF $\times$ subgroup interaction was assessed by likelihood-ratio tests against the corresponding nested models. The intraclass correlation coefficient (ICC) was computed from the random-effect variance to quantify the proportion of total variance attributable to between-patient differences.

3. Results

All treatment plans were normalised to ensure that 95% of the PTVTotal received at least 95% of the prescribed dose, with a normalisation ratio maintained close to unity across all AF levels (range: 1.029–1.032). Because every plan was renormalised to the same coverage criterion (PTVTotal V95% = 95%), target coverage—including the superficial component retained by the 2 mm skin-subtracted PTV margin—was held constant across plans by construction; the quantity that therefore reflects the influence of AF thickness on coverage is the normalisation ratio, i.e., the scaling required to satisfy that criterion. Its near-unity, essentially constant value across AF0–AF15 indicates that the optimiser reached full target coverage with equivalent efficiency at every AF thickness, without AF-dependent rescaling. The study population included 20 patients in three different clinical treatment groups: 12 patients treated with whole breast irradiation (with or without tumor bed boost), 5 patients treated with breast and regional lymph node irradiation, and 3 patients treated with chest wall and regional lymph node irradiation. The mean PTVTotal volumes were 985.9 ± 403.1 cc, 1358.5 ± 464.5 cc, and 560.8 ± 125.9 cc for entire breast, breast with regional lymph nodes, and chest wall with regional lymph nodes groups, respectively, reflecting the broad anatomical and volumetric diversity inherent to this patient population. Across the full cohort, the median PTVTotal was 920.5 cc (mean 1015.3 ± 453.1 cc; range: 418.0–1942.8 cc), underscoring the heterogeneous clinical conditions under which AF thickness optimisation was evaluated. This heterogeneity is clinically significant as it facilitates the assessment of the five evaluated AF thicknesses (0, 3, 5, 10, and 15 mm) across a representative range of target geometries and treatment complexities encountered in standard breast and chest wall radiation practice.

3.1. Effect of Automatic Skin Flash Thickness on Plan Complexity Metrics

3.1.1. Modulation Factor (MF)

The mean modulation factor was within a narrow range across the five AF settings (Table 3), with values of 4.77 ± 1.01 MU/cGy for AF0, 4.64 ± 0.93 MU/cGy for AF3, 4.65 ± 0.79 MU/cGy for AF5, 4.66 ± 0.68 MU/cGy for AF10, and 4.54 ± 0.86 MU/cGy for AF15. Mean MF was highest at AF0 and lowest at AF15, although the overall variation remained small relative to inter-patient differences. As shown in Figure 2a, MF exhibits substantial inter-patient variability across all AF thickness levels. Despite this variability, the median MF remains stable from AF0 to AF15, with no systematic shift across increasing AF margins.

Table 3. Plan complexity metrics across autoflash thicknesses (mean $\pm$ SD; $n = 20$ per group).

Metric	AF0	AF3	AF5	AF10	AF15
MF (MU/cGy)	4.77 ± 1.01	4.64 ± 0.93	4.65 ± 0.79	4.66 ± 0.68	4.54 ± 0.86
Total MU	1087 ± 185	1070 ± 212	1082 ± 162	1069 ± 144	1039 ± 168
Segments (n)	124.3 ± 39.4	124.0 ± 38.4	122.8 ± 38.5	124.0 ± 36.4	125.3 ± 38.3
MU efficiency	0.86 ± 0.12	0.85 ± 0.11	0.85 ± 0.11	0.84 ± 0.11	0.82 ± 0.12
Normalisation ratio	1.032 ± 0.023	1.031 ± 0.020	1.032 ± 0.022	1.030 ± 0.021	1.029 ± 0.018

MF: modulation factor (total MU per cGy of prescribed dose). MU efficiency: ratio of optimised to delivered MU as reported by the planning system. Normalisation ratio: post-optimisation scaling factor applied to satisfy PTV V95% $\geq 95\%$.

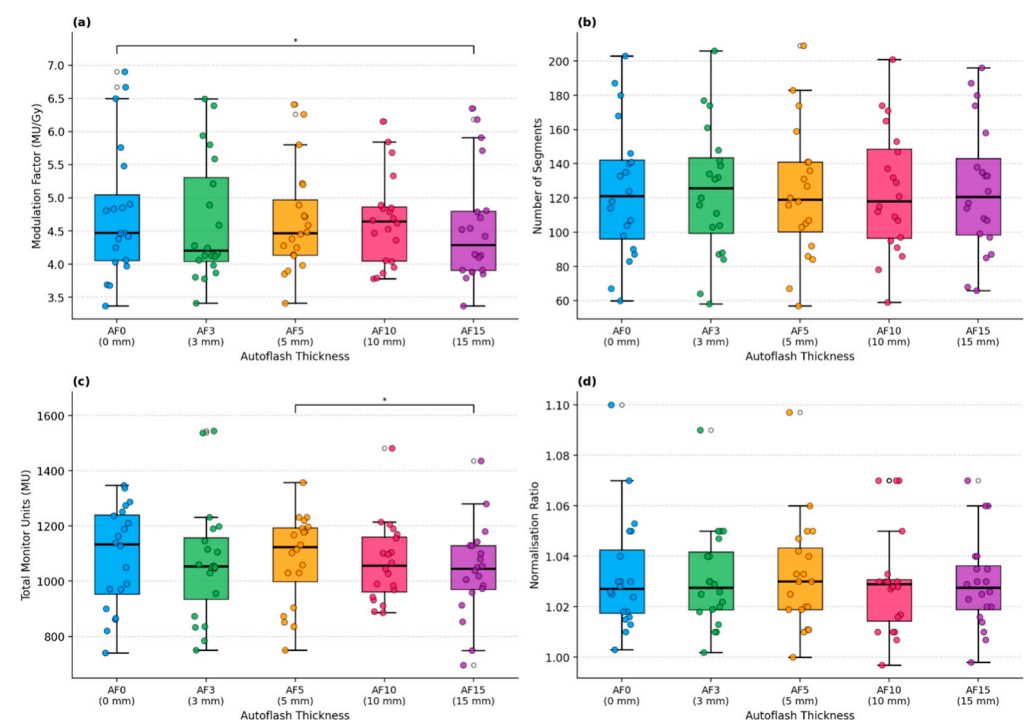


Figure 2. Distribution of plan complexity metrics across AF thicknesses (AF0–AF15). Box plots show median, interquartile range, and $1.5 \times$ IQR whiskers; points represent individual plans. (a) Modulation factor (MU/cGy), (b) number of segments, (c) total MU, and (d) normalisation ratio. * $p < 0.05$ for bracketed comparisons.

The Friedman test across the five AF levels did not reach statistical significance ($\chi^2 = 6.67$, $p = 0.154$; $n = 20$), indicating no overall effect of AF thickness on MF. Pairwise Wilcoxon signed-rank tests showed that only the AF0–AF15 comparison reached the conventional 0.05 threshold ($p = 0.040$, uncorrected); all remaining nine pairwise comparisons were non-significant (all $p \geq 0.165$), and this result did not survive correction.

for multiple comparisons. Pairwise comparisons against the clinically used AF10 reference were likewise non-significant for MF at every AF setting. Spearman rank correlation between AF thickness and MF was negligible and non-significant ($\rho = -0.06$, $p = 0.57$), consistent with the absence of a monotonic trend. Among the intermetric associations across all 100 plans, MF was most strongly correlated with total MU ($\rho = +0.67$, $p < 0.001$); this association is partly definitional, as MF is computed from total MU.

3.1.2. Segment Number and Total MU

The number of segments and the total MU showed minimal variation across the five AF settings (Table 3 and Figure 2b,c). Mean total MU varied only modestly across the AF range, from 1087 ± 185 at AF0 to 1039 ± 168 at AF15, corresponding to a cohort-mean spread of approximately 48 MU, which lies well within the inter-patient standard deviation (144–212 MU). The mean number of segments was even more stable, ranging from 122.8 ± 38.5 (AF5) to 125.3 ± 38.3 (AF15), a maximal inter-group difference of only 2.5 segments compared with inter-patient standard deviations of 36–39 segments. The box plots in Figure 2b,c show substantial overlap across AF levels, while the corresponding trends in Figure 3b,c show no systematic change with increasing AF thickness. The wide error bars in both representations confirm that target volume, anatomical complexity, and the inclusion of regional nodal irradiation drive these metrics far more than the AF margin.

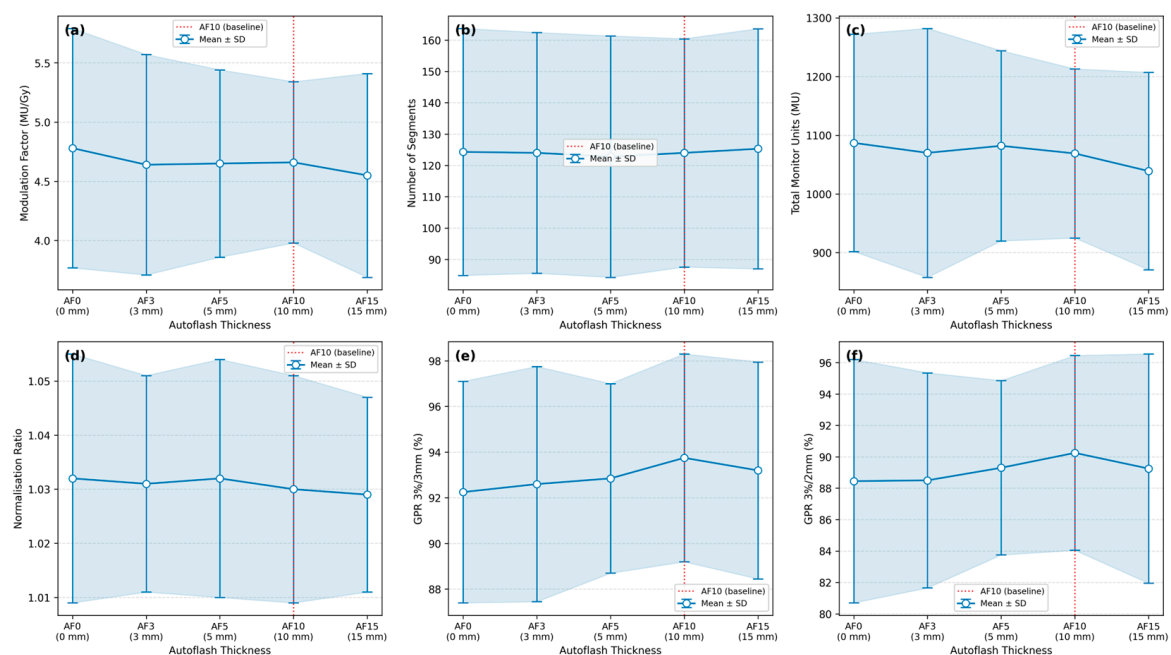


Figure 3. Distribution Mean $\pm$ SD trends across AF thickness levels. Plan complexity metrics (a–d) and PSQA GPR (e,f) are shown as a function of AF margin (AF0, AF3, AF5, AF10, AF15 mm; $n = 20$ per level). Open circles indicate cohort means, error bars and shaded bands represent ± 1 standard deviation, and the red dotted line marks AF10, the clinically used baseline. (a) Modulation factor (MU/cGy); (b) number of segments; (c) total MU; (d) normalisation ratio; (e) GPR at 3%/3 mm; (f) GPR at 3%/2 mm. Plan complexity metrics remain stable across all AF thicknesses, whereas GPRs increase up to AF10 and plateau at AF15.

Statistical testing supported the absence of an AF-dependent effect on either metric. The Friedman test across the five AF levels was non-significant for both Total MU ($\chi^2 = 3.50$, $p = 0.478$) and segment number ($\chi^2 = 4.02$, $p = 0.403$). Spearman rank correlations between AF thickness and the two metrics were correspondingly negligible and non-significant (Total MU: $\rho = -0.10$, $p = 0.317$; segments: $\rho = +0.001$, $p = 0.989$; Figure 4), confirming the absence of any monotonic trend across the 0–15 mm range. Pairwise Wilcoxon signed-rank tests

were non-significant for every AF pair for both metrics (Total MU: all $p \geq 0.06$; segments: all $p \geq 0.18$), including all comparisons against the clinically used AF10 reference. Across all 100 plans, Total MU was strongly correlated with the modulation factor ($\rho = +0.67, p < 0.001$), while segment number was inversely correlated with MU efficiency ($\rho = -0.45, p < 0.001$) and, to a lesser extent, with the normalisation ratio ($\rho = -0.31, p = 0.002$), reflecting the fact that plans with more apertures distribute dose over smaller fractional contributions per segment.

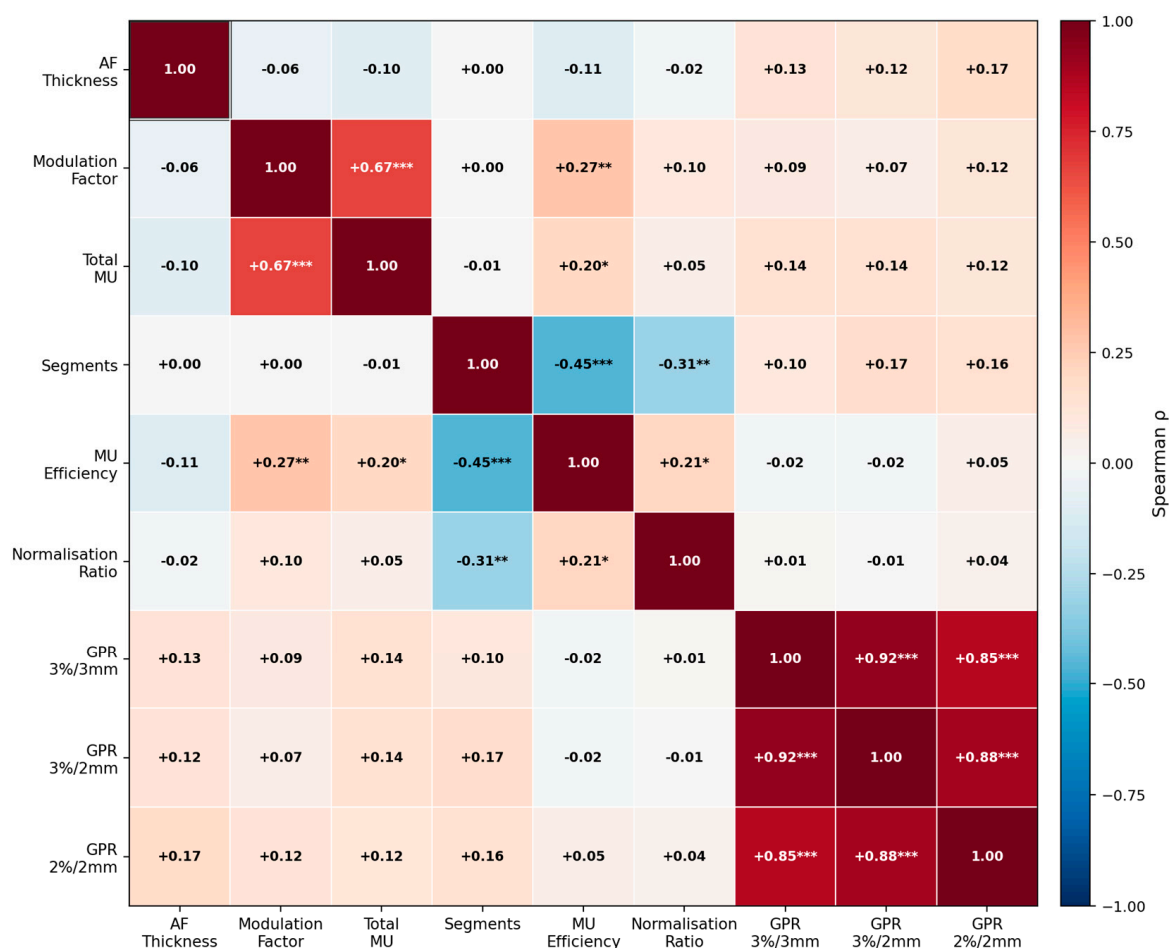


Figure 4. Spearman rank correlation matrix across all 100 VMAT plans (20 patients $\times$ 5 AF thicknesses). Cell values show Spearman ρ ; asterisks indicate significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). AF thickness shows no significant correlations with any plan complexity metrics ($|\rho| \leq 0.11$). The strongest inter-metric association is between modulation factor and total MU ($\rho = +0.67, p < 0.001$).

MU Efficiency and Normalisation Ratio: MU efficiency, defined as the ratio of optimised to delivered MU reported by the planning system, showed a small monotonic decrease with increasing AF thickness, from 0.86 ± 0.12 at AF0 to 0.82 ± 0.12 at AF15 (Table 3). The post-optimisation normalisation ratio applied to satisfy the PTV V95% $\geq 95\%$ coverage criterion was likewise nearly constant, ranging from 1.029 ± 0.018 (AF15) to 1.032 ± 0.023 (AF0), with all five cohort means falling within 0.003 of one another (Table 3; Figures 2d and 3d). A normalisation ratio of 1.030 corresponds to a uniform 3.0% upward rescaling of the optimised dose to bring the PTV V95% to 95%; values close to unity indicate that the optimiser achieved the coverage objective directly, with rescaling acting only as a small correction. All five cohort means fell within the institutional acceptable range (typically ≤ 1.05), with no AF level requiring substantial post hoc rescaling. For both

metrics, inter-patient variability dominated any AF-related differences, and the trend lines in Figure 3d remain essentially flat across the AF range.

The Friedman test across the five AF levels did not reach statistical significance for either metric (MU efficiency: $\chi^2 = 7.94$, $p = 0.094$; normalisation ratio: $\chi^2 = 4.26$, $p = 0.373$), and Spearman rank correlations with AF thickness were correspondingly negligible (MU efficiency: $\rho = -0.11$; normalisation ratio: $\rho = -0.02$; both $p > 0.05$; Figure 4). For the normalisation ratio, no pairwise Wilcoxon comparison reached significance (all $p > 0.30$); for MU efficiency, three uncorrected pairwise comparisons against AF15 reached the 0.05 threshold (lowest $p = 0.006$ for AF0 vs. AF15) but only one survived Bonferroni correction ($\alpha_{\text{corr}} = 0.005$), and none of the comparisons against the clinical AF10 reference reached significance for either metric (Figure 5d,e). Taken together, these results indicate that AF thickness has no clinically relevant effect on either MU efficiency or the normalisation ratio.

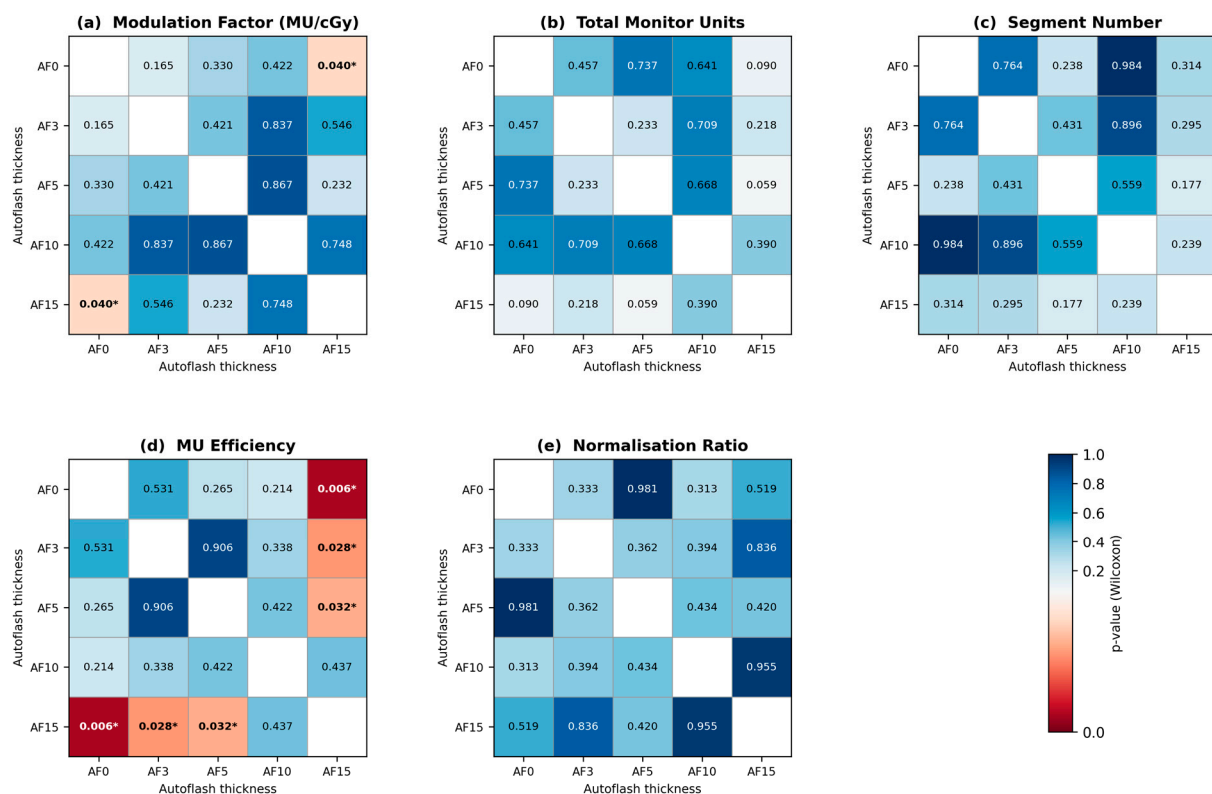


Figure 5. Pairwise Wilcoxon signed-rank test p -values for the five plan complexity metrics across AF thickness levels ($n = 20$ paired observations). (a) Modulation factor (MU/cGy); (b) total MU; (c) segment number; (d) MU efficiency; (e) normalisation ratio. Cell colour scales with significance (red = significant, green = non-significant); asterisks and bold font mark $p < 0.05$. None of the comparisons against the clinical AF10 baseline reached significance, indicating no meaningful effect of AF thickness on plan complexity.

Pairwise Wilcoxon signed-rank test p -values for all five complexity metrics are visualised in Figure 5. Across the fifty pairwise comparisons performed (ten per metric), only four reached the uncorrected 0.05 threshold (modulation factor: AF0 vs. AF15, $p = 0.040$; MU efficiency: AF0 vs. AF15, $p = 0.006$, AF3 vs. AF15, $p = 0.028$, AF5 vs. AF15, $p = 0.032$), and only the AF0–AF15 comparison for MU efficiency would survive Bonferroni correction. None of the comparisons against the clinically used AF10 reference reached significance for any complexity metric, confirming that AF thickness has no clinically meaningful effect on plan complexity over the 0–15 mm range.

3.2. Effect of Automatic Skin Flash Thickness on PSQA

Patient-specific QA was performed for all 100 plans using the IBA Matrixx Resolution (IBA Dosimetry, Schwarzenbruck, Germany) detector array, and GPRs were evaluated under three criteria of increasing strictness: the standard 3%/3 mm criterion (institutional clinical acceptance threshold $\geq 95\%$), the intermediate 3%/2 mm criterion ($\geq 90\%$), and the strict 2%/2 mm criterion (used for sensitive assessment of delivery accuracy).

Mean GPR increased with AF thickness up to AF10 and then plateaued at AF15 across all three gamma criteria (Table 4; Figures 3e,f and 6). This behaviour indicates that increasing AF thickness improves delivery robustness up to an optimal threshold (AF10), beyond which further increases do not yield additional dosimetric benefit.

Table 4. GPR across AF thicknesses under three evaluation criteria (Mean $\pm$ SD, $n = 20$).

AF Thickness	GPR 3%/3 mm (%)	GPR 3%/2 mm (%)	GPR 2%/2 mm (%)
AF0 (0 mm)	92.28 $\pm$ 4.83 §	88.46 $\pm$ 7.72 †	78.21 $\pm$ 7.69 †
AF3 (3 mm)	92.63 $\pm$ 5.11 ‡	88.54 $\pm$ 6.84 †	79.33 $\pm$ 7.88
AF5 (5 mm)	92.86 $\pm$ 4.14 †	89.35 $\pm$ 5.56	80.07 $\pm$ 6.82
AF10 (10 mm)	93.73 $\pm$ 4.55	90.32 $\pm$ 6.26	80.92 $\pm$ 8.48
AF15 (15 mm)	93.23 $\pm$ 4.70	89.27 $\pm$ 7.29	80.49 $\pm$ 8.20

† $p < 0.05$ vs. AF10 (Wilcoxon, uncorrected); ‡ $p < 0.01$ vs. AF10 (Wilcoxon); § $p < 0.001$ vs. AF10 (Wilcoxon). Clinical thresholds: $\geq 95\%$ for 3%/3 mm; $\geq 90\%$ for 3%/2 mm.

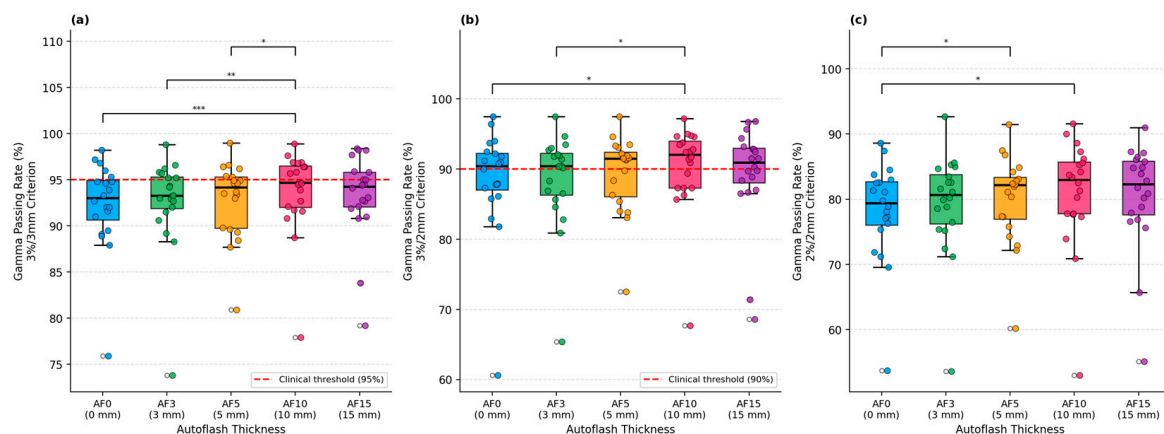


Figure 6. Effect of AF thickness on PSQA GPR across the five AF levels ($n = 20$ per level), evaluated under three criteria: 3%/3 mm (a), 3%/2 mm (b), and 2%/2 mm (c). Mean GPR increases from AF0 to AF10 and plateaus at AF15 across all three criteria; the AF0–AF10 and AF3–AF10 differences are statistically significant under the 3%/3 mm criterion after Bonferroni correction for ten pairwise comparisons. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

For the 3%/3 mm criterion the mean GPR rose from $92.28 \pm 4.83\%$ at AF0 to $93.73 \pm 4.55\%$ at AF10, with $93.23 \pm 4.70\%$ at AF15. The 3%/2 mm and 2%/2 mm criteria followed the same pattern (3%/2 mm: $88.45 \pm 7.72\%$ at AF0, $90.31 \pm 6.26\%$ at AF10, $89.27 \pm 7.29\%$ at AF15; 2%/2 mm: $78.21 \pm 7.69\%$ at AF0, $80.92 \pm 8.48\%$ at AF10, $80.48 \pm 8.20\%$ at AF15). Per-patient analysis showed that 17 of 20 patients improved their 3%/3 mm GPR between AF0 and AF10, with a mean within-patient gain of $+1.46\%$ (median $+1.35\%$, range -1.90% to $+3.70\%$). This consistent improvement across the majority of patients supports the robustness of the observed AF-dependent PSQA trend. The proportion of plans meeting the institutional clinical acceptance criterion of $\text{GPR} \geq 95\%$ at 3%/3 mm rose from 5/20 (25%) at AF0 to 9/20 (45%) at AF10 and 8/20 (40%) at AF15. When considering each patient individually, the AF thickness yielding the highest 3%/3 mm GPR was AF10 or AF15 in 14 of 20 patients.

The Friedman test confirmed a significant overall effect of AF thickness on GPR for all three criteria (3%/3 mm: $\chi^2 = 16.70$, $p = 0.002$; 3%/2 mm: $\chi^2 = 10.93$, $p = 0.027$; 2%/2 mm: $\chi^2 = 9.61$, $p = 0.048$; $n = 20$ paired observations per metric). Post hoc pairwise Wilcoxon signed-rank tests for the 3%/3 mm criterion identified AF0 vs. AF10 ($p = 0.0003$) and AF3 vs. AF10 ($p = 0.005$) as significant after Bonferroni correction for ten comparisons ($\alpha_{\text{corr}} = 0.005$); AF5 vs. AF10 ($p = 0.027$) and AF0 vs. AF15 ($p = 0.036$) reached the uncorrected 0.05 threshold but did not survive correction. Pairwise comparisons among AF5, AF10 and AF15 were non-significant for all three criteria, confirming the plateau of GPR at the higher AF settings (Figure 7). Spearman rank correlations between AF thickness and GPR were small and not statistically significant when AF was treated as a continuous variable, reflecting the plateau behaviour between AF10 and AF15, which is not well captured by rank-based monotonic analysis. The relationship between plan complexity and PSQA across all 100 plans is illustrated in Figure 4. Pearson product-moment correlations between complexity metrics and GPRs were small and consistently negative for total MU (MU vs. 3%/3 mm: $r = -0.21$, $p = 0.035$; MU vs. 3%/2 mm: $r = -0.21$, $p = 0.038$; MU vs. 2%/2 mm: $r = -0.20$, $p = 0.050$) and small and positive for segment number (segments vs. 3%/3 mm: $r = +0.20$, $p = 0.048$; segments vs. 3%/2 mm: $r = +0.25$, $p = 0.012$). Spearman rank correlations for the same pairs were uniformly small and positive ($r_s = +0.10$ to $+0.17$), and the apparent discrepancy between Pearson and Spearman correlations was driven by a single patient with consistently high MU and low GPR across all AF settings (Patient 8, MU = 1337–1544, GPR 3%/3 mm = 73.8–80.9%); excluding this patient brought the two correlation estimates into agreement, and overall trends across the cohort remained weak. The three gamma criteria were highly intercorrelated ($\rho = 0.85$ – 0.92 , $p < 0.001$ for all pairwise comparisons; Figure 4), confirming the internal consistency of the QA assessment. No systematic delivery failure was observed at any AF thickness, and all plans remained clinically deliverable.

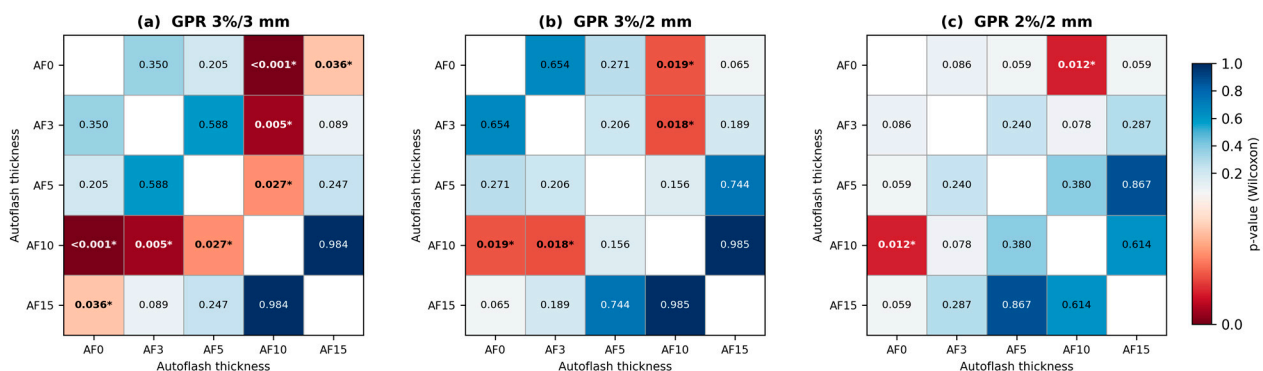


Figure 7. Pairwise Wilcoxon signed-rank test p -values for the three PSQA GPR criteria across AF thickness levels ($n = 20$ paired observations). (a) GPR 3%/3 mm; (b) GPR 3%/2 mm; (c) GPR 2%/2 mm. Cell colour scales with significance (red = significant, green = non-significant); asterisks and bold font mark $p < 0.05$. AF0 and AF3 differ significantly from the clinical AF10 baseline under the 3%/3 mm and 3%/2 mm criteria, while comparisons among AF5, AF10 and AF15 are non-significant under all three criteria, indicating a plateau of delivery accuracy at AF settings of 5 mm and above.

Taken together, these results provide evidence-based guidance for selecting the optimal AF margin in 4pVMAT breast and chest wall irradiation. From a clinical perspective, AF thickness should balance robustness and delivery accuracy as reflected by PSQA performance. Under the standard 3%/3 mm gamma criterion, improvements in GPR were observed up to AF10, beyond which no further benefit was evident. AF thicknesses of 5 mm and above achieved comparable compliance with the institutional acceptance threshold ($\geq 95\%$), whereas smaller margins (AF0–AF3) resulted in significantly lower passing rates

and may compromise delivery accuracy. A similar pattern was observed under stricter evaluation criteria (3%/2 mm and 2%/2 mm), which are relevant for more sensitive verification scenarios such as re-planning or accreditation. In these cases, a minimum AF thickness of 5 mm remains necessary, with AF10 consistently yielding the highest mean GPR across all criteria. As AF thickness showed no clinically meaningful impact on plan complexity metrics across the 0–15 mm range, margin selection can be guided primarily by delivery accuracy considerations.

Based on these findings, an AF margin of 10 mm is recommended as the optimal setting for 4pVMAT breast and chest wall plans using the Monaco Monte Carlo algorithm. A margin of 5 mm may be considered an acceptable lower bound, while margins below 5 mm should be avoided. Increasing the AF margin beyond 10 mm does not provide additional benefit in either delivery accuracy or plan complexity.

3.3. Consistency of the Autoflash Effect Across Clinical Subgroups

To address the heterogeneity of the cohort, which included breast-only ($n = 9$), breast plus regional nodal irradiation (RNI, $n = 8$) and chest wall plus RNI ($n = 3$) cases, we assessed whether the autoflash effect on PSQA differed across these clinical scenarios using a linear mixed-effects model on the primary endpoint (GPR 3%/3 mm). This analysis complemented the cohort-level Friedman and Wilcoxon tests reported in Section 3.2, which remain the primary statistical tests of the AF effect.

The model was fitted with patient as a random intercept and autoflash thickness, subgroup and their interaction as fixed effects. The intraclass correlation coefficient ($ICC = 0.84$) confirmed that between-patient variability dominated the data and justified the random-intercept specification. The direction and magnitude of the cohort-level AF effect were preserved in the mixed model. Although the omnibus AF main effect did not itself reach significance ($\chi^2 = 7.65$, $df = 4$, $p = 0.105$)—diluted by the AF15 plateau and the small AF3 and AF5 increments—the targeted AF10 vs. AF0 contrast remained significant: AF10 plans showed a +1.46% improvement in GPR over AF0 (95% CI +0.36 to +2.55, $p = 0.009$; Table 5), and AF15 plans a smaller non-significant gain (+0.96%, $p = 0.088$), consistent with the plateau described in Section 3.2.

Table 5. Linear mixed-effects analysis of GPR 3%/3 mm by AF thickness, stratified by clinical subgroup, ($n = 20$ patients, 100 plans).

AF Thickness	LMM Contrast vs. AF0 ($\Delta\%$, 95% CI)	Breast-Only ($n = 9$)	Breast + RNI ($n = 8$)	CW + RNI ($n = 3$)
AF0 (0 mm)	reference	92.88 $\pm$ 3.45	90.25 $\pm$ 6.16	95.87 $\pm$ 2.02
AF3 (3 mm)	+0.35 [−0.75, +1.45]	93.88 $\pm$ 2.70	89.85 $\pm$ 6.68	96.27 $\pm$ 2.30
AF5 (5 mm)	+0.59 [−0.51, +1.68]	93.76 $\pm$ 2.99	90.55 $\pm$ 4.74	96.33 $\pm$ 2.34
AF10 (10 mm)	+1.46 [+0.36, +2.55] §	94.87 $\pm$ 2.31	91.46 $\pm$ 6.17	96.37 $\pm$ 2.27
AF15 (15 mm)	+0.96 [−0.14, +2.05]	93.14 $\pm$ 4.42	92.42 $\pm$ 5.69	95.63 $\pm$ 2.71
AF10 − AF0 ($\Delta\%$)		+1.99 (9/9 improved)	+1.21 (6/8 improved)	+0.50 (2/3 improved)

Linear mixed-effects model fitted to all 100 plans: $GPR_{3,3} \sim AF \times Subgroup$, with patient as random intercept (maximum likelihood, statsmodels v0.14). 95% CI = estimate $\pm$ 1.96 $\times$ SE. Subgroup values are mean $\pm$ SD (%) of GPR 3%/3 mm. § LMM contrast vs. AF0 significant at $p < 0.01$. Likelihood-ratio tests: AF main effect $\chi^2 = 7.65$, $df = 4$, $p = 0.105$; subgroup main effect $\chi^2 = 4.12$, $df = 2$, $p = 0.127$; AF $\times$ subgroup interaction $\chi^2 = 10.90$, $df = 8$, $p = 0.207$ (non-significant interaction indicates no evidence that the AF effect differs between subgroups). Intraclass correlation coefficient (ICC) = 0.84.

The AF $\times$ subgroup interaction was non-significant ($\chi^2 = 10.90$, $df = 8$, $p = 0.207$), indicating that the AF effect did not differ statistically between subgroups; the subgroup main effect was likewise non-significant ($p = 0.127$). Per-subgroup AF10 − AF0 improvements were directionally consistent (Figure 8)—breast-only +1.99% (9/9 patients

improved), breast + RNI +1.21% (6/8 improved), CW + RNI +0.50% (2/3 improved)—with the smaller CW + RNI effect attributable to a ceiling at baseline (mean GPR at AF0 already $95.9\% \pm 2.0\%$).

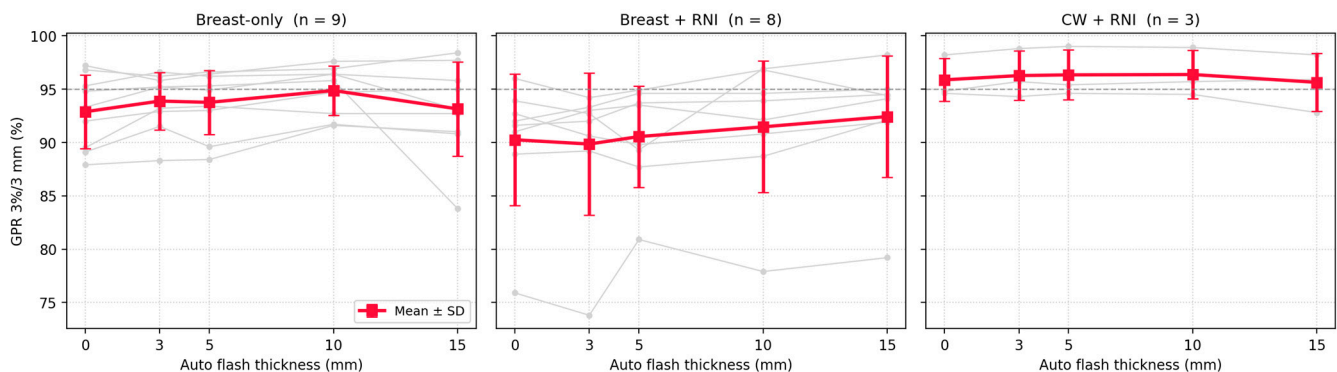


Figure 8. GPR 3%/3 mm versus autoflash thickness, stratified by subgroup (breast-only, $n = 9$; breast + RNI, $n = 8$; CW + RNI, $n = 3$). Grey lines show individual patients; red markers, mean $\pm$ SD; dashed line, the 95% threshold. Mean GPR rises to AF10 in all three subgroups.

These findings indicate that AF10 is the optimal flash margin across the clinical scenarios studied. Formal within-subgroup inference, particularly for CW + RNI ($n = 3$), remains underpowered; the mixed-effects analysis is therefore reported as the statistically appropriate test of subgroup consistency rather than as confirmatory evidence within each subgroup.

4. Discussion

The central question of this study was whether AF thickness, used purely as an optimisation construct in 4pVMAT breast and chest-wall planning with the Monaco Monte Carlo algorithm, has a measurable effect on plan complexity and on patient-specific delivery accuracy. Our results provide a clear, quantitative answer. Across AF thicknesses from 0 to 15 mm, plan complexity is statistically invariant—Friedman tests on modulation factor (MF), total MU, segment number, MU efficiency, and the post-optimisation normalisation ratio were all non-significant, and none of the pairwise Wilcoxon comparisons against the clinical AF10 reference reached significance for any complexity metric (Figure 5)—while GPR improves monotonically up to AF10 and plateaus thereafter (Figures 3 e,f and 6). To our knowledge, this is the first measurement-based confirmation of this behaviour in the Monaco/Versa HD/Agility MLC combination, and it has direct implications for how the AF margin should be selected in routine 4pVMAT breast and chest-wall planning.

VMAT for breast and chest-wall irradiation is most commonly delivered using one of two beam-design philosophies: a continuous partial arc covering the target from medial to posterior entry, or paired short arcs that mimic conventional tangential fields and form a “butterfly” geometry with avoidance sectors. The 4pVMAT geometry used here as illustrated in Figure 1, originally described by Poeta et al. [17] as a four-partial-arc split-VMAT technique with two medial arcs ($300^\circ \rightarrow 20^\circ$) and two lateral arcs ($90^\circ \rightarrow 180^\circ$) of 50° length each and a minimum 15° overlap between arcs, sits between these two approaches, combining the conformality of partial-arc planning with the contralateral sparing of tangential-style geometry. Although Poeta et al. originally introduced 4pVMAT to manage breath-hold time during deep inspiration breath-hold (DIBH) treatment without automatic beam-interruption devices, in our institution, the geometry has been adopted independently of DIBH on the basis of plan-quality and delivery-efficiency considerations. The choice of a partial-arc rather than full-arc geometry is consistent with the broader dosimetric concern, raised by Zhang et al. [7] in their h-IMRT versus VMAT comparison

for hypofractionated whole-breast irradiation, that wide full-arc geometries can spread substantial low-dose volumes into contralateral lung and breast that conventional indicators such as V20Gy or mean heart dose may not capture. Importantly, the four-partial-arc geometry constrains the angular range over which AF can affect aperture shape: each arc enters through the breast/chest-wall sector, so AF can only modify the apertures that face the skin surface. This is the geometric reason why, in our data, the optimiser absorbs the AF expansion through aperture-shape changes localised to the build-up region rather than through any global increase in modulation.

Within this beam-design context, the AF-dependent improvement in GPR was not accompanied by any deterioration in plan complexity. Mean total MU varied only modestly across the AF range, from 1087 ± 185 at AF0 to 1039 ± 168 at AF15, a cohort-mean spread well within the inter-patient standard deviation. Mean segment number was even more stable (122.8 ± 38.5 at AF5 to 125.3 ± 38.3 at AF15, a maximal inter-group difference of 2.5 segments against inter-patient standard deviations of 36–39). The post-optimisation normalisation ratio was nearly constant across the AF range (1.029–1.032), confirming that the optimiser achieved the PTV V95% $\geq 95\%$ coverage objective directly at every AF thickness without substantial post hoc rescaling. MU efficiency showed only a small monotonic drift from 0.86 ± 0.12 at AF0 to 0.82 ± 0.12 at AF15, with only the AF0–AF15 comparison surviving Bonferroni correction. Increasing the AF margin from 0 to 10 mm therefore improves delivery accuracy without any clinically meaningful cost in modulation, MU, segment count, MU efficiency, or normalisation behaviour. This decouples the historical concern that larger flash margins might inflate plan complexity, and supports AF10 as the optimal margin on both delivery-accuracy and plan-quality grounds. Hubley et al. [15] recommended a 7–10 mm flash region as a compromise between robustness and plan quality, but their assessment was based on plan-quality metrics alone, without patient-specific delivery verification. Our findings extend that work in two ways: we confirm directly with PSQA measurements that AF10 yields the highest GPRs across all three evaluation criteria— $93.73 \pm 4.55\%$ at 3%/3 mm, $90.32 \pm 6.26\%$ at 3%/2 mm, and $80.92 \pm 8.48\%$ at 2%/2 mm—and we show that this improvement is achieved without any measurable impact on conventional complexity metrics.

A plausible mechanism for the observed AF-dependent improvement in GPR without complexity inflation lies in how inverse optimisers handle the build-up region in superficial targets. As Arbor et al. [27] demonstrated using a GATE/Geant4 Monte Carlo toolkit for VMAT breast surface dosimetry, target volumes close to the patient surface can drive the optimiser to generate very high fluences in the build-up region in order to satisfy coverage constraints—precisely the region where treatment-planning-system algorithms have known limitations and where measurement–calculation agreement is most uncertain. Extending the optimisation envelope into a virtual flash region relieves the optimiser of this constraint: the fluence used to cover the superficial part of the target is shaped beyond the skin, and the steep fluence gradients that would otherwise sit on the build-up region are pushed outside the patient. The aperture sequence required to achieve this is not appreciably more modulated, which is exactly what we observed—MF, total MU, segment count, MU efficiency, and the normalisation ratio are all statistically indistinguishable across AF0–AF15. AF therefore re-shapes apertures locally at the breast surface without forcing the optimiser to introduce additional modulation elsewhere in the arcs. The same mechanism explains the GPR plateau beyond AF10: once the AF margin exceeds the geometric extent over which the optimiser was attempting to push fluence into the build-up region, further extension of the virtual envelope no longer changes the aperture-shape decisions that drive PSQA performance. The optimiser is no longer fluence-limited at the surface, and additional virtual margin only modifies fluence in regions that are anyway removed before final

dose calculation. This interpretation is quantitatively consistent with Lizondo et al. [11], who showed that beyond a virtual bolus thickness equal to the CTV–PTV margin plus 5 mm, additional thickness produces diminishing returns in robustness and increasing penalties in normalisation behaviour. For the effective skin-direction margin used here (3 mm CTV–PTV expansion reduced 2 mm from skin), this corresponds to a virtual flash of approximately 10 mm, in close agreement with our PSQA-determined optimum. The mechanism is further reinforced by Oh et al. [28], who analysed 285 breast VMAT plans on the same combination of treatment platform (Monaco TPS, Versa HD linac, Agility 5 mm MLC) used in our study and found that most conventional complexity metrics correlate only weakly with GPR; only aperture-shape descriptors such as leaf sequence variability, plan-averaged beam area, and edge-area metrics emerged as moderate predictors. In that context, our finding that AF improves PSQA without modifying conventional complexity metrics is consistent with the broader behaviour of this TPS/MLC combination: PSQA performance is determined more by aperture geometry than by aggregate modulation. The methodological reliability of our PSQA setup is also supported by Yasser et al. [29], who reported GPR above 95% at 3%/3 mm for mono-isocentric breast VMAT plans verified with the same IBA Matrixx Resolution detector, the same Monaco TPS, and the same Agility MLC used in the present study.

Our PSQA-based identification of AF10 as the optimum converges with multiple independent dosimetric studies in different planning environments. Nicolini et al. [12] originally formalised the pseudo-skin-flash strategy by extending the optimisation envelope into a 10 mm soft-tissue-equivalent virtual region around the breast and removing it before final dose calculation; Lizondo et al. [11] refined the approach by varying both bolus thickness and Hounsfield Unit (HU) assignment, recommending a CTV–PTV margin plus 5 mm thickness with HU values around -500 to -400 for robustness against shifts of up to 10 mm in the breathing direction. Tyran et al. [30] confirmed dosimetric safety by recalculating plans on second CT scans acquired during treatment, showing that V95% of the breast CTV remained 98.9% with virtual bolus versus 92.6% without, and that the virtual bolus did not degrade the initial plan. Ugurlu et al. [31] further validated the approach with thermoluminescent dosimetry on an anthropomorphic phantom, observing that 0.5 cm virtual boluses preserved plan quality while progressively thicker boluses elevated PTV maximum dose without compensating dosimetric benefit. Hubley et al. [15] reached a strikingly similar conclusion using a different planning architecture (static-angle ports within a VMAT arc, RAD platform) and a different HU value (-350), recommending 7–10 mm of automatic skin flash. Our PSQA-based finding that AF10 maximises GPR while leaving complexity unchanged provides independent measurement-based confirmation in a third planning environment—Monaco with 4pVMAT geometry—and connects the optimisation-side observations of these prior studies to the delivery-side reality of patient-specific QA. An important methodological clarification arises from He et al. [32], who reported that VMAT with 10 mm physical bolus is less robust to 3 mm setup errors than VMAT with 5 mm physical bolus and recommended a 5 mm bolus for postmastectomy VMAT. That conclusion is specific to physical bolus, which remains in the beam path during dose calculation and during delivery. AF is a virtual structure used only during optimisation and removed before final dose calculation; the physical attenuation and obliquity effects that drive the He et al. result therefore do not apply to AF, which can be increased to 10 mm without the robustness penalty associated with a thicker physical bolus. Our results also complement broader breast-VMAT robustness analyses showing that no-flash VMAT plans lose target coverage even under modest CBCT-based anatomical changes [33], and that robustly optimised VMAT outperforms manual-flash VMAT both in robustness against simulated organ motion and in planner-time efficiency [34]. The AF

mechanism evaluated here can be regarded as an automated counterpart to these strategies, providing the geometric robustness of a flash margin without the planner-dependent variability of manual MLC editing—and our data show that this automation is achieved without inflating plan complexity. Our work also aligns with the Dutch national consensus on breast radiotherapy plan evaluation reported by Hurkmans et al. [35], which defines target evaluation criteria of $D_{98\%} \geq 95\%$, $D_{2\%} \leq 107\%$, and Dmean within 99–101% of the prescription, all of which were maintained across AF0–AF15 in the present study. The Dutch consensus achieves robustness against the build-up region by clipping CTV and PTV 5 mm below the skin and reporting on the clipped volume; the AF approach used here addresses the same problem from the opposite direction, by extending the optimisation envelope outwards rather than restricting evaluation inwards. Both strategies converge on the same operational principle: the build-up region should not be allowed to drive the optimiser.

The combination of unchanged plan complexity and improved GPR up to AF10 supports adopting 10 mm as the standard AF margin for routine 4pVMAT breast and chest-wall planning in Monaco. $AF \geq 5$ mm is the practical lower bound consistent with acceptable PSQA performance; margins below 5 mm—including AF0 and AF3—show measurably degraded GPRs, particularly under the more stringent 3%/2 mm and 2%/2 mm criteria, with 17 of 20 patients improving their 3%/3 mm GPR between AF0 and AF10 and the proportion of plans meeting the institutional acceptance criterion of $GPR \geq 95\%$ at 3%/3 mm rising from 25% at AF0 to 45% at AF10. Beyond AF10, additional flash thickness up to 15 mm did not yield further improvement in PSQA and produced a borderline downward drift in MU efficiency that did not survive correction for multiple comparisons; AF15 should therefore be regarded as the upper bound rather than as a routine setting. These recommendations are relevant to both the 50 Gy/25 fraction and 40 Gy/15 fraction schedules included in our cohort, the latter of which has become the international standard for early-stage breast cancer following the FAST-Forward 5-year analysis confirming non-inferiority of moderate hypofractionation in efficacy and toxicity [36]. Plan-complexity stability is particularly desirable for hypofractionated regimens in which each fraction delivers a larger biologically effective dose, because any per-fraction PSQA failure carries proportionally greater dosimetric consequence. It should be emphasised that the gamma passing rate quantifies delivery accuracy, namely the agreement between calculated and delivered dose, and is not in itself a measure of geometric or dosimetric robustness against setup error and motion, nor of clinical outcome; the AF-thickness optimisation reported here should therefore be interpreted within the delivery-accuracy domain. The absolute GPR values observed here (mean 92–94% at 3%/3 mm) are characteristic of breast and chest-wall VMAT rather than indicative of delivery problems. Patient-specific QA for these sites is intrinsically demanding: the target abuts the skin and build-up region, the anatomy is large, curved and adjacent to low-density lung, and a 2D detector array has limited angular and spatial sampling for tangential, obliquely incident arcs. Interpreted against the AAPM TG-218 [37] framework—a universal tolerance limit of 95% (prompting investigation) and an action limit of 90% (below which a plan is clinically unacceptable)—18 of 20 AF10 plans (90%) exceeded the 90% action limit, with only two below it, and no plan showed systematic delivery failure. The clinical relevance of these results therefore lies not in the absolute passing rate, which is detector- and site-dependent, but in the internally controlled, relative effect of AF thickness: increasing AF to 10 mm raised the mean GPR and shifted the distribution upward (plans meeting $\geq 95\%$ rising from 25% to 45%, and plans meeting the $\geq 90\%$ action limit from 75% to 90%), moving more plans toward and above the acceptance threshold without any complexity penalty.

This study has several limitations. The results are specific to one TPS algorithm (Monaco v5.11 X-ray Voxel Monte Carlo) and one MLC model (Elekta Versa HD with 5 mm Agility leaves), and Oh et al. [28] showed that complexity–GPR relationships are sensitive to the specific TPS/MLC combination. Multicentric validation across different dose calculation algorithms (e.g., collapsed-cone convolution, Acuros XB, analytical anisotropic algorithm) and different AF calculation methods on various TPS platforms would strengthen the generalizability of the AF10 recommendation. Because the optimal AF thickness depends on the interplay between the dose-calculation algorithm and the MLC delivery characteristics, the specific value identified here (AF10) should be re-established on other platforms, even though the qualitative pattern of improved delivery accuracy up to a plateau without increased plan complexity is expected to be platform-transferable. We did not perform in vivo skin-dose measurements; the present analysis quantifies the impact of AF on plan complexity and PSQA but does not directly measure skin dose at the patient surface, which remains challenging to compute accurately because conventional TPS algorithms have known limitations in the build-up region [27]. The present study likewise did not directly evaluate plan robustness against setup error or respiratory motion, which we identify as future work. We also excluded IMN coverage from the cohort; AF margin behaviour for IMN-inclusive plans, where PTV_IMN often includes lung and where the Dutch consensus permits a relaxed $D_{98\%} \geq 90\%$ criterion [35], remains an open question. PSQA was performed with a 2D detector array, which cannot fully capture three-dimensional dose discrepancies. Future work should extend this analysis across multiple TPS/algorithm combinations with different AF implementations, include plans with IMN coverage, incorporate in vivo skin-dose verification, and explicitly evaluate plan robustness against simulated setup errors using the methodology described by He et al. [32] and Chan et al. [34]. Finally, the subgroup sample sizes—particularly chest wall plus RNI ($n = 3$)—were too small to achieve the power needed for formal within-subgroup inference, so the Section 3.3 mixed-effects analysis was designed to test the consistency of the autoflash effect across subgroups via the $AF \times \text{subgroup}$ interaction term rather than through three separate underpowered tests. The non-significant interaction therefore indicates an absence of evidence for subgroup heterogeneity rather than proof of a uniform effect, and confirmation will require larger, multi-institutional cohorts.

5. Conclusions

This study systematically evaluated the impact of AF thickness on plan complexity and delivery accuracy in 4-partial-arc VMAT breast and chest-wall radiotherapy. Increasing AF thickness significantly improves delivery accuracy while having no clinically meaningful effect on plan complexity metrics. The 10 mm AF margin emerges as the optimal compromise: it provides the highest GPR across all three evaluation criteria without inflating modulation, MU, segment count, MU efficiency, or normalisation behaviour. Its superiority is consistent across three independent planning environments—Eclipse with virtual bolus [11], the RAD platform with automatic skin flash [15], and Monaco with AF in 4pVMAT geometry (the present study)—supporting AF10 as a robust, platform-transferable recommendation for routine breast and chest-wall VMAT planning. A 5 mm AF margin can be considered an acceptable lower bound, while margins below 5 mm should be avoided as they significantly compromise PSQA delivery accuracy without offering any compensating advantage in plan complexity.

Author Contributions: Conceptualization, investigation, and writing—original draft preparation, S.S. and C.C.; methodology, software, validation, formal analysis, resources, data curation, writing—review and editing, visualization, supervision, and project administration, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This retrospective dosimetric study involved re-planning analysis of anonymised treatment planning data without use of patient images or identifiable information. Approval was granted by the hospital scientific committee.

Informed Consent Statement: Individual informed consent was waived for this retrospective dosimetric study as it involved computational re-analysis of anonymised treatment planning parameters without patient images or identifiable data, as approved by the hospital scientific committee.

Data Availability Statement: The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

VMAT	Volumetric-modulated arc therapy
AF	Automatic flash
PSQA	Patient-specific quality assurance
4pVMAT	Four-partial-arc volumetric-modulated arc therapy
MLC	Multileaf collimator
MF	Modulation factor
MU	Monitor units
GPR	Gamma passing rate
IMRT	Intensity-modulated radiation therapy
3D-CRT	Three-dimensional conformal radiation therapy
OAR	Organ at risk
PTV	Planning target volume
CT	Computed tomography
RT	Radiotherapy
LAD	Left anterior descending artery
CTV	Clinical target volume
CW	Chest wall
RN	Regional nodes
QA	Quality assurance
HU	Hounsfield unit
TPS	Treatment planning system
ALARA	As low as reasonably achievable
DIBH	Deep inspiration breath-hold
IMN	Internal mammary nodes

References

1. Cozzi, L.; Lohr, F.; Fogliata, A.; Franceschini, D.; De Rose, F.; Filippi, A.R.; Guidi, G.; Vanoni, V.; Scorsetti, M. Critical appraisal of the role of volumetric modulated arc therapy in the radiation therapy management of breast cancer. *Radiat. Oncol.* **2017**, *12*, 200. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Fogliata, A.; De Rose, F.; Franceschini, D.; Stravato, A.; Seppälä, J.; Scorsetti, M.; Cozzi, L. Critical appraisal of the risk of secondary cancer induction from breast radiation therapy with volumetric modulated arc therapy relative to 3D conformal therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2018**, *100*, 785–793. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Hunte, S.O.; Clark, C.H.; Zyuzikov, N.; Nisbet, A. Volumetric modulated arc therapy (VMAT): A review of clinical outcomes—What is the clinical evidence for the most effective implementation? *Br. J. Radiol.* **2022**, *95*, 20201289. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Kendall, R.; Robinson, T.; Reed, V.; Kanke, J.; Sosa, A.; Nelson, C.; Swanson, D.; Villa, M.; Bloom, E. Why is volumetric modulated arc therapy not considered the standard of care for locoregional radiation therapy for breast cancer patients? *Adv. Radiat. Oncol.* **2025**, *10*, 101728. [\[CrossRef\]](#) [\[PubMed\]](#)

5. Yu, P.C.; Wu, C.J.; Nien, H.H.; Lui, L.T.; Shaw, S.; Tsai, Y.L. Tangent-based volumetric modulated arc therapy for advanced left breast cancer. *Radiat. Oncol.* **2018**, *13*, 236. [[CrossRef](#)] [[PubMed](#)]
6. Liu, H.; Chen, X.; He, Z.; Li, J. Evaluation of 3D-CRT, IMRT and VMAT radiotherapy plans for left breast cancer based on clinical dosimetric study. *Comput. Med. Imaging Graph.* **2016**, *54*, 1–5. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, M.; Zhang, F.X.; Yang, X.L.; Liang, Q.; Liu, J.; Zhou, W.B. Comparative dosimetric study of h-IMRT and VMAT plans for breast cancer after breast-conserving surgery. *Transl. Oncol.* **2024**, *47*, 102012. [[CrossRef](#)] [[PubMed](#)]
8. Bartolucci, L.; Adrien, C.; Goudjil, F.; Amessis, M.; El Amine, W.; Fourquet, A.; Mazal, A.; Kirova, Y. Dosimetric comparison of four high performance techniques for irradiation of breast cancer patients. *Cancer Radiother.* **2021**, *25*, 254–258. [[CrossRef](#)] [[PubMed](#)]
9. Yoo, S.; Blitzblau, R.; Yin, F.F.; Horton, J.K. Dosimetric comparison of preoperative single-fraction partial breast radiotherapy techniques: 3D CRT, noncoplanar IMRT, coplanar IMRT, and VMAT. *J. Appl. Clin. Med. Phys.* **2015**, *16*, 183–207. [[CrossRef](#)] [[PubMed](#)]
10. Michalski, A.; Atyeo, J.; Cox, J.; Rinks, M. Inter- and intra-fraction motion during radiation therapy to the whole breast in the supine position: A systematic review. *J. Med. Imaging Radiat. Oncol.* **2012**, *56*, 499–509. [[CrossRef](#)] [[PubMed](#)]
11. Lizondo, M.; Latorre-Musoll, A.; Ribas, M.; Carrasco, P.; Espinosa, N.; Coral, A.; Jornet, N. Pseudo skin flash on VMAT in breast radiotherapy: Optimization of virtual bolus thickness and HU values. *Phys. Med.* **2019**, *63*, 56–62. [[CrossRef](#)] [[PubMed](#)]
12. Nicolini, G.; Fogliata, A.; Clivio, A.; Vanetti, E.; Cozzi, L. Planning strategies in volumetric modulated arc therapy for breast. *Med. Phys.* **2011**, *38*, 4025–4031. [[CrossRef](#)] [[PubMed](#)]
13. Fogliata, A.; Seppälä, J.; Reggiori, G.; Lobefalo, F.; Palumbo, V.; De Rose, F.; Franceschini, D.; Scorsetti, M.; Cozzi, L. Dosimetric trade-offs in breast treatment with VMAT technique. *Br. J. Radiol.* **2017**, *90*, 20160701. [[CrossRef](#)] [[PubMed](#)]
14. Virén, T.; Heikkilä, J.; Myllyoja, K.; Koskela, K.; Lahtinen, T.; Seppälä, J. Tangential volumetric modulated arc therapy technique for left-sided breast cancer radiotherapy. *Radiat. Oncol.* **2015**, *10*, 79. [[CrossRef](#)] [[PubMed](#)]
15. Hubley, E.; Koger, B.; Li, T.; Salerno, M.; Scheuermann, R.M.; Dong, L.; Teo, B.K.K. Automatic skin flash optimization in breast and chestwall VMAT with static angle modulated ports: Effect of HU and flash margin size on plan quality and robustness. *J. Appl. Clin. Med. Phys.* **2025**, *26*, e70036. [[CrossRef](#)]
16. Offersen, B.V.; Boersma, L.J.; Kirkove, C.; Hol, S.; Aznar, M.C.; Biete Sola, A.; Kirova, Y.M.; Pignol, J.P.; Remouchamps, V.; Verhoeven, K.; et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. *Radiother. Oncol.* **2015**, *114*, 3–10. [[CrossRef](#)] [[PubMed](#)]
17. Poeta, S.; Jourani, Y.; De Caluwé, A.; Van den Begin, R.; Van Gestel, D.; Reynaert, N. Split-VMAT technique to control the deep inspiration breath hold time for breast cancer radiotherapy. *Radiat. Oncol.* **2021**, *16*, 77. [[CrossRef](#)]
18. Nielsen, M.H.; Berg, M.; Pedersen, A.N.; Andersen, K.; Glavicic, V.; Jakobsen, E.H.; Jensen, I.; Josipovic, M.; Lorenzen, E.L.; Nielsen, H.M.; et al. Delineation of target volumes and organs at risk in adjuvant radiotherapy of early breast cancer: National guidelines and contouring atlas by the Danish Breast Cancer Cooperative Group. *Acta Oncol.* **2013**, *52*, 703–710. [[CrossRef](#)] [[PubMed](#)]
19. Chirilă, M.E.; Kraja, F.; Marta, G.N.; Neves Junior, W.F.P.; de Arruda, G.V.; Gouveia, A.G.; Franco, P.; Poortmans, P.; Ratos, I. Organ-sparing techniques and dose-volume constraints used in breast cancer radiation therapy—Results from European and Latin American surveys. *Clin. Transl. Radiat. Oncol.* **2024**, *46*, 100752. [[CrossRef](#)] [[PubMed](#)]
20. Duma, M.N.; Baumann, R.; Budach, W.; Dunst, J.; Feyer, P.; Fietkau, R.; Haase, W.; Harms, W.; Hehr, T.; Krug, D.; et al. Heart-sparing radiotherapy techniques in breast cancer patients: A recommendation of the breast cancer expert panel of the German Society of Radiation Oncology (DEGRO). *Strahlenther. Onkol.* **2019**, *195*, 861–871. [[CrossRef](#)] [[PubMed](#)]
21. Ghannam, Y.; Le Scodan, R.; Rivera, S.; Kirova, Y.; Pasquier, D.; Hennequin, C.; Bourgier, C.; Cutuli, B.; Richard-Tallet, A. Radiotherapy of breast cancer: 2025 update. *Cancer Radiother.* **2025**, *29*, 104767. [[CrossRef](#)] [[PubMed](#)]
22. Jimenez, R.B.; Abdou, Y.; Anderson, P.; Barry, P.; Bradfield, L.; Bradley, J.A.; Heras, L.D.; Khan, A.; Matsen, C.; Rabinovitch, R.; et al. Postmastectomy radiation therapy: An ASTRO/ASCO/SSO clinical practice guideline. *Pract. Radiat. Oncol.* **2025**, *15*, 549–571. [[CrossRef](#)] [[PubMed](#)]
23. Thomsen, M.S.; Berg, M.; Zimmermann, S.; Lutz, C.M.; Makocki, S.; Jensen, I.; Hjelstuen, M.H.B.; Pensold, S.; Hasler, M.P.; Jensen, M.B.; et al. Dose constraints for whole breast radiation therapy based on the quality assessment of treatment plans in the randomised Danish Breast Cancer Group (DBCG) HYPO trial. *Clin. Transl. Radiat. Oncol.* **2021**, *28*, 118–123. [[CrossRef](#)] [[PubMed](#)]
24. Zhang, Y.; Huang, Y.; Ding, S.; Yuan, X.; Shu, Y.; Liang, J.; Mao, Q.; Jiang, C.; Li, J. A dosimetric and radiobiological evaluation of VMAT following mastectomy for patients with left-sided breast cancer. *Radiat. Oncol.* **2021**, *16*, 171. [[CrossRef](#)] [[PubMed](#)]
25. De Rose, F.; Carmen De Santis, M.; Lucidi, S.; Ray Colciago, R.; Marino, L.; Cucciarelli, F.; La Rocca, E.; Di Pressa, F.; Lohr, F.; Vanoni, V.; et al. Dose constraints in breast cancer radiotherapy: A critical review. *Radiother. Oncol.* **2025**, *202*, 110591. [[CrossRef](#)] [[PubMed](#)]

26. Smith, B.D.; Bellon, J.R.; Blitzblau, R.; Freedman, G.; Haffty, B.; Hahn, C.; Halberg, F.; Hoffman, K.; Horst, K.; Moran, J.; et al. Radiation therapy for the whole breast: Executive summary of an American Society for Radiation Oncology (ASTRO) evidence-based guideline. *Pract. Radiat. Oncol.* **2018**, *8*, 145–152. [[CrossRef](#)] [[PubMed](#)]
27. Arbor, N.; Gasteuil, J.; Noblet, C.; Moreau, M.; Meyer, P. A GATE/Geant4 Monte Carlo toolkit for surface dose calculation in VMAT breast cancer radiotherapy. *Phys. Med.* **2019**, *61*, 112–117. [[CrossRef](#)] [[PubMed](#)]
28. Oh, D.H.; Choi, J.H.; Ryu, H.; Chun, M. Quantification of treatment plan deliverability in breast volumetric-modulated arc therapy with Agility multi-leaf collimator. *In Vivo* **2024**, *38*, 2254–2260. [[CrossRef](#)] [[PubMed](#)]
29. Yasser, R.; Rajaa, S.; Verma, T.; Pandey, V.; Yassine, H.; Mohammed, E.; Ezzahra, C.F.; Zahra, C.F.; Nihad, E.; Khalid, E.O.; et al. Pre-treatment verification of carcinoma breast VMAT plan based on mono-isocentric technique: Assessment of the combined fields feature of new 2D MatriXX arrays resolution. *Onkol. Radioter.* **2022**, *16*, 14–20.
30. Tyran, M.; Tallet, A.; Resbeut, M.; Ferre, M.; Favrel, V.; Fau, P.; Moureau-Zabotto, L.; Darreon, J.; Gonzague, L.; Benkemouche, A.; et al. Safety and benefit of using a virtual bolus during treatment planning for breast cancer treated with arc therapy. *J. Appl. Clin. Med. Phys.* **2018**, *19*, 463–472. [[CrossRef](#)] [[PubMed](#)]
31. Ugurlu, T.B.; Koksak Akbaş, C.; Ibis, K.; Becerir, H.B. The effect of using virtual bolus on VMAT plan quality for left-sided breast cancer patients. *Appl. Radiat. Isot.* **2022**, *189*, 110416. [[CrossRef](#)] [[PubMed](#)]
32. He, Y.; Chen, S.; Gao, X.; Fu, L.; Kang, Z.; Liu, J.; Shi, L.; Li, Y. Robustness of VMAT to setup errors in postmastectomy radiotherapy of left-sided breast cancer: Impact of bolus thickness. *PLoS ONE* **2023**, *18*, e0280456. [[CrossRef](#)] [[PubMed](#)]
33. Gleeson, I. Comparing the robustness of different skin flash approaches using wide tangents, manual flash VMAT, and simulated organ motion robust optimization VMAT in breast and nodal radiotherapy. *Med. Dosim.* **2022**, *47*, 264–272. [[CrossRef](#)] [[PubMed](#)]
34. Chan, R.C.K.; Ng, C.K.C.; Hung, R.H.M.; Li, Y.T.Y.; Tam, Y.T.Y.; Wong, B.Y.L.; Yu, J.C.K.; Leung, V.W.S. Comparative study of plan robustness for breast radiotherapy: Volumetric modulated arc therapy plans with robust optimization versus manual flash approach. *Diagnostics* **2023**, *13*, 3395. [[CrossRef](#)] [[PubMed](#)]
35. Hurkmans, C.; Duisters, C.; Peters-Verhoeven, M.; Boersma, L.; Verhoeven, K.; Bijker, N.; Crama, K.; Nuver, T.; van der Sangen, M. Harmonization of breast cancer radiotherapy treatment planning in the Netherlands. *Tech. Innov. Patient Support Radiat. Oncol.* **2021**, *19*, 26–32. [[CrossRef](#)] [[PubMed](#)]
36. Brunt, A.M.; Haviland, J.S.; Wheatley, D.A.; Sydenham, M.A.; Alhasso, A.; Bloomfield, D.J.; Chan, C.; Churn, M.; Cleator, S.; Coles, C.E.; et al. Hypofractionated breast radiotherapy for 1 week versus 3 weeks (FAST-Forward): 5-year efficacy and late normal tissue effects results from a multicentre, non-inferiority, randomised, phase 3 trial. *Lancet* **2020**, *395*, 1613–1626. [[CrossRef](#)] [[PubMed](#)]
37. Miften, M.; Olch, A.; Mihailidis, D.; Moran, J.; Pawlicki, T.; Molineu, A.; Li, H.; Wijesooriya, K.; Shi, J.; Xia, P.; et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218. *Med. Phys.* **2018**, *45*, e53–e83. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.