Investigation on the impact of the leaf trailing effect using the Halcyon integrated platform system

José Fernando Pérez Azorín
Medical Physics and Radiation Protection Department, Gurutzeta-Cruces University Hospital, E-48903 Barakaldo, Spain
Biocruces Health Research Institute, E-48903 Barakaldo, Spain

Jordi Saez
Department of Radiation Oncology, Hospital Clínic de Barcelona, 08036 Barcelona, Spain

Luis Isaac Ramos Garcia
Department of Oncology, Clínica Universidad de Navarra, University of Navarra, E-31008 Pamplona, Spain

Victor Hernandez
Department of Medical Physics, Hospital Sant Joan de Reus, IISPV, 43204 Tarragona, Spain
Universitat Rovira i Virgili, Tarragona, Spain

Version typeset March 22, 2022

Author to whom correspondence should be addressed. email: jfpazorin@gmail.com

Abstract

Purpose: The double-stacked design of the Halcyon multileaf collimator (MLC) presents new challenges for treatment planning systems (TPSs). The leaf trailing effect has recently been described as the result of the interplay between the fluence transmitted through the leaf tip ends of each MLC layer. This effect makes the dosimetric leaf gap (DLG) dependent on the distance between the leaves of different layers (trailing distance) and is not adequately modeled by the Eclipse TPS. The purpose of our study was to investigate and report the dose discrepancies produced by these limitations in clinical plans and to explore how these discrepancies can be mitigated and avoided.

Methods: The integrated platform with the Halcyon v2 system, Eclipse and Aria v15.6, was used. The dose discrepancies were obtained with EPID images and the portal dosimetry software and validated using radiochromic film dosimetry. The results for the AIDA commissioning test and for nine selected clinical beams with the sliding window intensity modulated radiotherapy (dIMRT) technique were thoroughly analyzed and presented. First, the DICOM RT plans were exported and the fluences were computed using different leaf tip models, and then were compared. Second, the detailed characteristics of the corresponding leaf sequences were investigated. Finally, modified DICOM RT plans were created in which the non-collimating (backup) leaves were retracted 2 mm to increase the leaf trailing distance, the modified plans were imported back into the TPS and the measurements were repeated. Dedicated in-house tools were developed in Python to carry out all analyses.

Results: Dose discrepancies greater than 10% and regions of gamma failure were found in both the AIDA test and clinical beams using static-gantry dIMRT. Fluence analysis highlighted that the discrepancies were due to limitations in the MLC model implemented in the TPS. Analysis of leaf sequences indicated that regions of failure were associated with very low leaf speeds and virtually motionless leaves within the beam aperture. Some of these discrepancies were mitigated by increasing the trailing distance of the non-collimating leaves without affecting the beam aperture, but this strategy was not possible in regions where the leaves from both layers actively defined the beam aperture.

Conclusions: Current limitations of the MLC model in Eclipse produced discrepancies between calculated and delivered doses in clinical beams that caused plan-specific quality assurance failures and interruptions in the clinical workflow. Careful evaluation of the clinical plans produced by Eclipse for the Halcyon is recommended, especially for static gantry dIMRT treatments. Some characteristics of leaf sequences are problematic and should be avoided in clinical plans and, in general, a better leaf tip model is needed. This is particularly important in adaptive radiotherapy treatments, where the accuracy and reliability of TPS dose calculations are of the utmost importance.

The table of contents is for drafting and refereeing purposes only. Note that all links to references, tables and figures can be clicked on and returned to calling point using cmd[on a Mac using Preview or some equivalent on PCs (see View - go to on whatever reader).

Contents

I. Introduction

A new platform for intensity modulated radiotherapy (IMRT) treatment delivery, the Halcyon™system, was introduced by Varian (Varian Medical System, Inc. Palo Alto, California, USA) in 2017. The system features a unique design with a ring-mounted linear accelerator that provides an efficient delivery of IMRT and volumetric modulated arc therapy (VMAT) treatments because of its high leaf and gantry speed (5 cm/s and 24 deg/s, respectively)[1,2,3]. For beam shaping and fluence modulation, the Halcyon incorporates an innovative dual-layer MLC with stacked and staggered rounded leaves with a width of 10 mm at the isocenter plane. This design provides an effective resolution of 5 mm at the isocenter while simultaneously minimizing interleaf leakage. Several authors have reported that the system complies with recommendations from international guidelines[4,5,6,7,8,9] and, in general, good agreement has been found in patient-specific quality assurance (QA) tests, end-to-end verifications, and external audits[2,3,10]. However, stacking two banks and using rounded leaf ends creates new challenges for proper modeling in treatment planning systems (TPSs)[11].

Several authors have reported that dose calculation accuracy depends on a proper beam characterization and a good MLC model[7,8,12,13]. In fact, the current MLC model in the Eclipse TPS (Varian Medical Systems) is dependent on three parameters: MLC transmission (T), dosimetric leaf gap (DLG) and tongue and groove width[14]. Dosimetric characterization of the Halcyon's dual-layer MLC system has been investigated by various authors[11,15,16]. Kim et al. (2019)[15] reported that the measured DLG was different for each layer and also when the aperture was defined by both layers simultaneously. Similarly, Lim et al. (2019)[16] studied the leaf end effect of the distal layer and reported a DLG of -0.19 mm at a depth of 10 cm. The same authors assessed the clinical accuracy of the leaf tip model for a selection of ten clinically representative plans in two different versions of the Eclipse TPS (Varian Medical Systems, Inc. Palo Alto, CA, USA) by comparing the calculated doses with point doses measured using an ion chamber. They obtained a mean dose discrepancy of -1% for plans evaluated with the latest TPS version.

A noteworthy aspect of the Eclipse beam model for the Halcyon system is that it is supplied with preconfigured beam data including fixed values for MLC transmission (0.47% per layer) and DLG (0.1 mm) applied to both layers, as well as tongue and groove widths (0.40 mm for the distal layer and 0.56 mm for the proximal layer). None of these values can

be modified by the user[17].

Hernandez et al.(2021)[11] recently investigated the interplay between the two layers and showed that the distance between the leaves of each MLC layer, i.e., the trailing distance, has a high impact on the photon fluence transmitted through the leaf tips and hence on measured DLG values. They found that measured DLG is dependent on the trailing distance: it shows a sharp increase for low values and finally levels out for trailing distances around 5 mm. This produces dose deviations as great as 10% for low trailing distances and sweeping gaps of 5 mm calculated with Eclipse. Miyasaka et al.(2022)[18] recently evaluated sequences of clinical VMAT plans and found no dosimetric consequences associated to the trailing effect. However, to the best of our knowledge, the impact of this effect on clinical dMLC plans has not yet been investigated.

In this work we present several cases of discrepancies between calculations in clinical dMLC plans produced by the Eclipse TPS and measurements that can be explained by poor modeling of the leaf tip and the leaf trailing effect. The goals of the study were to describe the situations where such discrepancies are found, to investigate their causes, and to discuss how these situations can be mitigated and solved.

II. Materials and Methods

II.A. Halcyon framework and description of test cases

In this investigation we report on the Halcyon v2 with SX2 MLC system and on version 15.6 of the Analytical Anisotropic Algorithm (AAA), smart Leaf Motion Calculator (smartLMC), and Photon Optimizer (PO) of the Eclipse TPS. Since the Halcyon system was commissioned in September 2019, more than one thousand patients have been treated, with approximately 70% of the treatments being delivered with VMAT and the remaining 30% with dIMRT.

We present the discrepancies found in a representative clinical beam and a test case. The clinical beam was a dIMRT beam that corresponding to a field used in a breast cancer treatment with involvement of supraclavicular nodes using the sliding window technique. Our analysis and discussion focus on this individual clinical case. However, eight additional clinical cases, corresponding to dIMRT treatments with similar discrepancies, are provided

as Supporting Information.

The test case presented is the AIDA test, which is routinely used for commissioning the portal dosimetry package (PDIP)[14]. It consists of an optimal fluence map provided by Varian with five rectangular slabs of different widths, a height of 3 cm and a separation of 2 cm between them. Fluence intensity is set to 1 inside the slabs and 0 outside. This optimal fluence for the AIDA test was imported into Eclipse ,the leaf sequence was calculated by the smartLMC algorithm, and the test was delivered with 370 MU and a dose rate of 600 MU/min (leaf speeds $<= 5\ cm/s$).

All plans were evaluated using the portal dosimetry algorithm (PDIP)[14], which is the AAA algorithm for Halcyon. Comparisons between predicted and measured dose distributions were performed with a local gamma metric of 2%-2 mm. Finally, to rule out any dose-response effects of the EPID, the AIDA test was also measured using radiochromic film dosimetry[19] and compared with the calculated dose distribution. The radiochromic film was placed at the depth of maximum dose in a water-equivalent phantom.

II.B. Comparison of fluence maps computed with different leaf tip models

In order to investigate if the observed dose differences could be associated with limitations in the MLC model, two different fluence maps were produced for each case considering different leaf tip models.

The first fluence map was with a leaf tip model equivalent to the one implemented in Eclipse. Thus, a constant $\mathrm{DLG_{const}} = 0.1$ mm was assumed and all leaf positions were retracted by half the DLG value (0.05 mm). To improve the resolution between control points (CP), new CPs were added by splitting the interval between CPs in 100 equal parts and using linear interpolation. The fluence was then computed at each interpolated CP by assigning a fluence value of 1 to the open regions within the beam aperture, 0.004 to regions shielded by a single MLC and zero to regions below both MLC layers. The total fluence map, φ_{const}, was finally computed by summing up all the partial fluences from each interpolated CP.

The second fluence map was computed using a variable DLG defined as a function of the

trailing distance t. For that purpose, the distance between leaves in different layers (trailing distance t) was first calculated at each CP and for each effective 5 mm leaf. Next, the same procedure described above was applied. This fluence map was named φ_{var}.

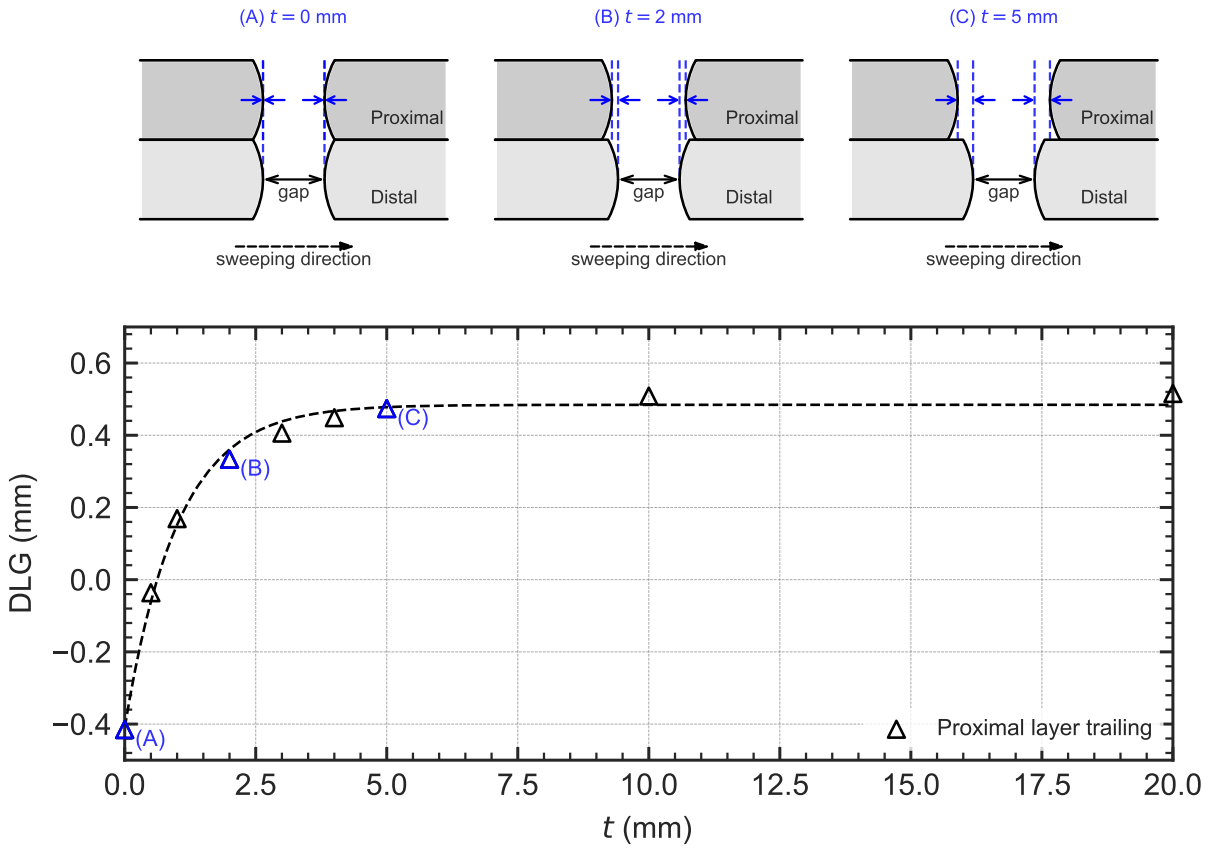


Figure 1: Sketch showing an MLC gap with three trailing distances t for proximal leaves acting as trailing (or backup) leaves (upper plot). The lower plot shows the dosimetric leaf gap (DLG) plotted as a function of the trailing distance. Letters (A, B, C) are used to identify each leaf arrangement and its corresponding DLG value.

The dependence of the DLG on the trailing distance between the two MLC layers (traling effect) is illustrated in Figure 1 . The trailing distance affects the photon transmission through the leaf tips and hence the measured DLG value, which depends on the distance between leading and back-up leaves (trailing distance).

To obtain the DLG as a function of the trailing distance t, the agreement with results reported for other Halcyon systems was verified measuring the trailing sweeping gap tests[11] and an analytical fit of the DLG was obtained using the empirical equation:

$$\mathrm{DLG}(t) = \alpha\left(1 - e^{-\beta t}\right) + \gamma \quad , \tag{1}$$

where α, β and γ are fitting parameters that depend on the layer acting as a collimating layer. The values obtained for the proximal layer were $\alpha = 0.89(6)\,\mathrm{mm}$, $\beta = 0.99(1)\,\mathrm{mm}^{-1}$ and γ = -0.41(6) mm and for the distal layer were $\alpha = 0.77(5)\,\mathrm{mm}$, $\beta = 1.2(2)\,\mathrm{mm}^{-1}$ and γ = -0.41(4) mm. This function provided a very good fit (shown in Supporting Materials), with all residuals $< 0.03\,\mathrm{mm}$.

Finally, the two fluence maps were compared using the percentage of differences defined as:

$$\mathrm{Diff}(\%) = \frac{\varphi_{\mathrm{var}} - \varphi_{\mathrm{const}}}{\varphi_{\mathrm{const}}} \quad . \tag{2}$$

Note that the only difference between fluence maps φ_{const} and φ_{var} was the leaf tip model used. This comparison therefore indicated the impact of the leaf tip model on TPS calculations , which is useful to investigate whether experimental discrepancies could be explained by limitations in the leaf tip model.

II.C. Analysis of leaf speeds

To analyze the characteristics of the leaf sequence, the RT plan was exported from the TPS in DICOM format and an in-house Python program was created to extract the leaf positions from both MLC layers, the gantry angle and the meterset weight at each control point. This information was used to determine the treatment time between each pair of CPs [20]. The leaf speed between each pair of CPs was then calculated and a 2D map of leaf speeds, $S(x, y)$, was computed for each leaf bank as:

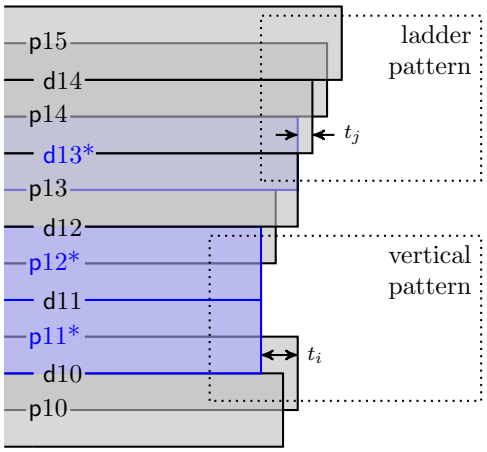
$$S(x, y) = \sum_{i=1}^{N-1} v_i(x, y) \quad , \tag{3}$$

where N is the total number of CPs, v_i is the speed (cm/s) between CPs i and $i+1$ and x, y are the spatial coordinates swept by each leaf with the y axis taken to be perpendicular to leaf motion.

II.D. Modified RT plan

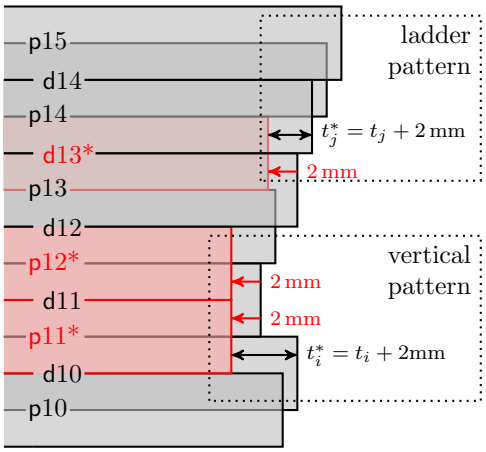
With the information obtained from the exported DICOM RT plans, two types of leaves were identified with respect to their role in primary fluence collimation: *collimating* and *non-collimating* leaves. A given leaf in any layer is considered to be a non-collimating or trailing leaf when its DICOM position is more retracted than its pair of overlapping of leaves from the other layer. A sketch illustrating this concept is depicted in Figure 2. In the latest implementation of the Halcyon, v2, both distal and proximal leaves can be collimating and non-collimating.

(a) Original plan.



* Non-collimating leaves.

(b) Modified plan.



* Non-collimating leaves shifted to increase trailing distance.

Figure 2: Sketch illustrating a leaf pattern from an original plan and the modified plan after applying the algorithm that detects and retracts non-collimating leaves. In the original plan (a) the backup or non-collimating leaves that did not define the beam aperture were identified. In the modified plan (b) the non-collimating leaves were further retracted to increase the trailing distance by 2 mm.

To investigate the effect of the trailing distance t on the delivered dose, trailing distances for non-collimating leaves were modified using an in-house Python program. The program identified non-collimating leaves at each CP and further retracted their position by 2 mm with respect to the leading leaves. This shift was selected because, as can be seen in Figure 1 , 2 mm is sufficient to avoid the steep increase in the DLG(t) curve and would produce a DLG close to its plateau. Since only non-collimating leaves were moved, the beam aperture

in the modified plan remained unaltered. This procedure is depicted in Figure 2. After modifying the RT plan, the new plan was imported back into Eclipse and the predicted portal image was recalculated, portal images were reacquired and the analysis was repeated.

III. Results

We found evident dose discrepancies between calculations and measurements in approximately 10-12% of the sliding window plans, but not in the VMAT plans. On average, the problematic dMLC plans had six beams and two of them were affected. The total percentage of sliding window beams with dose discrepancies was therefore around 3−4%. The differences in a test case (the AIDA test) and a representative clinical beam are reported and analyzed below. Additional clinical cases are provided as Supporting Material.

III.A. Analysis of the original plans

The gamma analysis results for both the AIDA test and the selected clinical case are shown in Figure 3. For the AIDA test, vertical bands indicating gamma failures ($\gamma > 1$) were clearly identified on the four rectangles at approximately similar distances from their right border (see Figure 3a). A crossline dose profile taken through the lower rectangle shows that the Eclipse TPS overestimated the dose by as much as 18% (see Supplementary Figure S3) in the region of gamma failure. Measurements for this test were repeated with radiochromic film and the same pattern with vertical bands within the homogeneous fluence region was observed, which confirmed the results obtained with portal dosimetry (shown in Supporting Information Figure S2).

The gamma map for the clinical beam displays regions of gamma failure with two distinct patterns. First, vertical straight bands with measured doses lower than calculated doses (cold spots, as in the AIDA test). Second, areas with a ladder-like pattern, where the measured doses were greater than calculated doses (hot spots). This is illustrated with a dose profile that includes both types of regions (Figure 3b). These dose discrepancies were as great as ± 20% (see Supplementary Figure S4).

A strict local gamma 2%-2 mm criterion was used to better identify the failing regions. However, most dose discrepancies were in high dose regions and were greater than 10%;

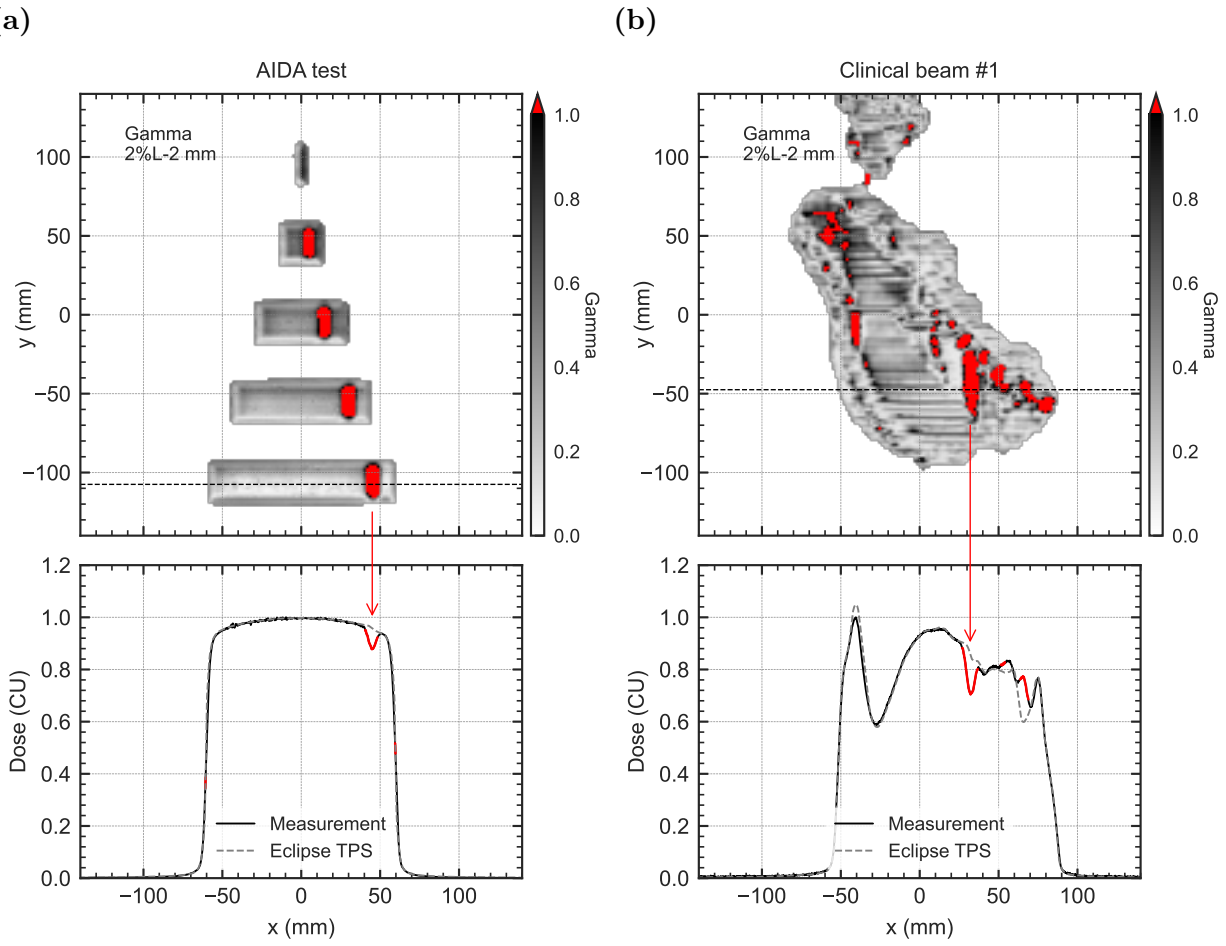


Figure 3: Comparison between the Eclipse and portal dosimetry system. Gamma maps (2% local, 2 mm) and dose profiles along the dotted lines are given for (a) the AIDA test and (b) the clinical case. Red regions indicate gamma values greater than 1.

hence, they would also fail with global gamma 3%-2mm.

III.B. Comparison of fluence maps computed with different leaf tip models

The constant DLG value used to generate φ_{const} was 0.1 mm, the same fixed value used in the Eclipse TPS. The fluence maps φ_{var} were computed using the DLG_{var} given in Eq. (1). Figure 4 shows the comparisons between both fluences, including fluence difference maps as described in Eq. (2) (upper row), and fluence profiles along the dashed lines (lower row).

Interestingly, the gamma map of φ_{var} and φ_{const} for a 2% local and 1 mm criteria, shown in Figure 4, closely reproduced the experimental gamma maps illustrated in Figure 3. The

only difference between the fluence maps φ_{var} and φ_{const} was the leaf tip model used in the fluence computation. Therefore, this high spatial correspondence clearly points out to the leaf tip model as the reason for the experimental discrepancies found.

For the AIDA test, the fluence values obtained using DLG_{var} were lower than those obtained using $\text{DLG}_{\text{const}}$ in the same locations where vertical bands of failure were measured (cold spots in Figure 4a, middle row). The fluence difference was around 18% and had a similar shape but larger discrepancies than the cold spots shown in Figure 3a. For the clinical case (Figure 4, middle row), the fluence difference map was more complex and showed DLG_{var} created regions with either higher or lower values than the ones found with $\text{DLG}_{\text{const}}$. These regions appeared in the difference map as large hot and cold spots (in ladder patterns and vertical bands, respectively). In the cold vertical bands, differences of around −30%

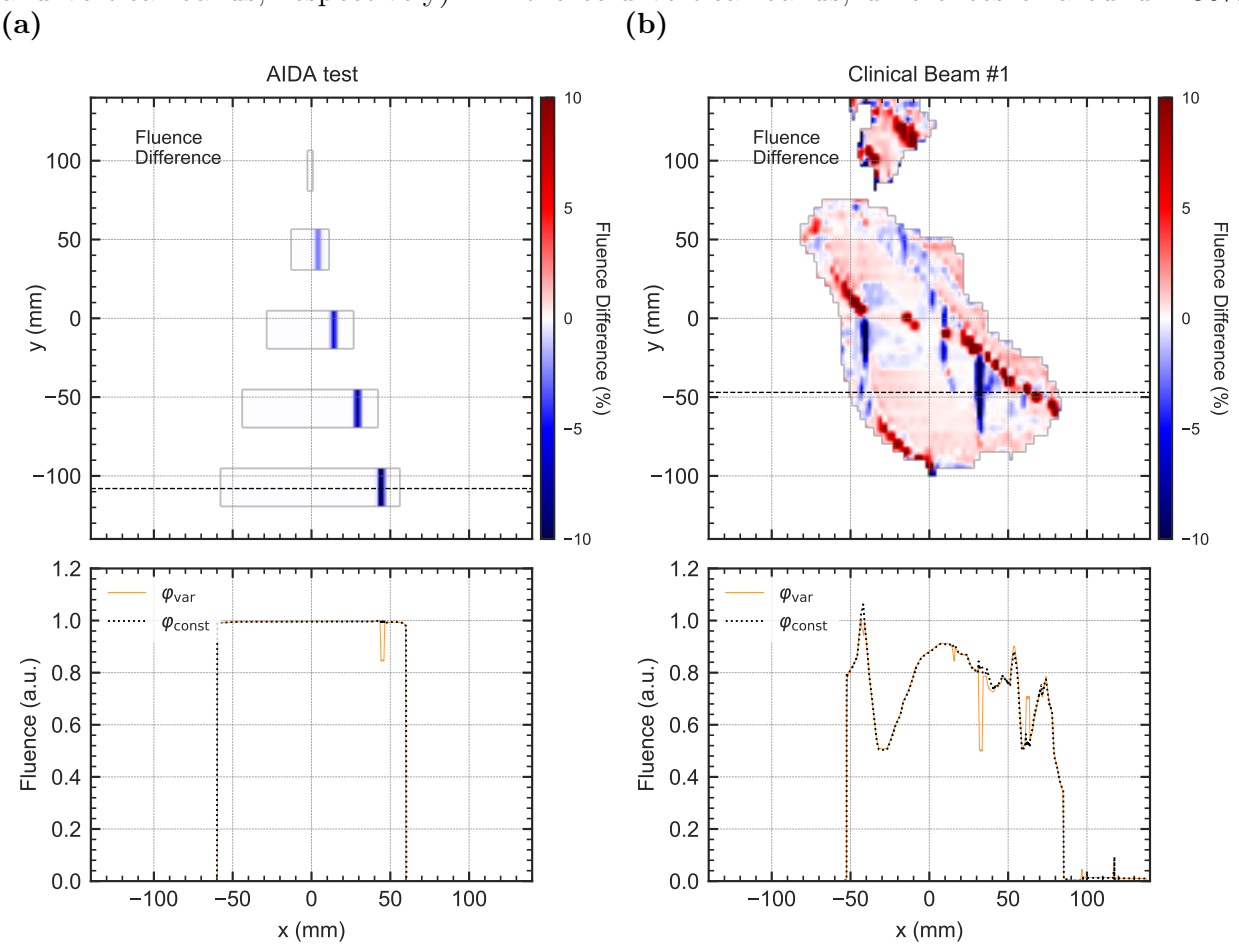


Figure 4: Comparisons between the fluence computed with a fixed ($\text{DLG}_{\text{const}}$) and a variable DLG ($\text{DLG}_{\text{var}}$). From top to bottom: fluence difference maps, and fluence profiles along the dashed profiles for (a) the AIDA test and (b) the clinical case.

were obtained, whereas in the hot ladder regions, differences reached +20%. In both cases, φ_{const} behaved similarly to Eclipse and φ_{var} behaved similarly to measurements. The fluence differences were greater than the experimental discrepancies shown in Figure 3b, which can be explained because only direct fluence without any spatial convolution was considered and this produces steeper dose variations.

III.C. Analysis of leaf speeds

The leaf speed was computed for each leaf and control point and the leaf speed map $S(x, y)$ for each leaf bank is shown in Figure 5. The speed maps reveal that the dose differences caused by the limitations in the leaf tip model coincided with the positions at the lowest leaf speeds, in this case speeds $< 0.5\,\text{cm/s}$. Correspondence with dose differences is observed for all regions of failure, i.e., in both vertical bands and ladder-like patterns. For instance, note the agreement between the positions with low leaf speeds and the positions where the gamma test failed as illustrated in the lower plots of Figure 5 where both the leaf speed and gamma map have been simultaneously represented for a single leaf pair. A high spatial correspondence exists for both leaf banks because the regions with low speeds are practically the same for banks A and B, meaning that opposing leaves were practically motionless in the same positions. The speed maps also show some low-speed areas that did not fail the gamma analysis, mostly at the periphery of the beam aperture,where steep dose gradients are present. Small differences in these regions pass the gamma analysis due to the distance-to-agreement criterion but would not be clinically relevant.

III.D. Modified plans

After determining that the reported dose differences were due to limitations in the leaf tip model implemented in the TPS (i.e., a fixed DLG), we investigated whether it was possible to minimize the impact of such limitations by increasing the trailing distance for non-collimating leaves. For this purpose, the trailing distances for the non-collimating leaves were increased by 2 mm, the modified plans were recalculated and measurements were repeated.

The gamma analysis and dose profiles for both the modified AIDA test and the modified clinical case are shown in Figure 6. In the AIDA test, the discrepancies were greatly improved

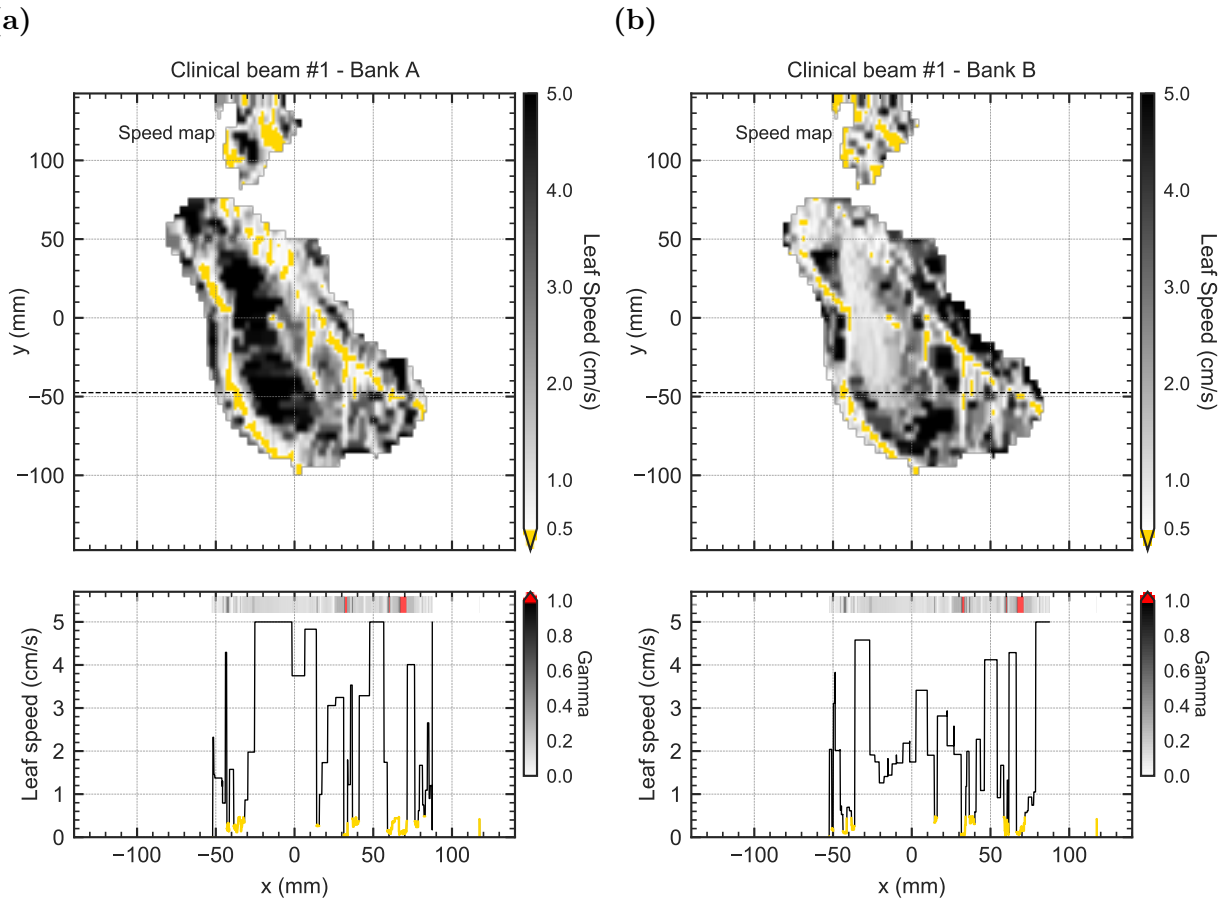


Figure 5: Speed maps S(x,y) indicating the distribution of leaf speeds for the clinical case for each MLC bank, (a) bank A, (b) bank B. The profiles of leaf speeds along the dashed lines in (a) and (b) are plotted below. A 1D representation of the gamma values obtained along the same line is attached to the upper part of the profile plot. Speeds lower than 0.5 cm/s and gamma values greater than 1 are represented in yellow and red, respectively.

and the vertical regions of gamma failure either disappeared or became much smaller. The dose differences with the modified plans were about 3%, much lower than the 10% differences obtained with the original plans.

Similarly, increasing the trailing distance in the clinical case also improved the dose agreement in the vertical bands, with a reduction from 20% measured for the original plans to 7% for the modified plans with increased trailing distances. The ladder-like regions of failure, however, remained unaffected, with cold spots and dose differences of around 20%. This strategy thus improved the global agreement but was not effective in all the regions of failure.

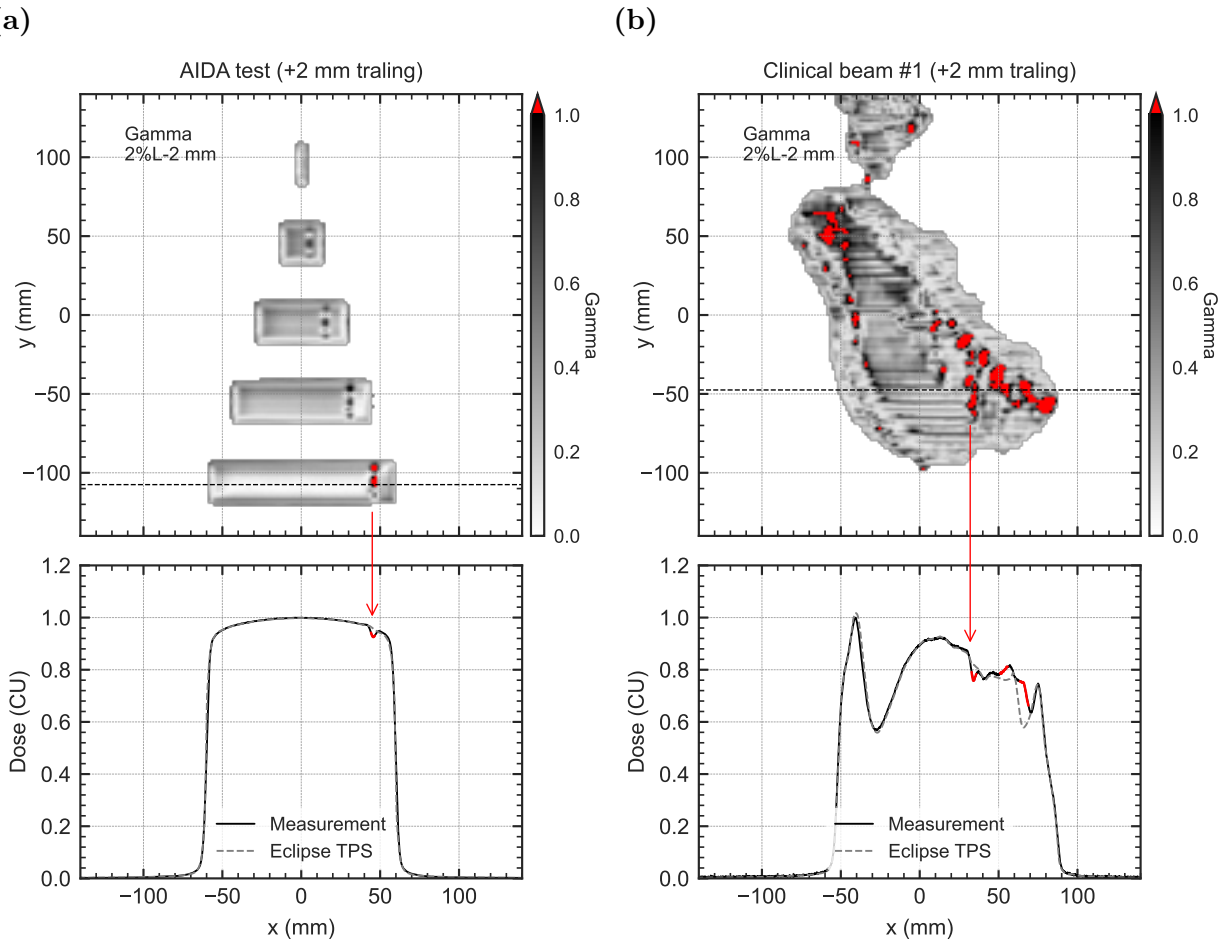


Figure 6: Comparison between Eclipse and the portal dosimetry system for the modified plans with increased trailing distances. Gamma maps (2% local, 2 mm) and dose profiles along the dotted lines are given for (a) the modified AIDA test and (b) the modified clinical case. Red regions indicate gamma values greater than 1.

The analysis performed in the previous sections was also applied to the eight additional clinical beams, which exhibited similar behavior. A complete set of plots for all cases is provided as Supporting Material, including, for completeness, the two cases presented in the main manuscript.

IV. Discussion

Dose discrepancies in test and clinical beams delivered with the Halcyon system caused by poor modeling of the leaf tip (trailing effect) were reported and analyzed. The discrepancies were found in dMLC beams with static gantry angles and were clearly detected with EPID

and film dosimetry. We did not find any evident discrepancies in VMAT treatments, probably because in VMAT the leaves move faster and perform multiple sweepings across the beam aperture. In VMAT treatments, dose discrepancies in the patient would also smear out during gantry rotation [18]. In dMLC plans, on the contrary, there is no gantry rotation while the beam is on and the leaves move slower and in only one direction, which increases the risk of dose discrepancies accumulating in the same region.

To investigate these discrepancies, we focused on a test case (AIDA) and a clinical beam from a breast treatment plan. The comparison of fluence maps computed with different leaf tip models showed differences that replicated the experimental dose differences and gamma maps, thus indicating that limitations in the leaf tip model were responsible for the dose discrepancies observed. Additionally, analysis of leaf speeds indicated that these discrepancies took place in regions where the leaves moved slowly ($<0.5\,\mathrm{cm/s}$). Low leaf speeds translated into long permanence times and, consequently, the contribution of transmission through the leaf tip accumulated in the same region and resulted in dose discrepancies. We observed two distinct spatial patterns for these dose discrepancies: straight bands perpendicular to the leaf motion direction and ladder shapes formed by several leaf tips from both layers at a similar distance from each other. In both cases, the leaf speeds from both banks were low (high permanence times) in the same positions. In the straight band patterns, the leaves from both layers were almost in the same position, trailing distances were close to zero and cold spots were found (measured doses lower than computed doses). In the ladder patterns, the leaves from each layer were several millimeters apart (trailing distance were also several millimeters) and hot spots were obtained (measured doses higher than computed doses). This was also consistent with the differences expected due to poor modeling of the trailing effect [11].

Taking into account the previous observations, we attempted to implement a mitigation strategy with the goal of minimizing dose trailing effects while preserving the leaf tip model in the TPS. We therefore externally modified the RT plans in order to increase the trailing distance of the non-collimating leaves, which could be done without affecting the beam apertures. The agreement of the modified plans greatly improved in the regions of failure with straight bands, whereas it remained unaltered in the regions of failure with ladder patterns. This was due to the fact that, in regions with ladder patterns, both MLC layers defined the beam aperture and all rge leaves acted as collimating leaves, which meant that

the trailing distance could not be increased. On the contrary, in the straight band region, only the leaves of one layer acted as collimating leaves and the leaves of the other layer could be retracted and the trailing distance increased without affecting the beam aperture.

These discrepancies in clinical plans produced failures in pre-treatment verifications, which made it necessary to replan the treatment by tentatively changing beam orientations, collimator rotation angles, and optimization parameters until the issues were finally resolved. However, this is a time-consuming process that undesirably delays the start of the treatment course. A more efficient workflow would therefore be preferable. As we have shown, anticipating these situations is possible through analysis of the plan, i.e., by comparing the fluence maps computed with different MLC models and by direct analysis of leaf speeds. This analysis can be performed either externally (exporting the plan and using dedicated tools) or within the TPS itself (using scripting tools) to flag problematic beams, take early actions and prevent interruptions in the clinical workflow.

The detailed and thorough analysis carried out in this study allowed us to identify the causes behind the discrepancies found in pretreatment verifications of dMLC plans. Our results show that the reported dose discrepancies were caused by a combination of two factors: limitations in the MLC model and some peculiar characteristics of certain leaf sequences. Regarding the leaf sequences, it was surprising to observe that some leaves were virtually motionless in regions with a homogeneous fluence, both in clinical beams and the AIDA test. Increasing the trailing distance for non-collimating leaves improved agreement in some regions, but some dose discrepancies persisted because this strategy was ineffective in ladder-like leaf patterns. Another solution is therefore necessary.

This problem can be solved in two different ways. First, the leaf sequencer in the TPS (LMC algorithm) can be optimized to avoid very low leaf speeds within the beam aperture. This would greatly help reduce the problem and would be relatively simple to achieve in regions with a homogeneous fluence, especially because there is no limitation in terms of leaf span in the Halcyon system.Second, a better MLC model could be implemented in the TPS, taking into account the dependence of the DLG with the trailing distance, to improve the accuracy of dose calculations in all situations. As shown by Hernandez et al. (2021)[11], a more detailed model of the leaf tip is needed to tackle the interplay between leaf tip transmissions from different MLC layers and its dependence on the distance between leaf positions (leaf

trailing distance). The plans showing the greatest discrepancies (and failing pre-treatment verifications) were static gantry IMRT plans using the sliding window technique, but we believe that better modeling of these effects in TPSs will also reduce uncertainties in VMAT treatments and improve overall system reliability. This is remarkably important in adaptive radiotherapy treatments, where fewer pre-treatment verifications can be carried out and a high accuracy and reliability of TPS dose calculations are essential.

One limitation of this study is that the clinical impact of the discrepancies found was not assessed. The reason is that clinical impact is strongly dependent on each particular case. Dose discrepancies of up to 10-20% were found in individual beams, but clinical plans using the sliding window technique involve multiple beams, which means that dose discrepancies in composite plans will be reduced. The overall impact on a clinical plan will depend on the exact position of these regions and on their projection within the patient's anatomy. Assessing the potential clinical impact of such discrepancies is beyond the scope of this study, but we believe that it should be carefully evaluated in each particular case.

V. Conclusions

Modeling the double-stacked MLC used in the Halcyon system is challenging due to transmission through rounded leaf-ends and the change in this transmission depending on the distance of the leaf positions in each MLC layer. Careful evaluation of clinical plans produced by Eclipse for the Halcyon is therefore recommended, especially for static-gantry IMRT treatments.

In this study we reported failures in pre-treatment verifications of sliding window plans and carried out a thorough analysis that linked these discrepancies to limitations in the MLC model and to specific characteristics of leaf sequences. In particular, poor modeling of the leaf tip and very low leaf speeds were identified as the causes of QA failures. Based on these results, strategies to anticipate, mitigate and avoid these failures are proposed. In general, leaf sequences that include trailing distances close to zero and very low leaf speeds should be avoided to reduce dose discrepancies in clinical plans. Better modeling of the leaf tip is also needed. This is particularly important in adaptive radiotherapy treatments, where the

accuracy and reliability of TPS dose calculations are of the utmost importance.

References

[1] L. Cozzi, A. Fogliata, S. Thompson, C. Franzese, D. Franceschini, F. de Rose, S. Tomatis, and M. Scorsetti, Critical appraisal of the treatment planning performance of volumetric modulated arc therapy by means of a dual layer stacked multileaf collimator for head and neck, breast, and prostate, Technology in cancer research & treatment **17**, 1533033818803882 (2018).

[2] R. De Roover, W. Crijns, K. Poels, S. Michiels, A. Nulens, B. Vanstraelen, S. Petillion, M. De Brabandere, K. Haustermans, and T. Depuydt, Validation and IMRT/VMAT delivery quality of a preconfigured fast-rotating O-ring linac system, Medical Physics **46**, 328–339 (2019).

[3] T. Netherton et al., Experience in commissioning the halcyon linac, Medical Physics **46**, 4304–4313 (2019).

[4] M. B. Sharpe, IAEA Technical Reports Series No. 430: Commissioning And Quality Assurance Of Computerized Planning Systems For Radiation Treatment Of Cancer , Medical Physics (2006).

[5] I. J. Das, C.-W. Cheng, R. J. Watts, A. Ahnesjö, J. Gibbons, X. A. Li, J. Lowenstein, R. K. Mitra, W. E. Simon, and T. C. Zhu, Accelerator beam data commissioning equipment and procedures: report of the TG-106 of the Therapy Physics Committee of the AAPM, Medical Physics **35**, 4186–4215 (2008).

[6] E. E. Klein et al., Task Group 142 report: Quality assurance of medical accelerators, Medical Physics **36**, 4197–4212 (2009).

[7] G. A. Ezzell et al., IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119, Medical Physics **36**, 5359–5373 (2009).

[8] J. B. Smilowitz, I. J. Das, V. Feygelman, B. A. Fraass, S. F. Kry, I. R. Marshall, D. N. Mihailidis, Z. Ouhib, T. Ritter, M. G. Snyder, and L. Fairobent, AAPM Medical Physics Practice Guideline 5.a.: Commissioning and QA of Treatment Planning Dose Calculations - Megavoltage Photon and Electron Beams, Journal of Applied Clinical Medical Physics (2015).

[9] K. Smith, P. Balter, J. Duhon, G. A. White Jr, D. L. Vassy Jr, R. A. Miller, C. F. Serago, and L. A. Fairobent, AAPM Medical Physics Practice Guideline 8. a.: linear accelerator performance tests, Journal of Applied Clinical Medical Physics **18**, 23–39 (2017).

[10] A. Saini, C. Tichacek, W. Johansson, G. Redler, G. Zhang, E. G. Moros, M. Quyyum, and V. Feygelman, Unlocking a closed system: dosimetric commissioning of a ring gantry linear accelerator in a multi-vendor environment, Journal of Applied Clinical Medical Physics **x**, 1–14 (2021).

[11] V. Hernandez, J. Saez, A. Angerud, R. Cayez, C. Khamphan, D. Nguyen, L. Vieillevigne, and V. Feygelman, Dosimetric leaf gap and leaf trailing effect in a double-stacked multileaf collimator, **48**, 3413–3424 (2021).

[12] T. LoSasso, C. S. Chui, and C. C. Ling, Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy, Med. Phys. **25**, 1919–1927 (1998).

[13] T. Netherton, Y. Li, S. Gao, A. Klopp, P. Balter, L. E. Court, R. Scheuermann, C. Kennedy, L. Dong, J. Metz, D. Mihailidis, C. Ling, M. Y. Lee, M. Constantin, S. Thompson, J. Kauppinen, and P. Uusitalo, Experience in commissioning the halcyon linac, **46**, 4304–4313 (2019).

[14] Varian Medical Systems, Eclipse Photon and Electron Algorithms 15.5 Reference Guide, page 332 (2017).

[15] M. M. Kim, D. Bollinger, C. Kennedy, W. Zou, R. Scheuermann, B.-K. K. Teo, J. M. Metz, L. Dong, and T. Li, Dosimetric Characterization of the Dual Layer MLC System for an O-Ring Linear Accelerator, **18**, 153303381988364 (2019).

[16] T. Y. Lim, I. Dragojević, D. Hoffman, E. Flores-Martinez, and G.-Y. Kim, Characterization of the HalcyonTMmultileaf collimator system, **20**, 106–114 (2019).

[17] S. Thompson, M. Constantin, and P. Kokkonen, Eclipse Photon Beam Models for Halcyon, Technical report.

[18] R. Miyasaka, S. Y. Cho, T. Hiraoka, K. Chiba, T. Kawachi, T. Katayose, Y. Suda, and R. Hara, Investigation of Halcyon multi-leaf collimator model in Eclipse treatment planning system: A focus on the VMAT dose calculation with the Acuros XB algorithm, Journal of Applied Clinical Medical Physics **23**, 23–e13519 (2022).

[19] J. F. Pérez Azorín, L. I. Ramos García, and J. M. Martí-Climent, A method for multichannel dosimetry with EBT3 radiochromic films, Med. Phys. **41** (2014).

[20] V. Hernandez, J. Saez, M. Pasler, D. Jurado-Bruggeman, and N. Jornet, Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy, Physics and imaging in radiation oncology **5**, 37–43 (2018).