

8-2020

## Characterization of Treatment Planning System Photon Beam Modeling Errors in IROC Houston Phantom Irradiations

Mallory Glenn

Follow this and additional works at: [https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations](https://digitalcommons.library.tmc.edu/utgsbs_dissertations)



Part of the [Physics Commons](#), and the [Therapeutics Commons](#)

---

### Recommended Citation

Glenn, Mallory, "Characterization of Treatment Planning System Photon Beam Modeling Errors in IROC Houston Phantom Irradiations" (2020). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 1016.  
[https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations/1016](https://digitalcommons.library.tmc.edu/utgsbs_dissertations/1016)

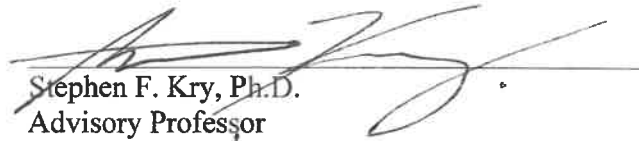
This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact [digitalcommons@library.tmc.edu](mailto:digitalcommons@library.tmc.edu).

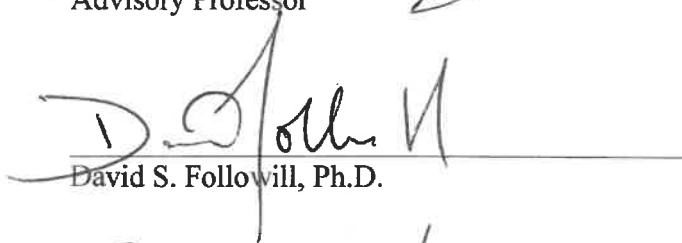
CHARACTERIZATION OF TREATMENT PLANNING SYSTEM PHOTON BEAM  
MODELING ERRORS IN IROC HOUSTON PHANTOM IRRADIATIONS

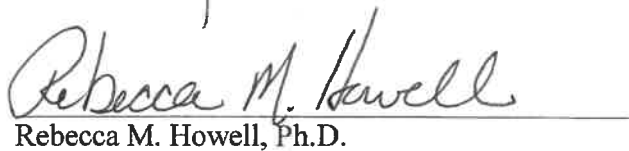
by

*Mallory Carson Glenn, B.S.*

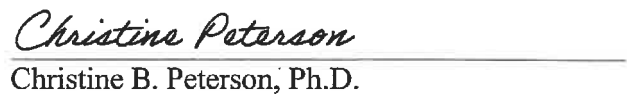
APPROVED:

  
Stephen F. Kry, Ph.D.  
Advisory Professor

  
David S. Followill, Ph.D.

  
Rebecca M. Howell, Ph.D.

  
Julianne M. Pollard-Larkin, Ph.D.

  
Christine B. Peterson, Ph.D.

APPROVED:

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

CHARACTERIZATION OF TREATMENT PLANNING SYSTEM PHOTON BEAM  
MODELING ERRORS IN IROC HOUSTON PHANTOM IRRADIATIONS

A

DISSERTATION

Presented to the Faculty of  
The University of Texas  
MD Anderson Cancer Center UTHealth  
Graduate School of Biomedical Sciences  
in Partial Fulfillment  
of the Requirements  
for the Degree of

DOCTOR OF PHILOSOPHY

by

Mallory Carson Glenn, B.S.

Houston, Texas

August, 2020

## **Dedication**

To my father.

Because I promised you I would never stop learning.

## Acknowledgements

Writing a dissertation is an arduous, mind-boggling, and herculean task. So to those of you who have prolonged my agonies with your unrelenting support and encouragement, well... this is all your fault. But all joking aside, the individuals who made this work possible deserve recognition far beyond what words can tell. Countless thanks go to:

- My advisor, Stephen Kry, for graciously supporting my ideas, professional growth, and desire for lunch meetings, *even though he never paid me*.
- The generous benefactors of the Rosalie B. Hite and the American Legion Auxiliary Fellowships in Cancer Research for inspiring and motivating my work, and *especially* for making the previous statement possible.
- My advisory committee, Drs. Followill, Howell, Pollard-Larkin, and Peterson, for kindly and eagerly scrutinizing this work every step of the way. After all, the purest gold is refined by fire.
- My poor husband Jacob, who deserves a Nobel Prize for putting up with my rants, tears, and long work days throughout this journey. Love you, kid.
- My friends and family, for misunderstanding my ideas, but nodding and smiling all the same.
- IROC Houston staff and students, who generously shared their knowledge and cake with me throughout the years.
- IROC phantom credentialing participants, for providing the wealth of data necessary to prove just how much work there is yet to be done in radiotherapy quality assurance.
- My dear friend Billie, for her limitless emotional support and always listening without judgment. Bow-wow, woof woof. Good girl. 🐾

# CHARACTERIZATION OF TREATMENT PLANNING SYSTEM PHOTON BEAM MODELING ERRORS IN IROC HOUSTON PHANTOM IRRADIATIONS

Mallory Carson Glenn, B.S.

Advisory Professor: Stephen F. Kry, Ph.D.

## **Abstract**

In radiation therapy, proper commissioning of the treatment planning system's (TPS) dose calculation algorithm is critical because any errors in this process impact all treatment plans prepared in the system. Previously, TPS errors have been identified as a major cause for poor phantom irradiation performance, which may also mean that patients are treated suboptimally. The purpose of this work was to investigate the TPS beam modeling developed by the radiotherapy community to understand where inconsistencies may arise, which variables are most susceptible to variations, and in what way changing these variables can alter dose calculations.

Using the Imaging and Radiation Oncology Core (IROC) Houston phantom credentialing framework, common observational characteristics among poor-performing phantoms were identified based on retrospective analyses of prior head and neck phantom performance. Next, treatment plan complexity, as defined by 16 different metrics, was considered and evaluated for relationships with treatment delivery accuracy for over 300 phantom irradiations. A survey was developed and deployed to the radiotherapy community to understand how institutions with similar linear accelerators (Linacs) establish their clinical beam models. From this survey information, a sensitivity analysis was completed on several head and neck phantom plans for parameters

modeling the multileaf collimator (MLC) characteristics in Eclipse and RayStation. Finally, previous phantom irradiation cases with concurrent survey results were investigated for relationships between beam modeling parameter choice and phantom performance accuracy.

The overwhelming majority of failing ( $>7\%$  error) and poor performing ( $>5\%$  error) irradiations were diagnosed as having systematic dose errors ( $>58\%$  of cases). Treatment plan complexity was completely non-predictive of phantom performance ( $p>0.01$ , Bonferroni-corrected) and all correlations between complexity and performance accuracy were weak (less than  $\pm 0.30$ ). The TPS beam modeling parameter survey generated 2818 responses from 642 institutions and revealed extensive variations in the modeling of MLC characteristics (leaf offset and transmission factor). These same parameters, namely Eclipse's dosimetric leaf gap and RayStation's MLC position offset, produced clinically significant dose changes when manipulated on 5 phantom treatment plans. Finally, the dosimetric leaf gap was associated with both poor-performing and failing phantom irradiations and correlated with TPS accuracy ( $r=0.397$ ,  $p=0.048$ ).

In conclusion, atypical beam modeling parameter values, specifically related to the representation of the MLC, are related to phantom performance and thus require careful attention in developing and performing quality assurance on the dose calculation.

## Table of Contents

Approval Page.....	i
Title Page .....	ii
Dedication .....	iii
Acknowledgements.....	iv
Abstract .....	v
List of Figures .....	xi
List of Tables .....	xv
Nomenclature .....	xix
Chapter 1: Introduction .....	1
1.1 Background and Significance.....	1
1.2 Hypothesis and Specific Aims .....	4
1.3 Dissertation Organization.....	6
Chapter 2: Sources of Error in IROC Phantom Irradiations .....	9
2.1 Introduction .....	9
2.2 Materials and Methods .....	10
2.3 Results .....	13
2.4 Discussion .....	22
2.5 Conclusions .....	25
Chapter 3: Complexity as a Contributing Factor of Phantom Performance .....	26
3.1 Introduction .....	26



3.2 Methods .....	28
3.2.1 Phantom Plans .....	28
3.2.2 Complexity Metrics .....	30
3.2.3 Data Analysis.....	32
3.3 Results .....	33
3.3.1 The relationship between TLD-based plan error and complexity metrics .....	33
3.3.2 The relationship between gamma-based plan error and complexity metrics .....	36
3.3.3 The relationship between poor performing phantom irradiations and complexity metrics .....	37
3.4 Discussion .....	38
3.5 Conclusions .....	42
Chapter 4: Survey of Photon Beam Modeling Parameters .....	44
4.1 Introduction .....	44
4.2 Acquisition and Validation Methods.....	46
4.2.1 Survey Creation and Implementation.....	46
4.2.2 Data Validation.....	47
4.2.3 Data Summarization .....	48
4.3 Data Format and Usage Notes.....	49
4.3.1 Data Format .....	49
4.3.2 Usage Notes .....	50
4.4 Discussion .....	51
4.5 Conclusions .....	57
Chapter 5: Sensitivity Analyses of Common Beam Modeling Parameters .....	58

5.1 Introduction .....	58
5.2 Methods and materials .....	60
5.2.1 TPS beam model creation and validation .....	60
5.2.2 Phantom plan development .....	61
5.2.3 Parameter manipulation (simulated beam model deviations) and evaluation .....	63
5.2.4 Parameter interplay .....	64
5.3 Results .....	65
5.3.1 Eclipse .....	65
5.3.2 RayStation .....	69
5.3.3 Dosimetric Parameters .....	74
5.4 Discussion .....	75
5.5 Conclusions .....	78
Chapter 6: Relationships between Beam Modeling and Phantom Performance .....	80
6.1 Introduction .....	80
6.2 Methods .....	82
6.2.1 Relationships between TPS beam modeling parameters and phantom irradiation performance .....	82
6.2.2 Impact of beam modeling variations on dose distributions .....	84
6.2.3 Estimating TPS dose calculation errors based on community data .....	85
6.3 Results .....	86
6.3.1 Relationship between TPS parameters and phantom performance .....	87
6.3.2 Impact of beam modeling parameters on dose distribution .....	88
6.3.3 Estimating dose error .....	93

6.4 Discussion .....	95
6.5 Conclusions .....	97
Chapter 7: Discussion & Conclusions .....	98
7.1 General Summary and Conclusions .....	98
7.1.1 Specific Aim 1: Identify common observational characteristics of poor phantom performance .....	98
7.1.2 Specific Aim 2: Determine critical beam modeling parameters for accurate dose calculation.....	100
7.1.3 Specific Aim 3: Characterize relationships between beam modeling choices and phantom performance .....	101
7.2 Evaluation of the Hypothesis .....	102
7.3 Future Research and Applications.....	103
7.4 Takeaway Messages .....	104
Appendix A: Supplement to Chapter 2 .....	106
Appendix B: Supplement to Chapter 3 .....	111
Appendix C: Supplement to Chapter 4 .....	121
Appendix D: Supplement to Chapter 5 .....	158
Appendix E: Supplement to Chapter 6 .....	178
Appendix F: Supporting Data for Mobius3D Beam Models .....	186
References.....	201
Vita.....	208

## List of Figures

Figure 2-1: IROC-H H&N phantom with block dosimetry insert. ....	11
Figure 2-2: Sample global error. ....	19
Figure 2-3: Sample systematic error. ....	20
Figure 3-1. Scatter plots of average absolute TLD error versus complexity metrics .....	34
Figure 3-2. Scatter plots of average gamma pass rate versus complexity metrics .....	37
Figure 3-3. Scatter plots of average TLD error versus complexity metrics for poor performing phantom irradiations .....	38
Figure 4-1. Distributions of survey responses .....	52
Figure 4-2. Survey responses for effective target spot size used to model standard 6 MV Varian Base class machines in Eclipse .....	53
Figure 4-3. Subset of survey responses for standard 6 MV Varian Base class and Elekta Agility Linacs that depict high variability in parameter value agreement .....	55
Figure 5-1. Axial slice of the 9-field IMRT plan developed for the IROC-H H&N phantom. ....	62
Figure 5-2. Average changes in dose calculated to target TLD structures for IROC-H head and neck phantom plans when parameters of interest are manipulated in Eclipse.....	67
Figure 5-3. Changes in average TLD dose calculated for each of the five IROC-H head and neck phantom plans following manipulation of the dosimetric leaf gap (DLG) in Eclipse.....	67
Figure 5-4. Interplay between dosimetric leaf gap (DLG) and MLC transmission (MLCT). ....	69
Figure 5-5. Average changes in dose calculated to TLD structures for IROC-H head and neck phantom plans when parameters of interest are manipulated in RayStation. ....	70
Figure 5-6. Changes in average TLD dose calculated for each of the five IROC-H head and neck phantom plans following manipulation of the MLC offset in RayStation.....	71

Figure 5-7. Select cases demonstrating the extent of interplay .....	73
Figure 5-8. Average changes in dose calculated to TLD structures for IROC-H head and neck phantom plans when dosimetric characteristics are manipulated in RayStation .....	75
Figure 6-1. Average changes in plan dose calculated to the primary and secondary PTVs following manipulation of (a) RayStation MLC leaf tip offset, (b) RayStation MLC transmission factor, and (c) Eclipse dosimetric leaf gap (DLG).....	90
Figure 6-2. Phantom cases of interest. ....	92
Figure 6-3. Comparison of estimated versus true dosimetric error for phantom irradiations performed using a Varian Base class Linac .....	94
Figure A-1. Proportions of common characteristics for failing phantoms .....	107
Figure A-2. Superior-inferior profile of a phantom irradiation exhibiting a global error.....	108
Figure A-3. Right-left profile of a phantom irradiation exhibiting a systematic error .....	108
Figure A-4. Sample setup/localization error. ....	109
Figure A-5. Sample local error. ....	110
Figure B-1. Histogram of VMAT plan complexity, as defined by MU. ....	118
Figure B-2. Histogram of IMRT plan complexity, as defined by MU. ....	119
Figure B-3. Average dose error versus complexity (MU) for all machines .....	119
Figure B-4. Average dose error versus complexity (MU) for failing phantom irradiations.....	120
Figure B-5. Average dose error versus complexity (MU) for phantom irradiations diagnosed as having a TPS error through Mobius3D recalculation .....	120
Figure C-1. Survey responses for the standard 6 MV Varian Base class for Eclipse AcurosXB129	
Figure C-2. Survey responses for the 6 MV Varian Base class for RayStation .....	130
Figure D-1. Axial slice of the IROC H&N phantom with the IMRT 5-field plan. ....	159

Figure D-2. DVH information for the IMRT 5-field plan .....	160
Figure D-3. Axial slice of the IROC H&N phantom with the IMRT 7-field plan. ....	161
Figure D-4. DVH information for the IMRT 7-field plan. ....	161
Figure D-5. Axial slice of the IROC H&N phantom with the IMRT 9-field plan. ....	162
Figure D-6. DVH information for the IMRT 9-field plan. ....	163
Figure D-7. Axial slice of the IROC H&N phantom with the VMAT 1-arc plan. ....	164
Figure D-8. DVH information for the VMAT 1-arc plan. ....	164
Figure D-9. Axial slice of the IROC H&N phantom with the VMAT 2-arc plan. ....	165
Figure D-10. DVH information for the VMAT 2-arc plan. ....	166
Figure D-11. Changes in average TLD dose calculated following manipulation of the effective target spot size in Eclipse.....	172
Figure D-12. Changes in average TLD dose calculated following manipulation of the MLC transmission factor in Eclipse .....	173
Figure D-13. Changes in average TLD dose calculated following manipulation of the source size in RayStation.....	173
Figure D-14. Changes in average TLD dose calculated following manipulation of the MLC transmission factor in RayStation .....	174
Figure D-15. Changes in average TLD dose calculated following manipulation of the leaf tip width in RayStation.....	174
Figure D-16. Changes in average TLD dose calculated following manipulation of the tongue and groove in RayStation.....	175
Figure D-17. Changes in average TLD dose calculated following manipulation of the MLC position offset in RayStation.....	175

Figure D-18. Changes in average TLD dose calculated following manipulation of the MLC position gain in RayStation.....	176
Figure D-19. Changes in average TLD dose calculated following manipulation of the MLC position curvature in RayStation.....	176
Figure D-20. Interplay between tongue and groove with MLC offset (top) and leaf tip width (bottom) in RayStation.....	177
Figure E-1. Case of interest #1: Irradiation using RayStation with high MLC Offset (92 <sup>nd</sup> percentile) .....	182
Figure E-2. Case of interest #2: Irradiation using RayStation with high MLC offset (100th percentile) .....	183
Figure E-3. Recalculated IMRT H&N phantom profiles for dosimetric modeling parameters.	184
Figure E-4. Recalculated IMRT H&N phantom profiles for RayStation’s leaf tip width. ....	185
Figure E-5. Recalculated IMRT H&N phantom profiles for Eclipse’s MLC transmission. ....	185

## List of Tables

Table 2-1. Institutional percentage pass rates for overall and individual criteria. ....	14
Table 2-2. Comparison of failures by test. ....	15
Table 2-3. Institutional gamma percentage pass rates for increasing percentage of pixels required for acceptance. ....	15
Table 2-4. Classification of institution TLD criteria failures by location in the phantom. ....	16
Table 2-5. Consensus description of causes of failure. ....	18
Table 3-1. Demographics of IMRT technique, treatment planning system (TPS), linear accelerator manufacturer, and linac-TPS combination ....	29
Table 3-2. Summary of Spearman correlations ( $r_s$ ) comparing TLD dose error and complexity metric value. ....	35
Table 4-1. Treatment planning system beam modeling parameters requested via survey. ....	47
Table 5-1. Summary of H&N phantom plans developed for testing. ....	62
Table 5-2. Beam modeling parameters investigated for sensitivity. ....	64
Table 5-3. Parameter changes implemented for the 6MV Varian Clinac-type machine simulated in Eclipse. ....	66
Table 5-4. Parameter changes implemented for the 6MV Varian Clinac-type machine simulated in RayStation. ....	70
Table 5-5. Dosimetric parameter changes implemented for the 6MV Varian Clinac-type machine simulated in RayStation. ....	74
Table 6-1. Treatment planning system beam modeling parameters requested via IROC Houston surveys and their range of dose effects, as previously determined by phantom dose recalculations using a common linac model (see Chapter 5). ....	84



Table 6-2. Summary of phantom irradiations. ....	87
Table A-1. Probabilities of phantom failure based on different reported TLD uncertainties and action criteria .....	106
Table B-1. Correlations between complexity metrics for all plans examined.....	112
Table B-2. Correlations between complexity metrics for dynamic IMRT plans.....	113
Table B-3. Correlations between complexity metrics for static IMRT plans. ....	114
Table B-4. Correlations between complexity metrics for VMAT plans.....	115
Table B-5. Correlations between complexity metrics for Eclipse plans.....	116
Table B-6. Correlations between complexity metrics for Pinnacle plans.....	117
Table B-7. Distribution summary of complexity (as defined by MU) for VMAT and IMRT plans sampled in this work. ....	118
Table C-1. Beam modeling summary statistics for 6MV beams modeled in Eclipse AAA.....	131
Table C-2. Beam modeling summary statistics for 6MV beams modeled in Eclipse AcurosXB. ....	132
Table C-3. Beam modeling summary statistics for 6MV beams modeled in RayStation. ....	133
Table C-4. Beam modeling summary statistics for 6MV beams modeled in Pinnacle. ....	135
Table C-5. Beam modeling summary statistics for 6 FFF beams modeled in Eclipse AAA. ....	137
Table C-6. Beam modeling summary statistics for 6 FFF beams modeled in Eclipse AcurosXB. ....	138
Table C-7. Beam modeling summary statistics for 6 FFF beams modeled in RayStation. ....	139
Table C-8. Beam modeling summary statistics for 6 FFF beams modeled in Pinnacle. ....	140
Table C-9. Beam modeling summary statistics for 10 MV beams modeled in Eclipse AAA....	141

Table C-10. Beam modeling summary statistics for 10 MV beams modeled in Eclipse AcurosXB. .....	142
Table C-11. Beam modeling summary statistics for 10 MV beams modeled in RayStation. ....	143
Table C-12. Beam modeling summary statistics for 10 MV beams modeled in Pinnacle. ....	144
Table C-13. Beam modeling summary statistics for 10 FFF beams modeled in Eclipse AAA.	146
Table C-14. Beam modeling summary statistics for 10 FFF beams modeled in Eclipse AcurosXB. .....	147
Table C-15. Beam modeling summary statistics for 10 FFF beams modeled in RayStation. ....	148
Table C-16. Beam modeling summary statistics for 10 FFF beams modeled in Pinnacle. ....	149
Table C-17. Beam modeling summary statistics for 15 MV beams modeled in Eclipse AAA..	150
Table C-18. Beam modeling summary statistics for 15 MV beams modeled in Eclipse AcurosXB. .....	151
Table C-19. Beam modeling summary statistics for 15 MV beams modeled in RayStation. ....	152
Table C-20. Beam modeling summary statistics for 15 MV beams modeled in Pinnacle. ....	153
Table C-21. Beam modeling summary statistics for 18 MV beams modeled in Eclipse AAA...	154
Table C-22. Beam modeling summary statistics for 18 MV beams modeled in Eclipse AcurosXB. .....	155
Table C-23. Beam modeling summary statistics for 18 MV beams modeled in RayStation. ....	156
Table C-24. Beam modeling summary statistics for 18 MV beams modeled in Pinnacle. ....	157
Table D-1. Plan information for the IMRT 5-field plan .....	159
Table D-2. Plan information for the IMRT 7-field plan .....	160
Table D-3. Plan information for the IMRT 9-field plan .....	162
Table D-4. Plan information for the VMAT 1-arc plan.....	163

Table D-5. Plan information for the VMAT 2-arc plan.....	165
Table D-6. Verification of the baseline beam model in Eclipse .....	166
Table D-7. Verification of the baseline beam model in RayStation .....	166
Table D-8. Comparison of dosimetric characteristics for the un-tuned Eclipse Varian Base 6X model.....	167
Table D-9. Dosimetric characteristics for the final tuned Eclipse Varian Base 6X model. ....	168
Table D-10. Dosimetric characteristics for the un-tuned RayStation Varian Base 6X model. ..	169
Table D-11. Dosimetric characteristics for the final tuned RayStation Varian Base 6X model.	170
Table D-12. Percentile information for PDD curve measurement errors .....	171
Table D-13. Percentile information for jaw-defined small field output factor measurement errors .....	171
Table D-14. Percentile information for off-axis measurement errors.....	172
Table E-1. Fisher’s exact test for typical/atypical parameters values versus poor- or well-performed irradiation status.....	179
Table E-2. Fisher’s exact test for typical/atypical parameters values versus passing or failing phantom status. ....	180
Table E-3. Counting of poor-performing and failing phantom cases with reported beam modeling parameters found important in dose calculation accuracy. ....	181

## Nomenclature

DTA	Distance to agreement
DVH	Dose volume histogram
H&N	Head and neck
IMRT	Intensity modulated radiation therapy
IROC Houston/ IROC-H	Imaging Radiation and Oncology Core Houston Quality Assurance Center
MCS	Modulation complexity score
MLC	Multileaf collimator
MU	Monitor unit
NCI	National Cancer Institute
OAR	Organ at risk
PDD	Percentage depth dose
PTV	Planning target volume
QA	Quality assurance
TLD	Thermoluminescent dosimeter
TPS	Treatment planning system
VMAT	Volumetric modulated arc therapy

## **Chapter 1: Introduction**

### **1.1 Background and Significance**

Of the nearly 1.7 million annual new cancer patients in the United States, radiation therapy is used to treat more than half of them.<sup>1</sup> For many of these cases, intensity modulated radiation therapy (IMRT) is considered the standard of care, especially for many hard-to-treat cases such as nasopharyngeal cancers. Both national and international scientific bodies maintain that such treatments should administer the radiation to the target within 5% of the intended dose to provide the optimal benefit to the patient.<sup>2,3</sup> If the dose is too low, the treatment is not curative; if it is too high, additional toxicities and complications can occur. However, recent clinical trial credentialing data have shown that nearly 30% of institutions fail to deliver radiation doses within that requisite 5% level, and approximately 10% of institutions are unable to perform within 7%, the current criterion for acceptability.<sup>4</sup> Thus, a substantial number of patients may be receiving clinically suboptimal treatment. Given this information, there is a critical need for quality improvement in radiation therapy.

One way in which radiation therapy quality can be determined and maintained is through the works of the Imaging and Radiation Oncology Core Houston Quality Assurance Center (IROC Houston). IROC Houston offers end-to-end testing through its anthropomorphic phantom program to ensure that institutions that treat cancer patients with radiation therapy and that participate in clinical trials can do so safely and accurately. These phantoms, resembling human anatomy with disease, are sent to institutions to be irradiated much like a real patient and returned to IROC Houston for analysis. The phantom undergoes the entire treatment delivery process: imaging the phantom to identify and delineate anatomy, generating a treatment plan following a typical protocol, setting up the phantom for treatment delivery, and finally delivering the radiation dose.

In order to track and record the dose delivered, the phantom contains thermoluminescent dosimeters (TLD) and radiochromic film, which are read upon return to IROC Houston.<sup>5-7</sup> To test the accuracy of the dose delivery, IROC Houston compares the measured dose to the planned dose reported from the institution's treatment planning system (TPS). If the planned dose and the delivered dose agree within criteria ( $\pm 7\%$ ), the institution is allowed to participate in clinical trials, as it demonstrates the ability to deliver dose accurately.

However, a large percentage of institutions do not correctly deliver dose to these phantoms, indicating the potential for problems in the treatment delivery process. Initial head and neck phantom results from 2003 demonstrated a phantom failure rate of approximately 35%.<sup>8</sup> In more recent years (as reported in 2013 and 2016) this rate has decreased and leveled off at approximately 10%.<sup>4,8</sup> This progress indicates, in part, that institutions are more able to deliver dose correctly, yet the rate of failure is still great considering the hundreds of institutions that undergo phantom credentialing each year. Furthermore, IROC Houston's current criteria for acceptability is more lax than is typically deemed acceptable in clinical practice, meaning the true rate of poorly performed irradiations could be greater than what is reported currently.

Although IROC Houston's phantom credentialing program has greatly assisted in identifying ways treatment delivery can be improved, determining the reason for failure is still a challenging process. Based on the phantom workflow, IROC Houston only knows the end result of the irradiation, i.e. whether the dose was delivered accurately. This leaves little indication for what mistakes or errors may happen over the course of the entire treatment delivery process. To compound the challenge, there exist multiple modes through which errors in the process may occur. Such problems can include incorrectly positioning the phantom, incorrect beam calibration, linear accelerator (Linac) mechanical issues, or poor dose calculations, among many other

potential causes. Beyond gross errors in localization, these modes remain generally unidentifiable in the IROC Houston phantom credentialing program. This problem is exacerbated by the inability of institutions to detect such errors using conventional IMRT quality assurance.<sup>9-13</sup> Together, these challenges underscore the need for new methods by which to identify, and ultimately rectify, errors in radiation therapy delivery.

While efforts have been made to improve error detection and correction, no one technique can identify all possible errors. Among errors that remains the most invisible is inaccurate TPS commissioning. Several studies have shown that deviations in beam modeling parameters can produce clinically-relevant differences between dose calculated and dose deposited.<sup>9,14,15</sup> This issue is serious because the commissioned TPS beam model is used to calculate the radiation dose for every treatment performed using that beam; should this model not best approximate the physical beam, patient outcomes can be negatively affected. Additionally, more recent work from IROC Houston has identified the TPS beam modeling as a major contributor to poor-performing phantom irradiations.<sup>16</sup> However, little is known about what factors in the TPS are most susceptible to causing inaccurate dose calculations and how such inaccuracies can manifest in clinical plans. This is grossly problematic, given that although IROC Houston can effectively identify suboptimal irradiations, it has no context by which to provide reasonable recommendations so that treatment delivery accuracy can be improved.

This work proposes to remedy the current unknowns by characterizing features that are common to failing phantom irradiations, especially those related to the dose calculation (and thereby related to the TPS beam modeling). This project serves as a substantial and viable step toward addressing the need for improved quality of radiation therapy by providing additional context for an identified subset of dose inaccuracies.

This work can improve the quality and safety of radiation therapy by working with IROC Houston to identify institutions that exhibit errors in dose calculation, and to uncover critical beam modeling parameters related to dose calculation errors. Such information can lead to corrections of the underlying beam model. This action would thereby improve the dose calculations for all patients treated by that beam, and thus systematically improve the accuracy and quality of therapy. Given this potential for direct impact on radiotherapy accuracy, this work is significant because characterizing TPS-related errors on a multi-institutional scale will lead to improvement in treatment accuracy and patient outcomes at radiotherapy facilities exhibiting errors in dose calculation.

The underlying data and infrastructure developed in this project will have extensive future applications for all of IROC Houston's phantom users and radiotherapy clinical trials in the United States, whose results impact cancer patients globally. This work will provide IROC Houston with additional contextual information regarding how clinical beam models should be developed in order to make knowledgeable recommendations when clinical dose calculations are identified as suboptimal. By association this work may further minimize the potential for systematic discrepancies in patient treatment among participating institutions, subsequently minimizing accrual requirements and maximizing the utility of clinical trials.<sup>17</sup> Additionally, regular auditing by IROC Houston using information found from this work will ensure that radiotherapy clinics continually provide the optimal level of care to all their patients.

## **1.2 Hypothesis and Specific Aims**

The central hypothesis of this dissertation is that inaccurate TPS commissioning (specifically beam model parameter selection) constitutes the primary phenomena related to



phantom dosimetric errors, with at least 50% of irradiations identified as having large dose disagreements reporting atypical values in one or more of the investigated parameters. To test this hypothesis, the following three specific aims were developed and completed:

*Specific Aim 1: Identify common observational characteristics of poor phantom performance.* Two studies were conducted to identify common traits among inaccurate phantom irradiations. First, a retrospective observational study was developed to determine predominating characteristics of phantom performance characteristics, as they may relate to dose calculation accuracy. Second, treatment plan complexity was assessed as a factor of phantom performance. In this work, multiple metrics were compared determine whether treatment plans with greater modulation had a greater propensity for poor irradiation performance. Our working hypotheses are that dosimetric characteristics will be a predominant factor in phantom performance, and complexity will be correlated with the degree of dosimetric error exhibited among phantom irradiations.

*Specific Aim 2: Determine critical beam modeling parameters for accurate dose calculation.* Two objectives were developed to understand the aspects of beam modeling that may be attributable to errors manifesting in phantom performance. First, a survey was developed to determine the consensus of clinical beam modeling among the greater radiotherapy community. Second, a sensitivity analysis of TPS model parameters (based upon survey results) was conducted to determine which parameters, if not modeled properly, will yield the greatest impact on the dose calculation for clinical radiotherapy treatments. A variety of treatment plans were evaluated to gain an overall understanding of how each parameter independently modifies the plan objectives (e.g. target and OAR dose). Our working hypothesis is that modeling parameters associated with

the multileaf collimators (e.g. dosimetric leaf gap) will be the most sensitive to changes and thereby present the greatest potential for changes in resultant dose calculations.

*Specific Aim 3: Characterize relationships between beam modeling choices and phantom performance.* The objective of this work was to corroborate sensitivity analysis results (from specific aim 2) with previous phantom irradiations to create a more complete narrative of TPS errors in IROC phantom performance and evaluate the hypothesis of this dissertation. For this objective, previous phantom irradiations with concurrently reported beam modeling parameters were compared based on TPS accuracy, and relationships among modeling parameters was assessed. Our working hypothesis is that phantom irradiations for which the clinical beam model had 1 or more beam modeling parameters outside the 10<sup>th</sup> and 90<sup>th</sup> percentiles will be more likely to contain clinically-significant dosimetric errors of greater than 5%.

In summary, the overarching goal of this study is to improve the identification of phantom performance errors related to the photon beam modeling inherent to all dose calculations.

### **1.3 Dissertation Organization**

This dissertation serves as a permanent record to document the work that was done to test the stated hypothesis and achieve the stated objectives. Chapters 2 through 6 are self-contained studies that each contain introduction, methods, results, discussion, and conclusions sections. Chapters 2 and 3 correspond to the work done for specific aim 1, Chapters 4 and 5 correspond to the work of specific aim 2, and Chapter 6 rounds out specific aim 3. The most pertinent and remarkable data are presented in each of these chapters, but supplementary figures and tabular data are included in the corresponding Appendices A through E.

Chapter 2 details the work of the first experiment for specific aim 1, namely, categorizing observational dosimetric characteristics of poor-performing phantom irradiations through a retrospective analysis of IROC phantom reports. In this work, phantom irradiations were categorized and assessed for commonalities in performance. Supplementary materials for Chapter 2 are contained in Appendix A.

Chapter 3 details the work of the second experiment for specific aim 2: understanding the role of treatment plan complexity in phantom irradiation performance. This work used a sample of 342 H&N phantom cases to determine the potential connection of overly-complex plans and failing phantom performance. Supplementary materials for Chapter 3 are contained in Appendix B.

Chapter 4 details the work of the first part of specific aim 2, which describes the results of a widely encompassing survey of TPS beam modeling parameters for the purpose of describing community consensus in beam modeling definition. Supplementary materials and links to the complete dataset for Chapter 3 are contained in Appendix C.

Chapter 5 details the work done for the second part of specific aim 2: a comprehensive sensitivity analysis of individual beam modeling parameters and dosimetric data inputs for the Eclipse and RayStation TPS platforms. This work describes the potential dosimetric effects of using parameter values that are not representative of the norm. Supplementary materials for Chapter 5 are contained in Appendix D.

Chapter 6 details the work of specific aim 3, which summarizes the potential connections between beam modeling variance and previous phantom performance. Supplementary materials for Chapter 6 are contained in Appendix E.

Chapter 7 summarizes the research project, including major results, overall conclusions, an evaluation of the hypothesis, and a discussion of the clinical implications of this work and future works that may be of particular interest.

## Chapter 2: Sources of Error in IROC Phantom Irradiations

This chapter is based upon the following publication:

**M. Carson**, A. Molineu, P. Taylor, D. Followill, F. Stingo, S. Kry. “Examining credentialing criteria and poor performance indications in IROC Houston phantom irradiations,” *Medical Physics* 43(12), 2016.

The permission for reuse of this material was obtained from John Wiley and Sons, Inc.

Additional materials regarding this work can be found in Appendix A.

### 2.1 Introduction

Since 2001, the Imaging and Radiation Oncology Core Houston Quality Assurance Center (IROC-H, formerly the RPC) head and neck (H&N) phantom has been used to credential institutions wishing to participate in National Cancer Institute (NCI) sponsored clinical trials utilizing intensity-modulated radiation therapy (IMRT) delivery techniques. The credentialing process ensures that the participating sites are capable of delivering complex treatment plans as intended. This limits the variability of data in the clinical trial and increases the quality of the study.<sup>18</sup> This process has also helped expose and resolve IMRT delivery errors (among other inaccuracies), which in turn helps improve an institution's treatment delivery as a whole.<sup>8,18</sup>

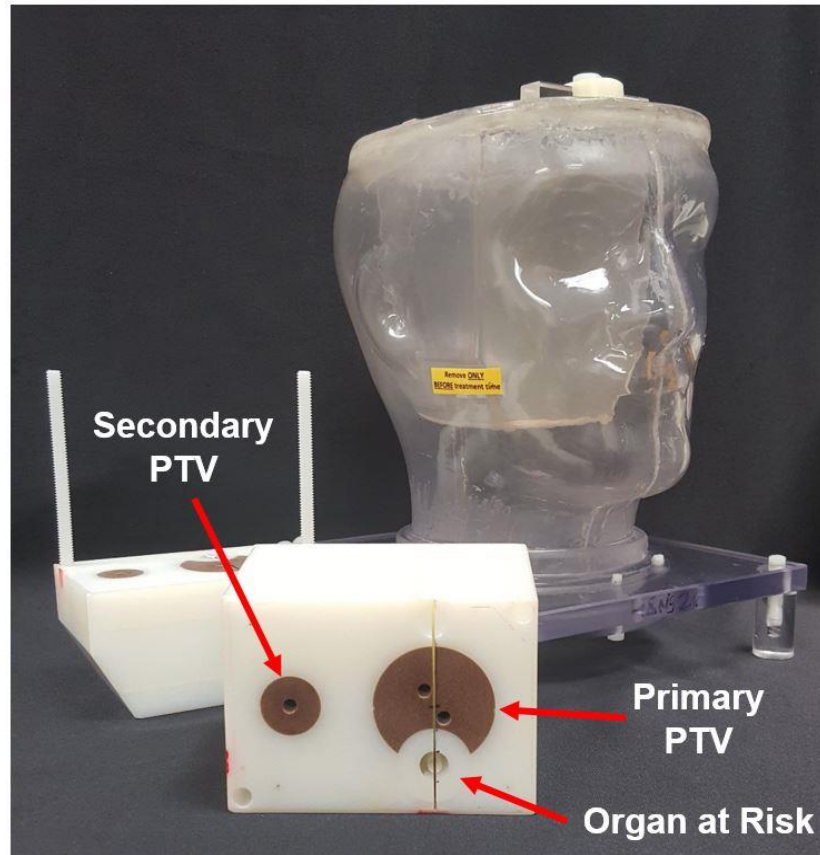
According to Molineu et al.,<sup>8</sup> since the introduction of the H&N phantom, the annual pass rate has increased from an initial low of 66% in 2001 to 88.5% in 2012 using criteria of  $\pm 7\%$  of the planned dose and  $\pm 4$  mm distance to agreement (DTA). This improvement is attributed to the

growing competency of advanced delivery techniques, modeling accuracy, and IROC-H feedback.<sup>8</sup> However, a substantial number of institutions still fail to meet the minimum criteria, thus warranting further investigation to determine the root cause. Previous works like that of McVicker et al.<sup>9</sup> have used the H&N phantom to determine the detectability of potential commissioning errors. While this work exposes some of the limitations of the phantom, it does not address the prevalence and detection of errors in multi-institutional performance. To date, no study has evaluated the predominant factors of failure in the H&N phantom credentialing results. Furthermore, as the pass rate continues to climb, it is possible that current acceptance criteria may be deemed unsuitably lax to correctly reflect the present accuracy of dose delivery.

The goal of the current study was to analyze the most recent results of IROC-H's IMRT H&N phantom irradiations with respect to thermoluminescent dosimeter (TLD) and film measured doses. Alternative criteria were applied to determine their effects on the pass rate, and failures were identified and categorized. Additionally, TLD measurement variations were analyzed to determine the probability of noise-induced failures at different criteria. This study also delineated the greatest contributions to institutional failure and examined the feasibility of revising IROC-H's acceptance criteria for credentialing.

## **2.2 Materials and Methods**

IROC-H's IMRT H&N phantom holds six double-loaded TLDs within two target structures (Figure 2-1). Radiochromic films are placed in a sagittal plane through the primary planning target volume (PTV) and in an axial plane through both PTVs and the organ at risk. The specific design and features of the H&N phantom have been described previously.<sup>5,6,8</sup>



**Figure 2-1:** IROC-H H&N phantom with block dosimetry insert.

Participating institutions are instructed to treat the phantom as they would a patient; this means the institution images the phantom, develops a treatment plan according to clinical practice, and treats the phantom. A dose of 6.6 Gy is to be administered to at least 95% of the primary PTV and 5.4 Gy is to be delivered to at least 95% of the secondary PTV. The organ at risk is to receive no more than 4.5 Gy.

IROC-H then analyzes the radiochromic films and TLDs as described by Molineu et al.<sup>6</sup> The TLD-100 capsules (Quantaflux, LLC, Oregonia, OH) are analyzed using the same technique as that used by IROC-H's mail-out dosimetry service, with a precision of 1%.<sup>19</sup> The radiochromic films (Gafchromic EBT2, Ashland, Wayne, NJ) are processed according to TG-55

recommendations<sup>20</sup> and have a localization accuracy of 1 mm and dosimetric uncertainty of 2.6%–3.6%.<sup>21</sup> For congruence, the planar dose is normalized to the TLD dose in the primary PTV. For acceptance, an irradiation must pass both TLD and film evaluations: all six TLDs must agree within a given percentage of the planned dose for each location, and both axial and sagittal films must meet the specified gamma analysis criteria.

In the current study, a total of 156 phantom irradiations between November 1, 2014 and October 31, 2015 were evaluated. Of those that failed, seven institutions repeated the phantom irradiation and six subsequently passed, and a single institution irradiated three times without passing in the timeframe of this study. All irradiations, including repeats, were included in this analysis. Uncertainties in the analyses, which were used for comparison between alternate criteria, were calculated using a binomial approximation of the variance.

Current acceptance criteria are that the measured TLD dose be within  $\pm 7\%$  of the planned absolute dose to the PTVs and that film measurements undergo gamma analysis for  $\pm 7\%$  dose and  $\pm 4$  mm DTA with  $\geq 85\%$  of pixels passing. These criteria, originally developed in conjunction with the National Clinical Trial Network, are based on the results of the first ten institutions to irradiate the H&N phantom. The current criteria were suitable such that 90% of the initial institutions could meet the criteria, and have since remained the standard for which IROC-H evaluates institution performance.<sup>8</sup> The phantoms considered in this study were re-evaluated using the standard IROC-H workflow but with the following more stringent criteria: (1) 5% TLD and 5%/4 mm, (2) 5% TLD and 5%/3 mm, (3) 4% TLD and 4%/4 mm, and (4) 3% TLD and 3%/3 mm. All gamma analyses were performed by comparing the measured film plane to the corresponding dose plane from the treatment planning system (TPS) and, unless otherwise specified, were performed with an acceptance criterion of at least 85% of pixels passing. Pass rates were also compared with



respect to percentage of passing pixels required for gamma analyses. Failure rates for varying criteria were also evaluated with respect to individual film and TLD performance by location in the phantom (e.g., film plane or primary versus secondary PTV location) in order to better elucidate the types of error manifestation.

Failure attributes were qualitatively assessed for the current acceptance criteria and for criteria 1 (5%/4 mm) by visually inspecting the dose profiles of these cases. Reports were characterized by consensus of IROC-H personnel in terms of systematic dosimetric errors (systematic dose shift of the delivered dose distribution by 3% or more on average), setup errors (systematic positional shift of the delivered dose distribution by >3 mm), global but nonsystematic errors (large-scale, nonsystematic deviation of the delivered dose distribution), and errors affecting only a local region that included irradiations in which only one TLD failed (small-scale, nonsystematic dose deviation of the delivered dose distribution). Phantom results were relatively clearly categorized into these four distinct categories.

## **2.3 Results**

Of the 156 phantom irradiations in this study, 140 (90%) met current acceptance criteria (Table 2-1). Values expressed in Table 2-1 are subdivided into the percentages of irradiations that passed both TLD and film criteria and at least one of the individual criteria. These values are shown with standard errors to facilitate comparison between the different criteria. The overall pass rate dropped 13% for criteria 1 (5%/4 mm) and continued to decline when tighter criteria were applied. A majority of participating institutions (63%) were still able to meet the 4%/4 mm acceptance criteria, for which 109 irradiations passed.

**Table 2-1.** Institutional percentage pass rates for overall and individual criteria.

Criteria	Overall pass <sup>a</sup>	TLD pass	Gamma pass <sup>b</sup>
7% TLD, 7%/4 mm	90 ± 2	93 ± 2	92 ± 2
5% TLD, 5%/4 mm	77 ± 3	80 ± 3	86 ± 3
5% TLD, 5%/3 mm	70 ± 4	80 ± 3	75 ± 3
4% TLD, 4%/4 mm	63 ± 4	67 ± 4	79 ± 3
3% TLD, 3%/3 mm	37 ± 4	49 ± 4	48 ± 4

<sup>a</sup> Overall pass rate describes the ratio of institutions that passed both gamma index and TLD criteria, thus fulfilling the requirements for acceptance.

<sup>b</sup> Gamma criterion requires  $\geq 85\%$  of pixels pass at the specified criteria.

Most (>44% for all criteria) of the failing irradiations resulted from both TLD and film (Table 2-2). Because the film dose is normalized by the TLD dose, this relationship is not surprising—if the TLD disagreed by >7%, the film would typically fail the 7%/4 mm gamma criteria when normalized to the measured TLD dose. The TLD point dosimeters identified 66%–90% of the failing cases for the changing criteria (found as the total failing cases identified by TLD, including irradiations failing both criteria). Gamma criteria did identify 55%–82% of failing cases, with the number of failures being higher when the DTA requirement was tightened to 3 mm.

**Table 2-2.** Comparison of failures by test.

Criteria	Both gamma & TLD (%)	Only TLD (%)	Only gamma (%)
7% TLD, 7%/4 mm	7 (44)	4 (25)	5 (31)
5% TLD, 5%/4 mm	17 (47)	14 (39)	5 (14)
5% TLD, 5%/3 mm	23 (49)	8 (17)	16 (34)
4% TLD, 4%/4 mm	26 (45)	26 (45)	6 (10)
3% TLD, 3%/3 mm	61 (62)	18 (18)	20 (20)

A total of 1560 gamma evaluations were calculated for this study (using both axial and sagittal films for each of the alternative criteria) and compared for different percentages of pixels required for acceptance. Increasing the percentage of pixels required yielded decreased acceptance for the tighter criteria (Table 2-3). Most of the institutions that we reviewed passed more than 90% of pixels at criteria 1 (5%/4 mm), and almost half of the institutions reviewed passed 85% pixels at the stringent criteria 4 (3%/3 mm). Pass rates dropped more than 10% when the DTA was restricted from 4 to 3 mm for the 5% criteria.

**Table 2-3.** Institutional gamma percentage pass rates for increasing percentage of pixels required for acceptance.

Criteria	≥ 85%	≥ 90%	≥ 95%
7%/4 mm	92 ± 2	90 ± 2	82 ± 3
5%/4 mm	86 ± 3	79 ± 3	58 ± 4
5%/3 mm	75 ± 3	62 ± 4	39 ± 4
4%/4 mm	79 ± 3	62 ± 4	42 ± 4
3%/3 mm	48 ± 4	28 ± 4	10 ± 2

Because each gamma analysis required both the axial and sagittal films to pass the phantom, we investigated rates of failure for the individual films. Despite the differences in film orientation, the results for the axial and sagittal films were not statistically different from one another for any of the criteria ( $p \geq 0.453$ , McNemar test), meaning that for a given poor-performing irradiation, it was equally likely for an error to be exhibited in either film. Additionally, most irradiations that did fail the film criteria failed both films. These observations are of interest because for most typical plans with no additional isocenter shift or collimator angle, the axial and sagittal films capture slightly different information: the axial film provides a detailed description of the dose distribution from a single leaf-pair, whereas the sagittal distribution provides less information on a given leaf pair, but includes information on all leaf pairs. It is interesting that neither approach shows a clear benefit over the other.

When TLD results were analyzed by location in the phantom, the majority of failures were related to the primary PTV (Table 2-4). However, the mean ratios (TLD dose to TPS-predicted dose) for the primary PTV and the secondary PTV TLDs were correlated ( $r = 0.768$ ), and this relationship was statistically significant ( $p < 0.001$ , 2-sided). This means that TLD results were consistent across the two targets, thus signaling a systematic issue.

**Table 2-4.** Classification of institution TLD criteria failures by location in the phantom.

<b>Criteria</b>	<b>Total fail (%)</b>	<b>At least PTV (%)</b>	<b>At least 2PTV (%)</b>	<b>Both PTVs (%)</b>
7%	11 (7)	9 (6)	5 (3)	3 (2)
5%	31 (20)	24 (15)	18 (12)	11 (7)
4%	52 (33)	45 (29)	35 (22)	28 (18)
3%	79 (51)	68 (44)	61 (39)	50 (32)

Note: PTV = planning target volume; 2PTV = secondary PTV

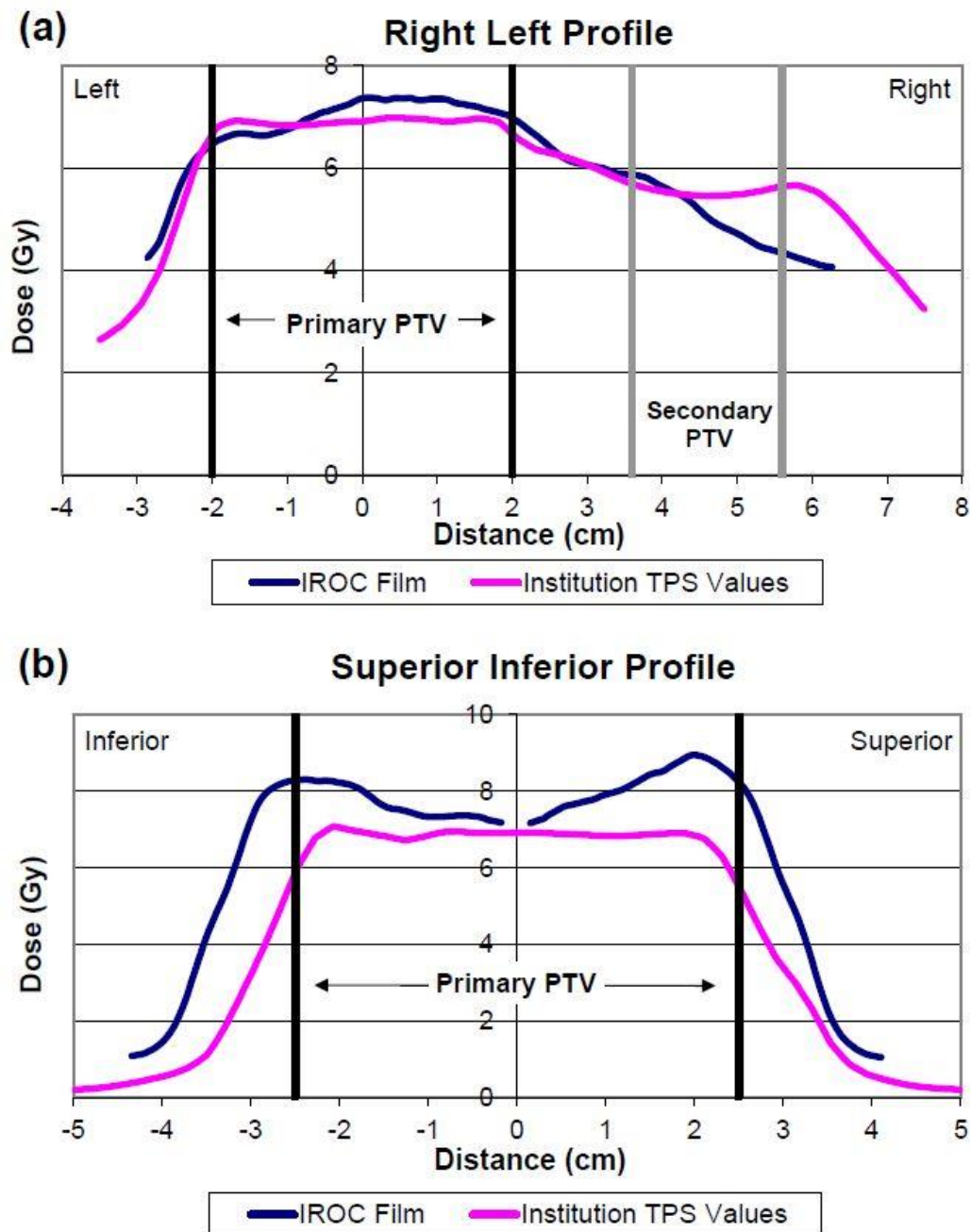
Qualitative observations of the planar dose distributions highlight predominant attributes found from failing institutions (Table 2-5). Data are presented only for 7%/4 mm and 5%/4 mm since these two criteria were found to be most practically achievable. An overwhelming majority of the failures at either criterion were determined to be due to systematic errors, with a greater extent being systematically low compared with the dose profile predicted by the TPS. That is, these profiles had the correct shape and position, but the magnitude of the dose was incorrect. Setup/positioning errors occurred in only 2 and 4 cases for the 7%/4 mm and 5%/4 mm criteria we analyzed, respectively (11%–13% of failing cases). These profile shifts were measured to be between 3 and 5 mm from the expected distribution. When further explored, it was determined that the four phantom irradiations designated as setup errors at the 5%/4 mm criteria were all cases that failed the gamma criteria only, which consisted of a total of five cases (Table 2-2). These irradiations were still within tolerance for the 5% TLD criterion. That is, almost all of the irradiations that failed only the gamma criteria but not the TLD criteria were setup/positioning errors. The third category of errors — local errors — were not observed at the 7%/4 mm criteria but did constitute a major contribution to failures at the 5%/4 mm criteria, consisting of eight failing cases (22%). Finally, we identified three global, nonsystematic failures as irradiations that either did not follow the TPS-predicted dose profile or exhibited multiple errors. Figure 2-2 and Figure 2-3 depict comparisons of film and calculated dose planes exhibiting global and systematic errors, respectively.

**Table 2-5.** Consensus description of causes of failure.

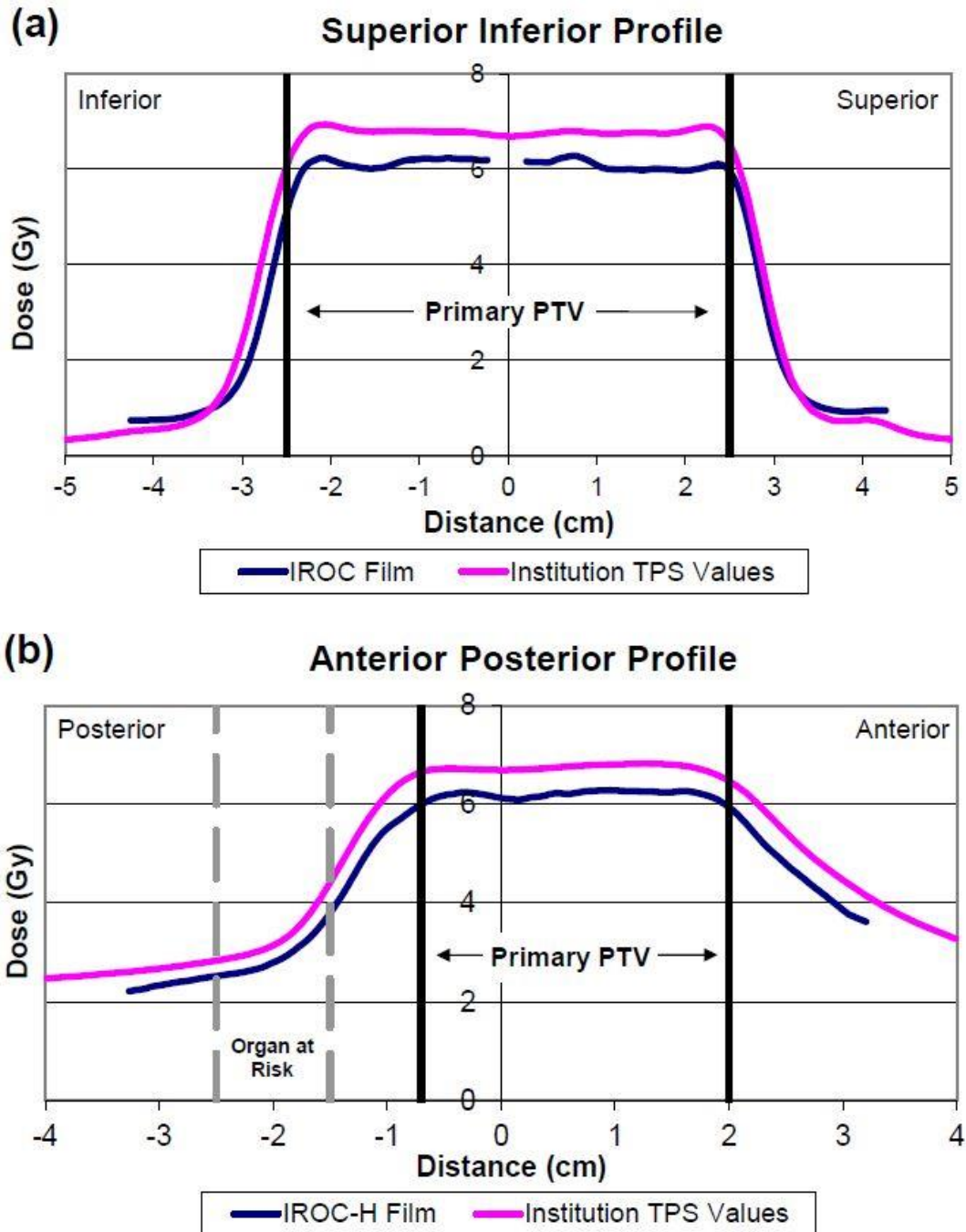
<b>Attribute</b>	<b>7% TLD, 7%/4 mm (%)</b>	<b>5% TLD, 5%/4 mm (%)</b>
Systematic	11 (69)	21 (58)
Low	9 (56)	17 (47)
High	2 (13)	4 (11)
Setup/position	2 (13)	4 (11)
Local <sup>a</sup>	0 (0)	8 (22)
Global <sup>b</sup>	3 (19)	3 (8)

<sup>a</sup> Local errors included tests in which local phenomena caused only 1 TLD to fail, typically by a small margin (~1%)

<sup>b</sup> Global errors were those that resulted from multiple errors or irregular dose distributions



**Figure 2-2:** Sample global error. These plots of (a) a right–left profile taken from an axial film and (b) a superior–inferior profile taken from a sagittal film are from an institution trial deemed to exhibit global nonsystematic errors in treatment planning and delivery. This plan has poor dose uniformity through both the primary PTV and the secondary PTV. No additional error patterns (rotation, shifting, etc.) were evidenced in this irradiation. PTV = planning target volume.



**Figure 2-3:** Sample systematic error. These plots of (a) a superior–inferior profile taken from a sagittal film and (b) an anterior–posterior profile taken from an axial film are from an institution trial deemed to exhibit systematic errors in treatment planning and delivery, as there is a systematic difference of ~8% between the measured and calculated dose distributions in the high-dose region. Such errors constitute the majority of errors observed in failing plans. PTV = planning target volume.



It is critical that tests such as the phantom irradiation test have a negligible probability of failing due to random noise in the dosimeters. This is a particular concern for those irradiations presenting local errors, which could potentially be manifestations of such an effect. Based on the known TLD variation<sup>21</sup> and assuming that phantom results are normally distributed (and that TLDs are not correlated), we calculated the statistical probability that at least one TLD result would fall outside of either the 7% or 5% criterion. The worst-case probability for a random failure of a perfectly delivered treatment (the ratio of IROC-H measured dose to TPS-planned dose being unity) was approximately 0.002% and 0.4% for the 7% and 5% criteria, respectively. Should the average measured dose be 2% different from the calculated dose, the probability of at least one random TLD failure would not exceed 9% for the 5% criterion. This rate is less than half the proportion of failures classified as local phenomena in Table 2-5 (22%), meaning these events cannot be described by dosimeter variability alone, and indicating a genuine difference between the measured and calculated doses. Furthermore, because this probability is conservative, we expect the true rate of noise-induced failures should be less than that reported here.

Nearly 70% (109/156) of the irradiations expressed a mean TLD ratio less than 1, meaning that the measured dose was less than the institution treatment plan prediction. The overall mean TLD ratio for all phantom irradiations was determined to be 0.98 (s.d. 0.03), and this mean was statistically different from unity ( $p < 0.001$ , t-test). Given that systematic biases have not been demonstrated in the TLD protocol,<sup>22</sup> these data indicate that the TPSs, in general, overestimated the dose that was actually delivered to the target. This interpretation is consistent with the observation that most phantom failures were classified as systematically low (Table 2-5).

## 2.4 Discussion

Considerable improvements have been made in the pass rate for the IMRT H&N phantom; yet, phantom irradiation failures still occur with current acceptance criteria. Therefore, it is imperative to identify and describe the causes of these failures so that corrective actions can be taken. Our study highlights that, in general, systematic differences exist between TPS calculations and the actual delivered dose, and that other errors are less prominent causes of credentialing failure. Tightening the passing criteria facilitated visualization of errors that could not be discerned at current criteria, especially local phenomena and less extreme cases of systematic discrepancies. Together these results assist in the diagnosis of potential causes of suboptimal performance in phantom credentialing.

The prominent cause of phantom failures was likely dosimetric, and inspection of failing cases likewise indicated such inaccuracy, especially underdosing (Table 2-5). Dosimetric errors may result from inaccuracies such as those in beam-modeling within the TPS software, inadequate commissioning for IMRT, or incorrect output factors (or the corresponding calibration). These faults have been identified previously in credentialing for clinical trials,<sup>8,18</sup> especially in accurate TPS modeling of the relative output of IMRT segments, where the multileaf collimator defines a small opening in a larger jaw field.<sup>23,24</sup> Consistently and systematically, the output from such an “IMRT segment” is overestimated by the TPS. This would be expected to cause an overestimation of the delivered dose from an IMRT treatment, consistent with the trends seen in this study.

One of the major challenges of the IMRT phantom program is that the phantom irradiation serves as an end-to-end quality assurance (QA) check. As such, it is limited in its ability to identify specific contributions to irradiation failure. Because the current work can only conjecture the scope

of possible errors, further work is necessary to develop methodologies to properly diagnose the source of these dosimetric errors.

For film measurements, Table 2-3 demonstrates that the current rate of agreement displayed in planar dose distributions is reasonable, that is, more than half of all irradiations passed even under the stringent  $\geq 95\%$  requirement for the 5%/4 mm criteria. This demonstrates the increasing ability of institutions to achieve accurate setup and to deliver dose to the regions of interest, which may be attributable to better image guidance. We found 85% to be sensitive enough to intercept gross errors without excessive specificity. To increase the passing pixel requirement beyond 90% may be feasible if current criteria are maintained, given that both the QA standards and possible clinical trial participation must be considered in the development of reasonable criteria.

While in many cases the TLD dose measurement can accurately determine errors in phantom irradiation, the film data are useful in the diagnosis of gross setup errors. As discussed previously, the cases in which setup error was presumed to be the mode of failure, film analyses indicated such. For these instances, film analysis was the sole failing criterion: the setup errors could not be accurately detected by TLD. This suggests that, for the criteria investigated here, when an accurate setup was achieved, the TLDs captured the extreme majority of the true irradiation conditions. This result is initially surprising (considering the seemingly large tolerances of both TLD and film criteria currently) but is actually consistent with IMRT QA results that have found point dosimeters to do a very good job of capturing the entire picture of an IMRT treatment and expose dosimetric errors that may not be discovered through planar dosimetry.<sup>11,12,25,26</sup> This result was also simulated in work by McVicker et al.,<sup>9</sup> where gamma analyses comparing plans calculated with and without commissioning errors could not reveal clinically severe effects, even

with stringent 2%/2 mm criteria. These data suggest that film dosimetry may have limited applicability in error detection for IMRT.

We expect that the introduction of new acceptance criteria may further improve the quality of results found from clinical trials. However, the decision to implement new criteria must also consider its impact on the cohort of institutions that would be permitted to participate (and consequently the number of patients who can be enrolled) in clinical trials. Although the credentialing process is meant to reduce the variability between institutions, the goal of study quality must not supersede the necessity for adequate participation.

Of the criteria considered in the current study, criteria 1 (5%/4 mm) appears to be the most plausible alternative due to its increased sensitivity relative to current standards, as well as its projected pass rate. Several visually detectable errors were identified at the  $\pm 5\%$  criterion that could not be distinguished with an action level of  $\pm 7\%$ . According to Molineu et al.<sup>8</sup> the current acceptance criteria had been criticized previously for being too lax, but adjusting the criteria was considered impractical because so many institutions still could not meet the standard. Now that more than 90% success for the H&N phantom has been achieved in the past few years, consideration of tighter acceptance criteria may be plausible. Here we demonstrate that 77% of all irradiations tested could meet criteria 1. In addition, criteria 1 would better adhere to the clinical threshold of  $\pm 5\%$  originally proposed by ICRU Report No. 24,<sup>3</sup> which is closer to the action level that is commonly accepted in clinical use.

Another concern in considering the adoption of a new acceptance standard is the possibility of noise-induced failure, especially as the criteria approach the uncertainty limit for current measurement techniques (e.g., film and TLD). As previously determined, the rates of failure due to variability for the 7% and 5% criteria are very low compared with the observed proportions.

Likewise, the response of EBT2 film is well-understood and should yield dose uncertainty of no more than 2.8%, considering all uncertainties.<sup>27</sup> Together, these findings suggest that the dosimetric differences observed were more likely a result of poor TPS calculations than of random variation, and differences should be discernable at a 5% action level. However, more work is necessary to determine the causes linked to poor performance, especially considering the degree of dosimetric errors observed.

## **2.5 Conclusions**

We have investigated the effect of tightening criteria on pass rates and have found  $\pm 5\%$  TLD and 5%/4 mm to be both theoretically and practically achievable, with 77% of institutions currently able to meet these criteria. We have also explored results from the different dosimeters and regions of interest. According to our observations, approximately half of all failures seen in credentialing tests are due to systematic underdosing, and these inaccuracies are identified primarily by TLD measurements. Other failure modes contributed to a lesser degree. While film analysis was less effective at detecting dosimetric discrepancies, it could diagnose gross setup errors. Local phenomena could only be determined at the 5% criteria, and global errors were uncommon.

Although errors are still widely present in radiotherapy credentialing and warrant attention, the extent of errors at a  $>7\%$  level is slowly declining. Tightening the criteria has the potential to increase the quality of clinical trials and to more closely reflect the precision to which institutions can currently perform. This study also highlighted that further work is warranted in identifying and resolving the continuing dosimetric errors in treatment planning.

## Chapter 3: Complexity as a Contributing Factor of Phantom Performance

This chapter is based upon the following publication:

**M. Glenn**, V. Hernandez, J. Saez, D. Followill, R. Howell, J. Pollard-Larkin, S. Zhou, S. Kry.  
“Treatment plan complexity does not predict IROC Houston anthropomorphic head and neck phantom performance,” *Physics in Medicine and Biology* 63(20), 2018.

The permission for reuse of this material was obtained from IOP Publishing.

Additional materials regarding this work can be found in Appendix B.

### 3.1 Introduction

Intensity modulated radiation therapy (IMRT), including volumetric modulated arc therapy (VMAT), is currently a standard of care technique for many disease sites. This delivery technique allows for better dose conformity than traditional 3D conformal radiation therapy while simultaneously sparing normal tissues from extraneous radiation dose. However, this technique also requires variations in multileaf collimator (MLC) motion, as well as gantry rotation speed and dose rate in some cases. Such sources of variability increase the plan ‘complexity’, a term describing the frequency and amplitude of fluctuations in IMRT dose distributions.<sup>28</sup> Thus, a simple IMRT treatment consists of large beam apertures of regular shapes, and complex IMRT beams tend to have small, narrow, or irregularly shaped apertures.

Many have previously reported that the degree of complexity (i.e. beam modulation) may be associated with greater uncertainties in radiation treatments.<sup>29–35</sup> This is a logical supposition

as high-complexity treatment plans include more challenging dose calculations and increased sensitivity to mechanical delivery performance, especially when using very small fields. The potential for delivery errors associated with highly complex plans has ushered the need to characterize and mitigate complexity in IMRT. To do so, researchers have developed several metrics as indicators of plan complexity, consisting of both fluence map-based and aperture-based metrics.<sup>29–35</sup> Fluence map-based metrics, such as the modulation index proposed by Webb, measure the variations in photon fluence between adjacent pixels in a fluence map.<sup>36</sup> Aperture-based approaches measure complexity by directly measuring the irregularity of the treatment field, as defined by the MLC, although some metrics also evaluate other plan parameters, such as leaf speed and variations of the dose rate and gantry speed.

Complexity metrics have also been suggested to be a time-efficient complement to current IMRT quality assurance (QA) methods, as they further inform the extent of beam modulation in the treatment and therefore may flag cases where modulation is higher than would normally be expected. This application is of particular interest to the Imaging and Radiation Oncology Core Houston (IROC) Quality Assurance Center. IROC seeks to confirm that institutions participating in National Cancer Institute sponsored clinical trials, including those utilizing IMRT, can calculate and deliver radiation doses consistently and accurately. For IMRT, this is done through the use of end-to-end anthropomorphic phantom irradiations whereby institutions irradiate an IROC phantom containing thermoluminescent dosimeters (TLD) and radiochromic film.<sup>6</sup> The measured dose distribution is then compared to the institution's calculated dose distribution. Yet, even with improvements in IMRT planning and delivery over time, and relatively lax dosimetric agreement criteria for the phantom (7%), a sizeable percentage of institutions continue to fail the phantom test; only 85%–90% of institutions have passed in recent years.<sup>8</sup> Of concern, dose calculation

inaccuracies have been shown to be a leading cause of treatment delivery error.<sup>4,16</sup> If complexity could be used to predict treatment accuracy, such analysis would aid in identifying the cause of phantom failures.

Therefore, the purpose of this study was to investigate the relationship between treatment plan complexity and treatment accuracy, with the aim of identifying which complexity metrics best predict planning and/or delivery errors and how much complexity contributes to dosimetric errors in IMRT delivery. To date, a comprehensive evaluation of a broad range of complexity metrics has not been done, particularly using a single, controlled patient geometry. This evaluation, as performed using IROC phantoms, has the potential to identify metrics related to the agreement between dose calculations and measurements. In addition, the information produced in this work may be used to better inform the treatment planning process or guide QA testing in order to mitigate potential errors.

## **3.2 Methods**

### *3.2.1 Phantom Plans*

A total of 343 IMRT and VMAT irradiations of IROC's head and neck (H&N) phantom (including 11 repeat irradiations) were performed by 312 different institutions between September 2011 and December 2016 as part of IROC's phantom credentialing program. The H&N phantom was chosen for evaluation because it is the most frequently irradiated phantom and is the default credentialing phantom for IMRT. The phantom contains two PTV targets and an organ at risk, and the dose was assessed with six double-loaded TLD and two sheets of film. Phantom performance was evaluated by comparing the dose calculated by the TPS with the dose actually delivered. Additional details on the phantom and analysis program are available in the literature.<sup>6</sup> Despite the



uniform geometry and planning objectives, the phantom irradiations were done with a broad cohort of delivery methods, and thereby a variety of different complexities. The demographics of these are detailed in Table 3-1. This cohort was limited to 6 MV photon treatments administered by Varian and Elekta linear accelerators, which account for the vast majority of H&N phantom irradiations. For all of these irradiations, institutions irradiated identical phantoms and were instructed to follow the same IROC protocol for phantom irradiation, thus achieving very similar dose distributions.<sup>6</sup>

**Table 3-1.** Demographics of IMRT technique, treatment planning system (TPS), linear accelerator manufacturer, and linac-TPS combination for the sample of this study.

	N	%
<b>IMRT technique</b>		
Dynamic MLC	93	27.1
Static MLC	43	12.5
VMAT	207	60.3
<b>Linear accelerator manufacturer</b>		
Elekta	39	11.4
Varian	304	88.6
<b>Treatment planning system</b>		
Eclipse	249	72.6
Pinnacle <sup>3</sup>	69	20.1
RayStation	9	2.6
Other*	16	4.7
<b>Linac-TPS combination</b>		
Elekta-Eclipse	1	0.3
Elekta-Pinnacle <sup>3</sup>	24	7.0
Elekta-RayStation	4	1.2
Varian-Eclipse	248	72.3
Varian-Pinnacle <sup>3</sup>	45	13.1
Varian-RayStation	5	1.5

\*Other TPS include XiO, iPlan, Monaco, and Oncentra

### 3.2.2 Complexity Metrics

In this study, sixteen identified measures of complexity were computed for each of the 343 phantom plans in order to provide a comprehensive view of complexity definitions, including both aperture-based and fluence map-based metrics. Here we considered both established measures of IMRT complexity from the literature, as well as several additional metrics describing variations within the MLC position, gantry position, and dose rate, thus allowing for a more well-rounded assessment of IMRT treatment delivery. For each of the metrics described herein, complexity was calculated for each beam or arc in a treatment plan, and subsequently averaged for all beams or arcs to yield the plan's average complexity. The following indices were evaluated:

- a) Total monitor units (MU) delivered. In general, a high degree of complexity is typically associated with a large number of MU; this has been used as a surrogate for treatment plan complexity previously, though correlations have not been definitive.<sup>10,31,37,38</sup>
- b) Modulation complexity score (MCS).<sup>29</sup> The MCS aims to characterize beam complexity in terms of the aperture shapes and area present throughout treatment. This metric was originally conceptualized for step-and-shoot delivery but was later adapted for sliding window and VMAT techniques.<sup>31</sup>
- c) Edge metric (EM).<sup>30</sup> This metric defines complexity as a ratio of MLC side length edge to aperture area. In this study the original recommendation for the input parameters ( $C1 = 0$  and  $C2 = 1$ ) was used. A larger EM index signifies larger positional differences between adjacent leaves.
- d) Plan irregularity (PI) and plan modulation (PM).<sup>33</sup> PI describes the non-circularity of the MLC apertures, averaged for all beams. PM indicates to what extent the beam delivery is delivered into smaller apertures; this metric is also averaged for all beams.

- e) Modulation indices ( $MI_{\text{speed}}$ ,  $MI_{\text{accel}}$ ,  $MI_{\text{total}}$ ).<sup>34</sup>  $MI_{\text{speed}}$  and  $MI_{\text{accel}}$  evaluate the extent of variation within the speed and acceleration of the MLC, respectively. In addition to these variations,  $MI_{\text{total}}$  also considers variations in gantry speed and dose rate to quantify the total delivery complexity.
- f) Leaf travel (LT).<sup>31</sup> LT indicates the average distance traveled by the MLC leaves. Because LT was originally designed for single full arc treatments, this metric is divided by the treatment's corresponding arc length to allow for comparisons with treatments with multiple arcs or partial arcs. Here the metric is denoted "LT/AL" to establish this modification.
- g) Mean dose rate variation. This metric is defined as the sum of dose rate variations from all control points, divided by arc length (to allow comparisons of plans with different numbers of control points).
- h) Mean gantry speed variation. Like mean dose rate variation, this index is the sum of variations in gantry speed, divided by arc length.
- i) Percentage of MLC gaps >10 mm. This metric describes the cumulative window width for MLC leaves. Plans with a small cumulative metric indicate the use of many small MLC gaps. Here we choose 10 mm as an appropriate threshold in order to delineate the difference between large and small leaf gaps.
- j) Mean tongue and groove index. This index indicates the fraction of MLC gaps adjacent to a consecutive leaf. This is calculated as the sum of the difference in positions from consecutive leaves for each control point, divided by the sum of gaps for the same control point.

- k) MLC interdigitation. This index characterizes the overlap between consecutive leaves from opposing banks with respect to the maximum interdigitation, taking into account the complete irradiated area outline of the MLC.
- l) Mean MLC speed variation. The mean variation in MLC speed is computed as the sum of MLC speed variations (i.e. MLC accelerations), divided by the total leaf travel.
- m) Mean Gap speed variation. The mean variation of gap sizes is computed as the sum of gap size variations, divided by the total leaf travel.

### *3.2.3 Data Analysis*

To quantify the previously described complexity indices, a MATLAB-based software called PlanAnalyzer was used to read the DICOM plans submitted by the institutions undergoing phantom credentialing.<sup>39</sup> These measures of plan complexity were compared against the dosimetric error found for each delivered plan. The average TLD error was defined as the average magnitude percentage difference between the TPS-calculated doses and the corresponding measured doses for the TLD in the H&N phantom (six TLD per phantom). Because point dosimetry may not fully characterize the irradiation conditions, plan error was also measured by the percentage pixels passing from radiochromic film gamma analysis, following IROC's protocol for analysis with the criteria of 7% dose agreement and 4 mm distance to agreement.<sup>8</sup>

Correlations between complexity metrics and phantom plan error were determined using Spearman's rank-order correlation coefficients (with Bonferroni corrections applied for multiple comparisons). For the purposes of this work, the strength of the association in absolute value was defined as follows: 0–0.19 was regarded as 'no correlation', 0.20–0.39 as 'weak', 0.40–0.59 as 'moderate', and 0.60–1 as 'strong'. Correlations were evaluated for the entire sample, as well as according to TPS (Pinnacle and Eclipse), machine type, and delivery technique, as delineated in

Table 3-1. Similarly, poor phantom results, those with at least one TLD measuring >5% error, were segregated, and the same analyses were applied to visualize whether such clinically underperforming plans had distinguishing features.

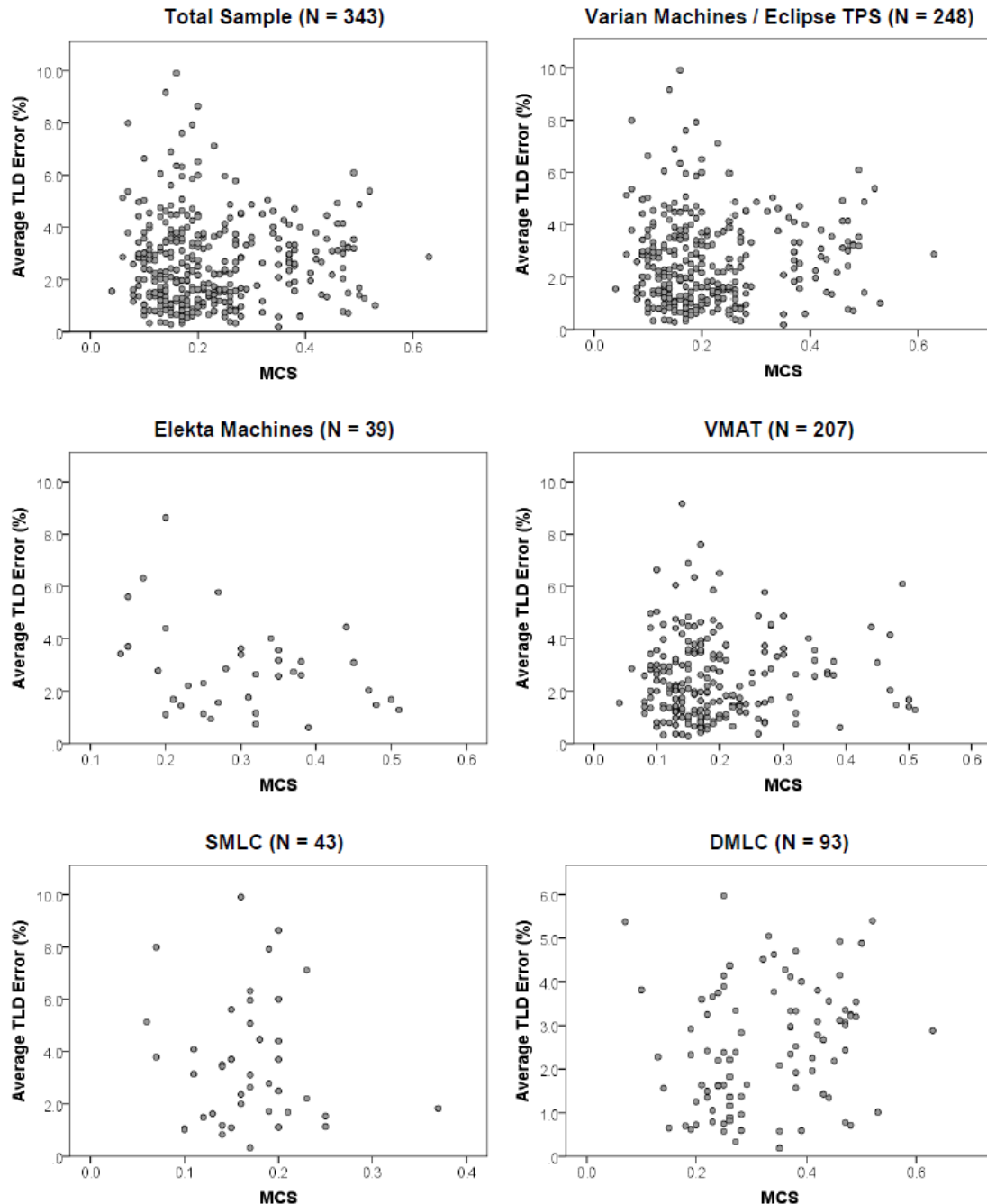
### 3.3 Results

#### 3.3.1 *The relationship between TLD-based plan error and complexity metrics*

Despite the uniformity of the phantom and dose objectives, the plans in this study had a comprehensive assortment of treatment complexities; for example, MU used in delivery ranged from 458 to 3358 with a mean of 1883. Figure 3-1 shows the distributions of the MCS and corresponding plan error for the total sample and multiple subsamples examined in this work. Visually, these distributions represent poor ability of complexity metrics to be utilized as a means of distinguishing irradiations prone to error, at least under the circumstances examined herein. Other complexity metrics appeared similarly and yielded indistinguishable relationships.

Relationships between complexity metrics and plan error are shown in Table 3-2. The only index to achieve significance was MLC interdigitation ( $p < 0.003$ ); no other complexity metric significantly predicted dosimetric inaccuracies in the calculation or delivery of the radiation dose. However, even this single significant relationship, found for the total sample and also only Varian machines, had a very low correlation strength ( $|r_s| < 0.18$ ), which indicated no clinically meaningful relationship by our criteria. The highest correlation was found between LT/AL and Elekta machines supported by Pinnacle ( $r_s = -0.395$ ,  $p = 0.116$ ), but this correlation coefficient was still classified as ‘weak’ and was not found to be a significant relationship, partly due to the small subsample size. From this data, it is evident that complexity metrics are not related to the

TLD error observed in IROC's H&N phantom practice, regardless of delivery technique, TPS, or machine manufacturer.



**Figure 3-1.** Scatter plots of average absolute TLD error versus complexity metrics for the whole sample and several subsamples in this work. The distributions showed no correlation between the measured TLD error and the assessed complexity metric, here shown as the MCS (smaller values of MCS correspond to more complex plans). SMLC = static MLC, DMLC = dynamic MLC.

**Table 3-2.** Summary of Spearman correlations ( $r_s$ ) comparing TLD dose error and complexity metric value for the subsamples described in this study (i.e. machine manufacturer, TPS, or delivery method).

		MU	MCS	EM	PI	PM	MI speed	MI accel.	MI total	LT/AL	Mean DR Var.	Mean GS Var.	Gap >10mm	Mean TG	MLC Inter- digitation	Mean MLC Speed Var.	Mean Gap Speed Var.
<b>All Machines</b>	$r_s$	-0.025	0.031	-0.119	-0.082	-0.018	-0.024	-0.028	-0.007	-0.006	0.025	-0.006	-0.018	-0.080	-0.176	-0.132	-0.134
N = 343	$p$	0.645	0.570	0.027	0.130	0.742	0.655	0.599	0.899	0.931	0.650	0.917	0.740	0.141	0.001*	0.015	0.013
<b>Varian Machines</b>	$r_s$	-0.022	0.040	-0.107	-0.065	-0.011	-0.010	-0.011	0.008	0.002	0.039	0.008	0.015	-0.065	-0.174	-0.120	-0.124
N = 304	$p$	0.699	0.484	0.063	0.256	0.853	0.864	0.853	0.896	0.976	0.500	0.890	0.798	0.256	0.002*	0.036	0.030
<b>Elekta Machines</b>	$r_s$	-0.016	-0.303	-0.150	-0.071	0.016	-0.241	-0.223	-0.283	-0.095	-0.237	-0.227	-0.044	-0.075	0.002	-0.130	-0.174
N = 39	$p$	0.924	0.061	0.363	0.668	0.922	0.140	0.173	0.081	0.636	0.147	0.165	0.791	0.649	0.991	0.429	0.289
<b>Pinnacle TPS</b>	$r_s$	0.208	-0.103	0.233	0.188	0.158	0.018	0.022	0.020	-0.173	0.002	0.049	0.200	0.171	0.089	0.140	0.159
N = 69	$p$	0.087	0.397	0.054	0.122	0.196	0.883	0.855	0.872	0.274	0.986	0.688	0.100	0.159	0.466	0.253	0.193
<b>Eclipse TPS</b>	$r_s$	0.025	0.017	-0.093	-0.058	0.055	-0.080	-0.066	-0.070	-0.059	-0.063	-0.075	0.029	-0.022	-0.110	-0.131	-0.104
N = 249	$p$	0.697	0.785	0.142	0.364	0.390	0.208	0.298	0.269	0.467	0.322	0.236	0.646	0.730	0.082	0.039	0.102
<b>Varian + Eclipse</b>	$r_s$	0.024	0.019	-0.094	-0.058	0.056	-0.078	-0.065	-0.070	-0.054	-0.065	-0.076	0.028	-0.022	-0.114	-0.131	-0.106
N = 248	$p$	0.708	0.768	0.141	0.362	0.382	0.221	0.306	0.274	0.509	0.309	0.236	0.659	0.725	0.073	0.039	0.096
<b>Varian + Pinnacle</b>	$r_s$	0.143	0.116	0.235	0.200	0.080	0.149	0.124	0.127	0.327	0.103	0.192	0.144	0.124	-0.042	0.231	0.242
N = 45	$p$	0.349	0.447	0.121	0.187	0.603	0.327	0.417	0.405	0.110	0.499	0.207	0.344	0.417	0.782	0.127	0.109
<b>Elekta + Pinnacle</b>	$r_s$	0.296	-0.446	0.131	0.120	0.225	-0.295	-0.282	-0.314	-0.395	-0.311	-0.187	0.244	0.195	0.203	-0.144	-0.111
N = 24	$p$	0.160	0.029	0.541	0.576	0.291	0.161	0.182	0.135	0.116	0.139	0.381	0.250	0.361	0.341	0.502	0.607
<b>VMAT</b>	$r_s$	0.005	0.021	-0.077	-0.008	0.040	0.100	0.089	0.148	-0.006	0.122	0.119	0.176	-0.008	-0.140	-0.159	-0.177
N = 207	$p$	0.945	0.765	0.272	0.903	0.571	0.151	0.200	0.033	0.931	0.079	0.086	0.011	0.904	0.044	0.022	0.011
<b>DMLC</b>	$r_s$	-0.218	0.240	-0.188	-0.079	-0.088							-0.052	-0.179	-0.220	0.065	0.160
N = 93	$p$	0.036	0.021	0.071	0.450	0.402							0.623	0.086	0.034	0.534	0.127
<b>SMLC</b>	$r_s$	0.212	0.057	-0.003	0.119	0.097							0.067	0.111	-0.031		
N = 43	$p$	0.172	0.719	0.987	0.446	0.535							0.670	0.480	0.846		

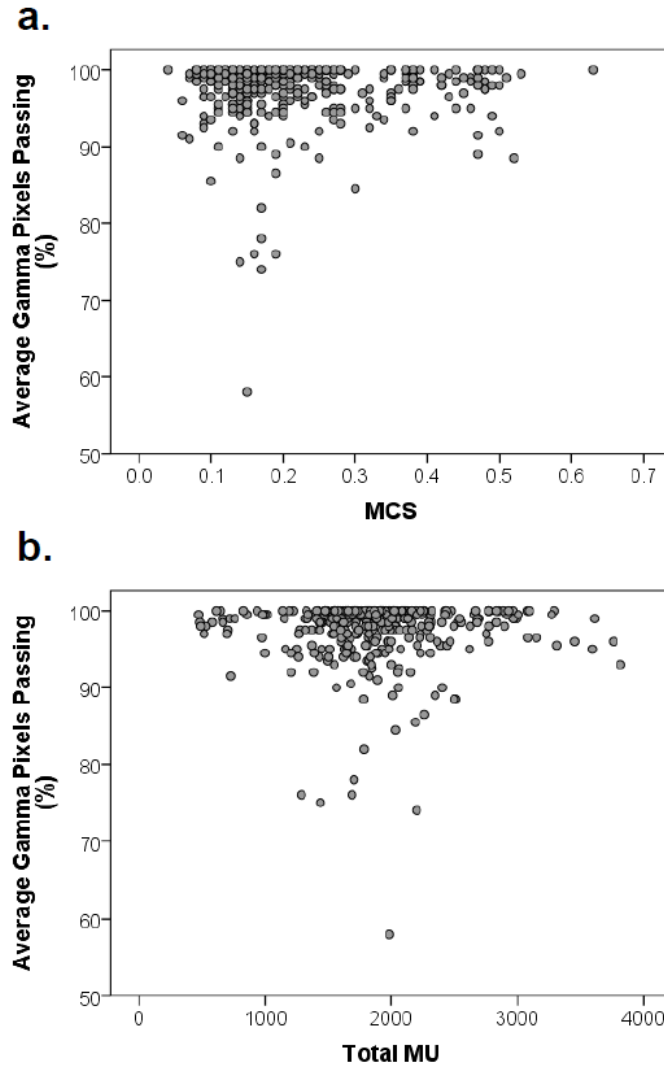
\* Correlation is significant at the 0.3% level (required for Bonferroni correction)

Note: DMLC = dynamic MLC, SMLC = static MLC, MU = monitor units, MCS = Modulation Complexity Score, EM = Edge Metric, PI = Plan Irregularity, PM = Plan Modulation, MI = Modulation Index, LT/AL = leaf travel per arc length, DR = dose rate, GS = gantry speed, TG = tongue and groove.

### *3.3.2 The relationship between gamma-based plan error and complexity metrics*

Figure 2 shows the distributions of treatment complexity and corresponding average film gamma percentage pixels passing for two prominent metrics, MCS and MU. Much like the results of Section 3.3.1, no significant relationships were evident, regardless of how the sample was broken down ( $r_s < 0.206$ ,  $p > 0.05$ ). Upon further inspection, this result is expected because the average absolute TLD error is correlated with the average gamma pass rate ( $r_s = -0.464$ ,  $p < 0.001$ ), meaning similar information is provided by both methods of plan error measurement.



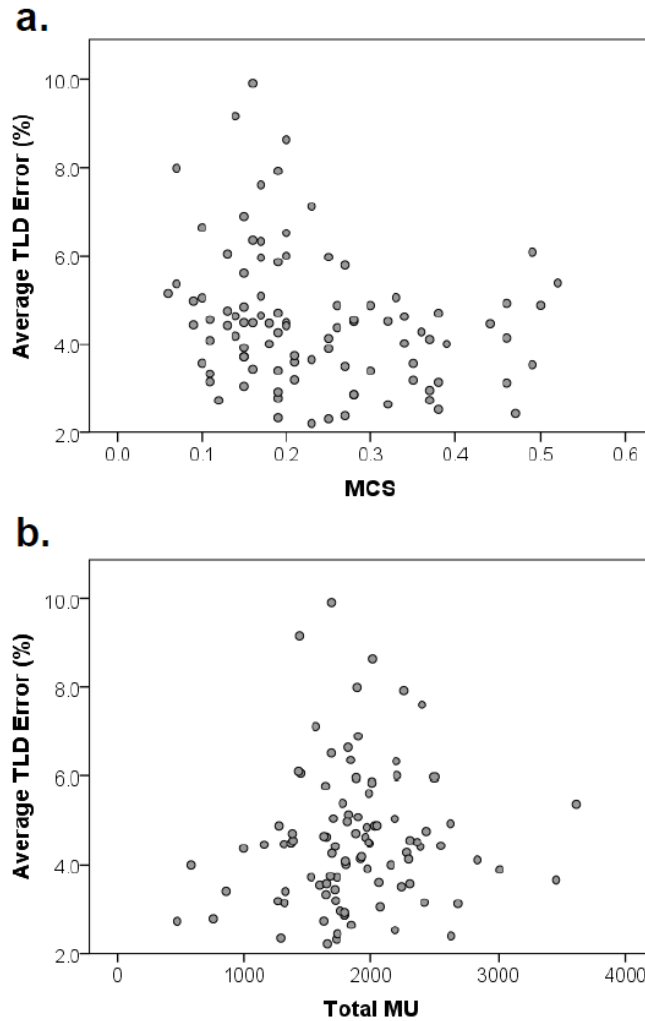


**Figure 3-2.** Scatter plots of average gamma pass rate versus complexity metrics for all irradiations examined in this work: (a) MCS, (b) total MU.

### 3.3.3 The relationship between poor performing phantom irradiations and complexity metrics

Of the 343 phantom irradiations initially analyzed, 96 cases were identified as ‘poor performers’ based on a threshold of 5% TLD dose error for any given TLD within the phantom. Figure 3-3 depicts two distributions of treatment complexity and corresponding plan error. Like

previous cases, no relationships were found to be significant, meaning that trends could not be distinguished in even the most concerning of irradiations.



**Figure 3-3.** Scatter plots of average TLD error versus complexity metrics for poor performing phantom irradiations: (a) MCS, (b) total MU.

### 3.4 Discussion

In this study, we examined several known measures of complexity, as well as additional plan metrics. These metrics generally have clear physical meanings that describe the beam aperture

or fluence. The rationale for examining sixteen metrics was because others have suggested that a single measure may not be able to reveal all the details of the complexity in IMRT plans, nor may an individual metric be suitable for all TPSs.<sup>33,39</sup> By evaluating several metrics, we quantify different aspects of complexity and generate a much clearer, more comprehensive picture of a treatment plan and its potential challenges. Here the use of complexity metrics allowed for the potential identification of specific relationships that can influence plan performance in IROC's uniform phantom program.

Unfortunately, the results of this work show that there are no observable correlations between complexity metrics and the observed plan error in recent IROC H&N phantom performance. These results are interesting because they corroborate well with preliminary work that evaluated the utility of the MCS with a small cohort of H&N phantom plans on a single machine.<sup>40</sup> It was expected that certain metrics would not produce strong correlations based on typical clinical practices: for example, variations of the dose rate are generally well-controlled, and many IMRT plans do not have any dose rate variation. Other studies have also described how some metrics, such as the MCS, do not have a large effect on IMRT QA performance, which may then translate to a lack of relationship in our work.<sup>29,41</sup> Additionally, certain metrics provide similar information, as was determined by Hernandez et al. in their comparisons of MCS, PI, and EM, meaning these indices should consequently produce similar results.<sup>39</sup> However, what is somewhat surprising from this work is that none of the sixteen complexity metrics even remotely produced viable relationships with the IROC H&N phantom results under the range of conditions evaluated. Our analyses show that complexity metrics were poor tools for predicting phantom performance and may have limited utility in determining the accuracy of treatment delivery.

Previous works examining IMRT complexity have shown mixed success in determining relationships with plan performance. Götstedt et al. observed strong correlations between several metrics and gamma pass rates for both EBT3 film and portal imaging using a variety of MLC aperture shapes.<sup>35</sup> Likewise, both Masi et al. and Agnew et al. determined that the MCS correlates with gamma analyses for patient-specific QA.<sup>31,37</sup> Building on these ideas, Crowe et al. suggested that there exist threshold complexity values that can be defined to identify plans that are likely to fail QA.<sup>32</sup> However, others, such as Du et al. and McNiven et al., conceded that their proposed complexity indices did not yield correlations with plan quality metrics (including IMRT QA) but could still have utility in limiting the uncertainty in IMRT performance.<sup>29,33</sup> Of particular interest, the work of McGarry et al., which observed QA phantom irradiations from multiple institutions, discovered weak but significant relationships between MU and plan quality, as well as MCS and plan quality, for all linear accelerators or Varian accelerators considered in their work.<sup>42</sup> The work presented here clearly shows no indication of significant relationships between complexity and plan performance based on a relatively large sample of irradiations performed using multiple TPS, linear accelerator models, and delivery techniques.

In previous works relating complexity and plan quality, patient-specific IMRT QA was typically used as a measure for plan performance accuracy. More recently, IMRT QA has come under scrutiny because of its inability to discriminate unacceptable plans.<sup>10,11,43,44</sup> Between different institutions, IMRT QA is completed using a plethora of devices, delivery techniques, and criteria for acceptability, which also limit the reproducibility and applicability of results derived from IMRT QA analysis. This work differs from previous studies of IMRT plan complexity in that it is among the first to examine complexity on a single patient geometry, the H&N phantom, using in situ dose measurement to characterize plan error for a multitude of institutions. Although both

IMRT QA and IROC phantoms are designed to verify the accurate delivery of IMRT, this distinction is important because the phantom provides a direct comparison between the dose that was planned and that which was delivered, whereas IMRT QA measurements serve as a proxy for treatment accuracy. The H&N phantom is advantageous because all the plans observed had similar treatment objectives, thus eliminating the variability found between patient plans. This phantom also has a conceptual advantage over IMRT QA because its analysis was designed with consistent dose delivery in mind: all its irradiations are processed, analyzed, and evaluated in a consistent manner, and the uncertainties of this process are documented and well controlled.<sup>6</sup> Such standardized treatment limits variability and allows for better understanding of overall performance trends in the radiotherapy community.

However, this approach is not without its own limitations. While the phantom test can control for many factors that other studies could not, such as patient geometry, this process also introduces other forms of variance, which can arise from the multitude of beam models used to calculate the dose distributions. Additionally, because the phantom is an end-to-end assessment of the treatment delivery process, it is possible that some of the plan errors observed here do not have a causal link with treatment delivery, but rather other external factors, such as phantom setup. Fortunately, based on IROC experience, incorrect setup does not contribute near as much to phantom errors as do systematic dosimetric inaccuracies.<sup>4</sup> There may also exist cases for which our methods are not sufficiently sensitive to characterize dose errors caused by excessively complexity plans, yet these would not be of clinical concern, as the measurement uncertainty for each double-loaded TLD is approximately 1.6%.<sup>22</sup> Lastly, another factor that was not examined, but would pose a valid concern for patient treatment, is the potential effect of motion. Longer

treatment times, as is common with high complexity treatments, may increase the sensitivity of dose accuracy to patient/target motion, but this could not be tested with a static phantom.

Though limiting the complexity of a plan may be good practice to limit some planning and delivery uncertainties, other factors may contribute to the degradation of plan accuracy. First and foremost of possibilities is the TPS calculation, which includes the beam modeling and inputs for beam characterization. The use of MLC-shaped beam segments, as is standard in IMRT, requires accurate modeling of several factors, including the leaf end, leaf transmission, and inter-leaf leakage.<sup>45</sup> If modeled improperly, the dose distributions delivered through MLC-defined apertures will have introduced error; systematic dosimetric errors have been documented for small fields.<sup>23</sup> Second, errors could be related to phantom or QA device positioning, which is user-dependent. Third, errors could be caused by inaccurate machine delivery characteristics, especially concerning the MLC positioning and dose rate accuracies. Because complexity measurement cannot encompass all potential failure modes, it is essential that these and other treatment delivery factors also be considered when assessing the potential for poor plan performance.

### **3.5 Conclusions**

This study evaluated IMRT treatment plan complexity metrics with the purpose of identifying those which best predicted irradiation errors. Surprisingly, existing complexity metrics were universally not predictive of dosimetric errors in the IROC H&N phantom irradiations. That is, all metrics evaluated in this study failed to show a statistically significant relationship between phantom performance and the degree of complexity of the treatment plan, regardless of delivery technique, machine model, or TPS. This is interesting, because unlike previous experiments evaluating complexity metrics, the irradiated geometry is constant and without the heterogeneities

or uncertainties found in real patient cases. These findings indicate that variations in beam complexity could not explain the disparities in phantom plan performance and that other factors affecting treatment delivery, such as beam modeling inaccuracies, dictate the accuracy of phantom treatment plans.

## Chapter 4: Survey of Photon Beam Modeling Parameters

This chapter is based upon the following publication:

**M. Glenn**, C. Peterson, D. Followill, R. Howell, J. Pollard-Larkin, S. Kry. “Reference dataset of users’ photon beam modeling parameters for the Eclipse, Pinnacle, and RayStation treatment planning systems,” *Medical Physics* 47(1), 2020.

The permission for reuse of this material was obtained from John Wiley and Sons, Inc.

Supplementary documentation and other materials from this work can be found in Appendix C.

### 4.1 Introduction

Constructing an accurate and robust linear accelerator (Linac) beam model is fundamental to providing high-quality radiation therapy. To do so, medical physicists must manage and define several dosimetric and non-dosimetric input parameters to create a model in the treatment planning system (TPS) that optimally agrees with the physical Linac output. It is expected that this model will then be suitable for a wide variety of clinical scenarios. The process of beam model creation often consists of several iterations in order to achieve the most robust solution, and the amount of adjustment or model tuning available to the user varies among TPS vendors.

The challenge in beam model creation, however, is that rapidly advancing technologies such as intensity-modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) can test the limits of these TPS algorithms, thus requiring that extra care and attention be given during the commissioning process. Several studies have determined that clinically



significant errors (>5%) can occur when certain factors are measured or employed improperly, thus underscoring the importance of beam model accuracy.<sup>9,15,46–48</sup> As such, both the approach and user knowledge needed to achieve good dosimetric commissioning of IMRT and VMAT are of paramount importance. Accordingly, it is the duty of the qualified medical physicist to understand both the limitations of the dose calculation algorithm and measurements used to validate the model to ensure its accuracy in clinical use.<sup>49</sup>

To provide assistance in beam model creation, professional organizations such as the International Atomic Energy Agency and the American Association of Physicists in Medicine have published recommendations for both the commissioning and quality assurance of treatment planning systems for modern applications, including IMRT.<sup>45,49–52</sup> These references provide guidance and several validation tests for the dose calculation algorithms used by the TPS. However, such guidance does not extend to specific methods by which commercially available TPS software and individual model parameters are to be evaluated, especially because each TPS manufacturer has different standards and specifications for clinical use.

Despite these additional resources for beam model development and testing, studies have determined that treatment errors related to the TPS calculation still exist. The Imaging and Radiation Oncology Core Houston Quality Assurance Center's (IROC Houston) recent works underscore the continued challenges of achieving accurate dosimetric commissioning for IMRT systems. While the percentage of institutions that pass the IROC Houston head and neck phantom irradiation has improved over time, a substantial number of institutions still fail to meet the relatively loose minimum criteria required for clinical trial credentialing.<sup>4,8</sup> Results published by Kerns, et al.<sup>16</sup> and Carson, et al.<sup>4</sup> strongly demonstrate the dominating dosimetric characteristics of phantom failures, which can originate from poor TPS dose calculations.

One way to help ensure accurate beam model commissioning is to understand how the radiotherapy community at large defines their beam models. Because of IROC Houston’s unique relationship with the radiotherapy community and the relative complexity associated with beam model creation, the goal of this work was to create a reference dataset for comparison of C-arm Linac models that may aid physicists in beam model creation and validation. Although several beam modeling parameters have been tested previously on multiple Linacs from different institutions and demonstrated their relative importance in beam modeling,<sup>14,15,47</sup> no large scale data source describing individual parameters is yet available. Here we provide distribution characteristics of several key input beam modeling parameters for several Linacs, TPS, and beam energies, so that physicists may evaluate their institution’s beam models in the context of the distribution of similar Linacs.

## **4.2 Acquisition and Validation Methods**

### *4.2.1 Survey Creation and Implementation*

In order to acquire data describing how the community defines its Linacs in their TPS environments, a survey was designed for Eclipse, Pinnacle, and RayStation users. These three TPS were chosen based upon the demographics of IROC Houston service users and their frequency of use today. The survey encompassed multiple beam modeling parameters, many of which have been found to be of interest from previous studies.<sup>14,47,53,54</sup> The requested beam modeling parameters are listed for each of the three TPSs examined in Table 4-1. Notably, these parameters model the behavior of the multileaf collimator (MLC), radiation source/spot size, and radiation field penumbra, all aspects of which are relevant for accurate IMRT and VMAT.

**Table 4-1.** Treatment planning system beam modeling parameters requested via survey.

<b>Eclipse</b>	<b>Pinnacle</b>	<b>RayStation</b>
Effective Target Spot Size X and Y [mm]	Effective Source Size X and Y [cm]	Primary Source X Width and Y Width [cm]
MLC Transmission Factor	Rounded Leaf Tip Radius [cm]	MLC Transmission
Dosimetric Leaf Gap [cm]	Tongue and Groove width [cm]	Tongue and Groove [cm]
	Additional Tongue and Groove Transmission	Leaf Tip Width [cm]
	MLC Transmission	MLC Position Offset [cm]
	Flattening Filter Gaussian Height and Width	MLC Position Gain
		MLC Position Curvature [1/cm]

Beam modeling parameter data were acquired via an online survey as part of the IROC Houston Facility Questionnaire, as well as by paper copy issued with phantom tests. The survey was available to all institutions monitored by IROC. To facilitate survey participation and minimize transcription errors, visual instructions (included in the supplemental materials) were provided on how to identify and record the requested beam modeling parameters found in each of the native TPS environments.

The survey was available in the online Facility Questionnaire from January 2018 through January 2019. During this timeframe users were allowed to edit their survey responses such that the most up-to-date information regarding the modeling process could be captured at the time of analysis.

#### *4.2.2 Data Validation*

In total, 2915 individual beam models from 699 institutions were recorded. The results of this survey serve as a broad representation of today's radiotherapy practice; the survey was available to all IROC service users, totaling over 2200 institutions globally, and responses were received from nearly one third of the institutions. Given this breadth of survey respondents and

that the types of responding institutions ranged from single-machine community clinics to large academic hospital centers, we expect that these data are representative of most radiotherapy institutions and that nonresponse bias does not contribute significantly to these survey results.

To ensure the most accurate representation of TPS modeling data, survey responses were examined for gross inconsistencies or typographical errors (for example, recording the MLC transmission value as “1.5”, intending to mean 1.5%, instead of the listed value in the native TPS, “0.015”). Each survey response was checked for the presence of atypical or missing decimal places (indicating wrong magnitude or reported units), illogical values (i.e. negative values for most parameters), or completely blank survey responses. If an unexpected result was obtained, the individual survey response was cross-checked with the institution’s most recent phantom test submissions, which may contain a paper copy of the survey with hand-written responses. These paper copies were assumed free of transcription errors. If corrections to any parameter value could be validated with the institution’s paper survey, the response was amended and included in the analyses; if not, the response was excluded. In a number of cases, institutions elected to submit only some of the applicable parameter values; these partial submissions, so long as the values provided were validated, were retained for analyses. In total, only 95 responses (3.3% of all survey responses) were excluded from analyses, resulting in 2818 usable responses from 642 institutions.

#### *4.2.3 Data Summarization*

Survey results summarized in the supplemental Excel workbook are categorized according to the following: Linac class (e.g. Varian Base, Varian TrueBeam, Elekta Agility), beam energy, MLC configuration (e.g. Millenium 120 versus HD120), and calculation algorithm (for Eclipse; i.e. AAA or AcurosXB). A given Linac class consisted of potentially several machine models that

were deemed to be dosimetrically equivalent, according to previous work performed by IROC Houston.<sup>17,18</sup> In addition to these stratifications, TPS version number was also evaluated as a potential factor for stratification based on the potential for substantive changes to the dose calculation process as version number changed. Ultimately, it was not necessary to separate out the version number because the usable survey responses represented TPS versions for which no substantive changes were made to the dose calculation engine (per the manufacturer) that could, for example, affect the basic modeling or require significant reconfiguration. The TPS version numbers described by this survey data were: Eclipse 10.0+, Pinnacle 9.10+, RayStation 4.7+.

The distributions of survey responses are presented in terms of 2.5th, 25th, 50th, 75th, and 97.5th percentiles to encompass both the interquartile range and the major breadth of parameter values. All data analyses was performed in SPSS Statistics 24 (IBM Corp., Armonk, NY). Percentiles were calculated using the “HAVERAGE” method, which provides an unbiased estimate of the population percentile. Note that for some combinations of Linac class, beam energy, and MLC type, some percentiles (typically the 2.5th and 97.5th) will be undefined because the method is based on a function of the number of cases present; that is, some percentiles may not be defined for smaller subsamples.

## **4.3 Data Format and Usage Notes**

### *4.3.1 Data Format*

The compiled list of survey responses is archived on Zenodo (DOI: 10.5281/zenodo.3357124) as well as the IROC Houston Quality Assurance Center website (<http://rpc.mdanderson.org/RPC/IROCReferenceData.htm>) as an Excel workbook file in the \*.xlsx format. The list is composed of spreadsheet rows, each of which corresponds to a single beam

model reported by an institution. Information included in this dataset include the machine model, beam energy, MLC model, TPS/algorithm, TPS version, and the numeric values associated with the parameter values investigated in Table 1.

Beam modeling parameter distributions are also included as a supplemental Excel workbook file in the \*.xlsx format. The survey results are segregated by beam energy into worksheets named “6 MV”, “6 FFF”, “10 MV”, “10 FFF”, “15 MV”, and “18 MV”. Worksheet titles including “FFF” describe models utilizing flattening filter free beams. Each worksheet is divided into sections by treatment planning system (and algorithm for Eclipse). Linear accelerator classes are differentiated in columns; Varian models are also separated according to MLC model (standard versus high definition MLC) as applicable.

While valid data has been collected, several models using uncommon machines (e.g. Siemens) and beam energies (e.g. 4 MV, 16 MV, 20 MV, etc.) are not presented in the summary distributions due to very limited survey responses; combined, such uncommon models represented only 3% of all survey responses.

#### *4.3.2 Usage Notes*

This data is intended for use as a comparison tool during TPS commissioning or validation studies. The information from this work describes the TPS parameter values that have been deemed clinically acceptable by centers throughout the community, and thus can only provide so much information about the most appropriate values to adopt, particularly in the context of modeling an individual Linac. Through individual measurements and testing, one may determine that the optimal parameter values deviate from the median values shown here. However, even within this limitation, this dataset can serve as a quality check for those who wish to determine whether their

model reasonably follows what others expect for use of a similar Linac/TPS combination based upon the distributions presented. Ideally, this data may also highlight when gross measurement errors occur, such as those associated with determining the dosimetric leaf gap (DLG) or MLC transmission factor.

However, this data is not intended to be used as a reference by which to shortcut the TPS commissioning process. Likewise, the use of popular parameter values may not be the most appropriate for those using treatment specific beam models (e.g. stereotactic radiosurgery, VMAT-dedicated units, etc.). Physicists should be careful to recognize when parameter values presented here may not be ideal for the Linac's intended purpose. Similarly, parameters for which there were few survey responses should also be viewed with greater skepticism. Ultimately, the best test of Linac model value is the agreement between the TPS dose calculation and dosimetric measurement.

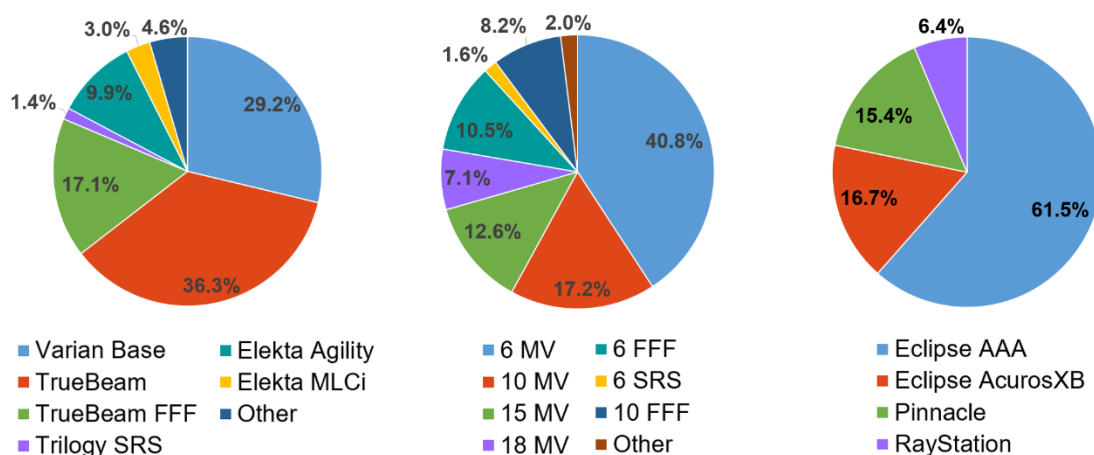
#### **4.4 Discussion**

The goal of this study was to present the community's consensus for the TPS modeling of photon beams (median values and descriptions of the distributions) for the different Linac models and energies that are currently used clinically. To facilitate data interpretations, different MLC models were separately combined, and linear accelerator models were combined into dosimetrically equivalent classes according to reference dosimetry data from IROC-Houston site visit data.<sup>55,56</sup> Reference data was compiled so that physicists can compare their own input values to those shown in this work.

Because the dataset presented here is for informational purposes, this work does not include interpretation or hypotheses. This work does not, therefore, include assessment of the impact of

different parameters on underlying dose calculation accuracy, nor does it attempt to identify “unsuitable” parameter values. This work simply includes raw survey responses as well as general descriptions and trends observed within the survey results.

As depicted by the distributions in Figure 4-1, the greatest proportion of survey responses were for Eclipse (78.2%), a Varian product (84%), or 6 MV treatment beams (40.8%). These proportions followed what was expected based on IROC Houston service user demographics.

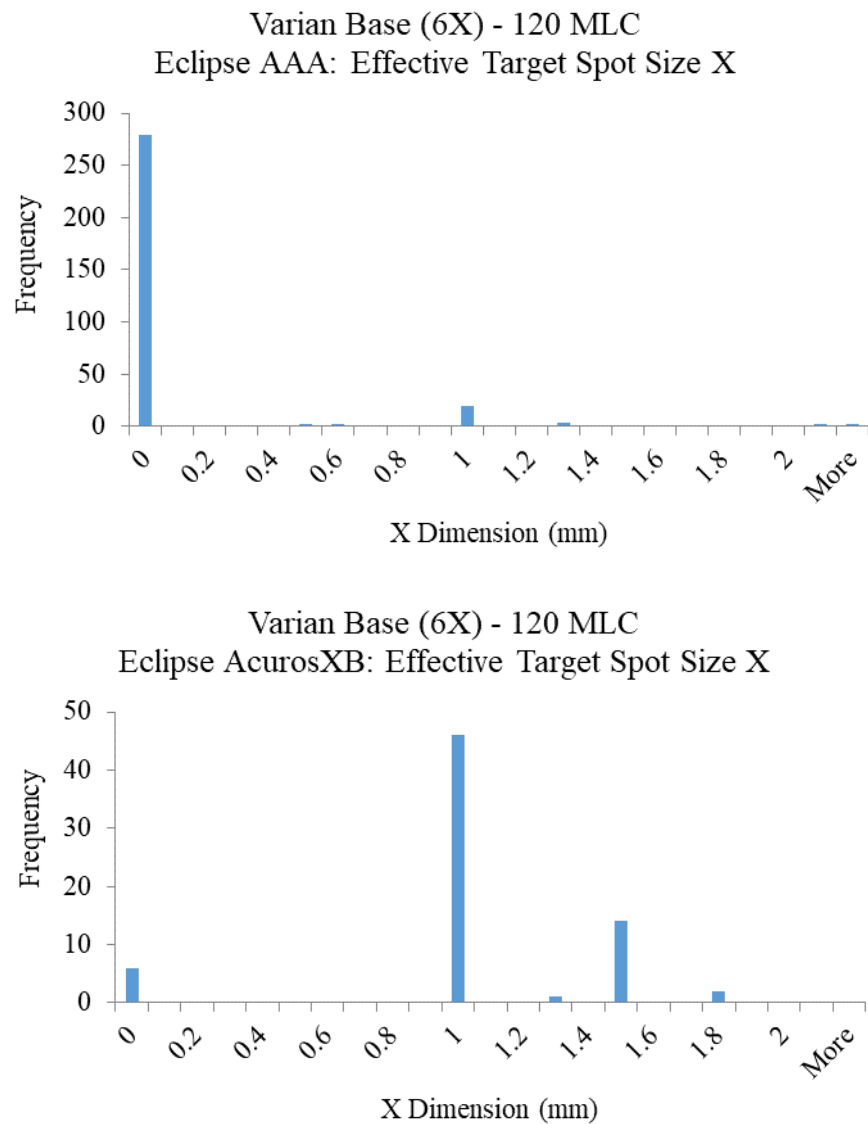


**Figure 4-1.** Distributions of survey responses according to linear accelerator class (left), beam energy (center), and TPS/algorithm (right).

Interestingly, there exist cases where the radiotherapy community shows substantial agreement, particularly concerning the spot size in Eclipse. Nearly all participants using Varian machines and the AAA algorithm (regardless of beam energy) opted to model a point source (0 mm x 0 mm), and AcurosXB users modeled the spot size as 1 mm x 1 mm (Figure 2). This extensive agreement may be due to, in part, by the proprietary auto-modeling features found in Eclipse, which pre-fill model parameter values based on pre-measured dose profiles. This is in direct contrast to Pinnacle, where users individualize the source size and many other parameters

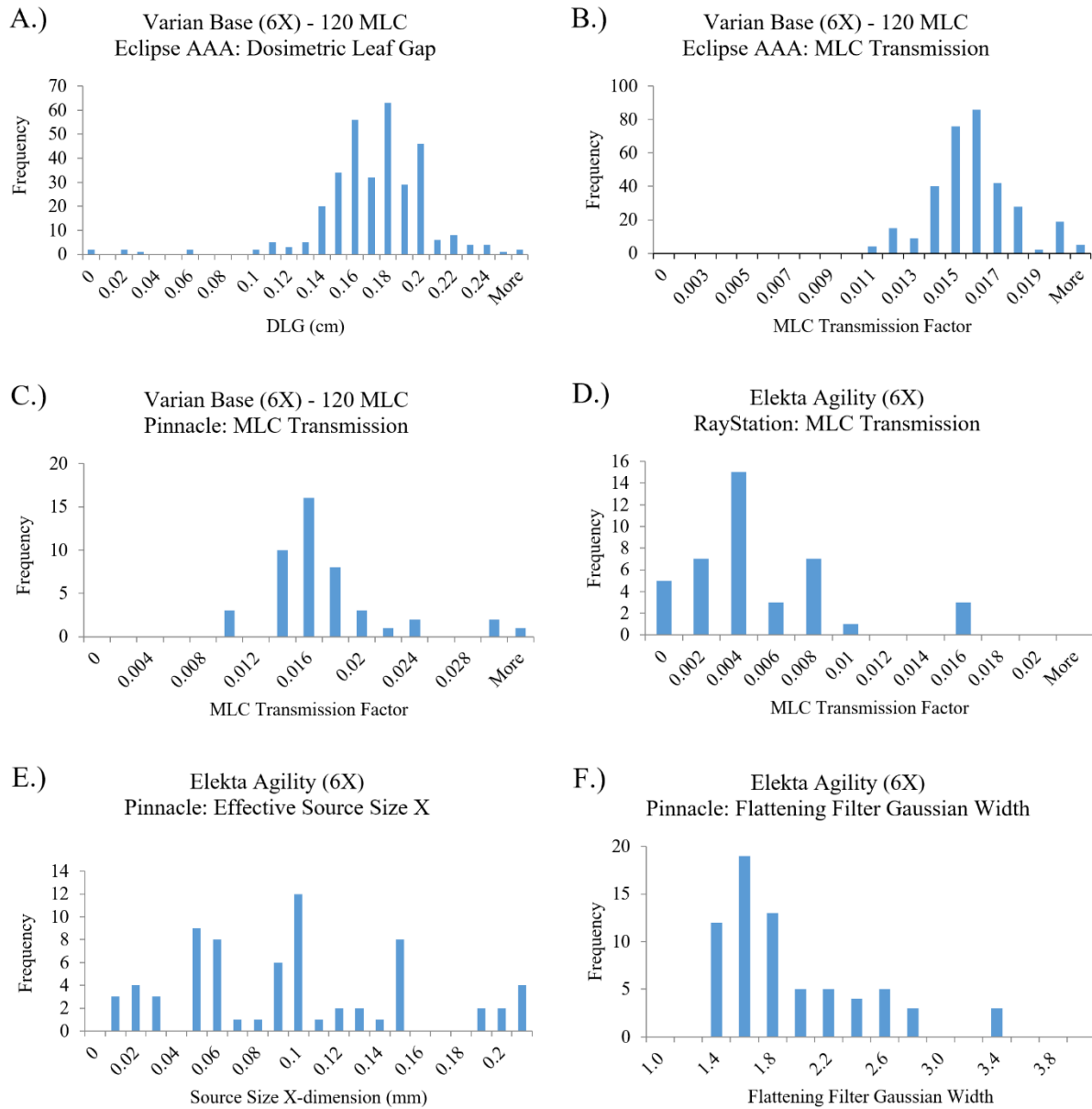


from the start. This result suggests that computer-aided decisions may decrease the variations exhibited, although the value or risk associated with such a process is unclear from these data.



**Figure 4-2.** Survey responses for effective target spot size used to model standard 6 MV Varian Base class machines in Eclipse (top: AAA; bottom: AcurosXB). Both algorithms show uniformity of response.

The histograms of Figure 3 demonstrate where the radiotherapy community shows the least consensus in the development, commissioning, and validation of their beam models. Here the DLG (from Eclipse), MLC transmission (for all three TPS), and effective source size and flattening filter Gaussian width (for Pinnacle) show notable variations, even among similar Linacs. Some of this variation in Pinnacle can be attributed to the many degrees of freedom that physicists have in model creation. Other variations can be explained by the way the DLG and MLC transmission are developed; physicists often physically measure these factors following different protocols and using different equipment, leading to different measured values. For example, measurement of DLG, a parameter found to be critical in the accuracy of IMRT,<sup>9,15,57</sup> has been found to vary based on the size of ion chamber used to measure it.<sup>58</sup> This underscores the necessity for physicists to critically analyze the measurement and validation process to ensure that, ultimately, the value used in their model is most adequate for their dose calculations.



**Figure 4-3.** Subset of survey responses for standard 6 MV Varian Base class and Elekta Agility Linacs that depict high variability in parameter value agreement: a) Eclipse AAA DLG, b) Eclipse AAA MLC transmission, c) Pinnacle MLC transmission, d) RayStation MLC transmission, e) Pinnacle source size X-dimension, f) Pinnacle flattening filter Gaussian width.

One aspect of beam modeling that cannot be properly assessed in this work is the potential interplay between different parameters. This may be of particular interest and may be more pronounced for TPSs with more parameters (i.e. more degrees of freedom) to explore in beam

model creation. Should interplay be present among the parameters presented, including the potential to offset or exacerbate errors, interpretation of these survey results may be more challenging. Thus, more work assessing such factors and their associated effects is warranted. In fact, such work may be of particular interest given that for institutions modeling a Varian Base class machine (with standard MLC) using Pinnacle, it was exceedingly uncommon for a user to report all parameter values within the interquartile range (25% - 75%); that is, nearly all Pinnacle beam models had at least one parameter value outside this range. This was not the case for common models like the Varian Base class for Eclipse or RayStation, where at least 25% of all responses reported all values within the defined interquartile range. However, such observations suggest that additional benefit may be derived from follow-up analyses determining how different parameter values and/or combinations can affect the overall beam model accuracy.

Some limitations of the current study are that not all survey results could be represented in this dataset because certain Linacs did not fit within the predefined classes (e.g. Siemens machines were excluded from presentation). Additionally, the set of parameters presented herein only represent a subgroup of all the modeling parameters available in TPS commissioning. Users should be mindful of factors that may affect the aforementioned parameters or properties that may substantially alter the overall model applicability, such as the tongue and groove effect.<sup>59</sup> Lastly, the survey data herein can only represent the most up-to-date interpretation of the TPS models through 2018. Should further TPS upgrades occur that drastically affect the way the TPS calculates dose, the data found in this set may be of historic interest, rather than prospective. Given the stability of the dose calculation algorithm in the current TPS algorithms, this data may retain utility for an extended period of time.

## 4.5 Conclusions

In this study, beam modeling parameters for the Eclipse, Pinnacle, and RayStation TPS were compiled and analyzed to create a reference dataset describing how the radiotherapy community assigns parameter values to generate its clinical beam models. Statistical metrics were provided so that physicists examining the commissioning of their TPS may recognize parameters that require greater attention and consideration to ensure the most accurate and robust models possible. This dataset can be used as a second check for physicists during the TPS commissioning process to detect what may contribute to anomalies in the beam model that could warrant further attention. This work also highlights considerable variations among several critical parameters used in beam modeling, thus providing additional caution.

## Chapter 5: Sensitivity Analyses of Common Beam Modeling Parameters

This chapter is based upon the following publication:

**M. Glenn,** C. Peterson, R. Howell, D. Followill, J. Pollard-Larkin, S. Kry. “Sensitivity of IROC phantom performance to TPS beam modeling parameters based on community-driven data.” Submitted for publication, 2020.

Additional materials regarding this work can be found in Appendix D.

### 5.1 Introduction

Radiation dose calculation accuracy is contingent upon how well the treatment planning system (TPS) mathematically represents the physical photon beam under the conditions used for radiation therapy. Good commissioning and validation of the beam model is fundamental, for once established, this model is used to calculate the dose for all treatments done with the radiation beam. However, modern technologies such as intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) pose a particular challenge for TPS dose calculations due to the modulation and constraints needed to provide highly conformal dose distributions. These additional challenges allow increased uncertainty in the dose calculation.

In order to properly assess the accuracy of IMRT, external validation tests are suggested for individual institutions, as well as for clinical trials.<sup>49,60,61</sup> The Imaging and Radiation Oncology Core Houston Quality Assurance Center (IROC-H) provides anthropomorphic phantom credentialing for National Cancer Institute-sponsored multi-institutional clinical trials using IMRT to ensure treatments are delivered as intended while minimizing uncertainty. Over the years,

IROC-H has observed a broad range of IMRT treatment performance.<sup>8</sup> In particular, recent works from IROC-H indicate a substantial presence of phantom results showing systematic dose errors,<sup>4</sup> and poor TPS dose calculations in failing phantom cases.<sup>16</sup> Additionally, there exists substantial evidence that standard quality assurance (QA) methods, including IMRT QA, fail to detect unacceptable plans and errors related to the TPS.<sup>10–12,43</sup>

Due to these challenges, interest has developed in understanding how beam modeling, and which specific factors within the model, can contribute to poor plan performance. Previous studies have investigated the relative errors that several modeling factors related to the multileaf collimator (MLC) can contribute to the overall accuracy, as well as the detectability of these errors.<sup>9,62,63</sup> While generally informative, such works have been relative to single clinical systems, and thus cannot provide wide-ranging context into other clinical scenarios. More problematically, the magnitude of change in each parameter (i.e., how much error is introduced into the MLC offset), and associated effect size, have not been based on clinically realistic values. That is, the ranges of values for modeling parameters used in these works are, in general, arbitrary and may not necessarily be relevant to current practice.

Instead, this study evaluated the impact of beam modeling errors (both dosimetric and non-dosimetric) that are consistent with the errors seen clinically, or are consistent with the range of values used in clinical practice. TPS errors in basic dosimetric data, such as percent depth dose (PDD) measurement, have been previously reported by IROC-H based on measurements of over 1000 linear accelerators (linacs).<sup>64</sup> Non-dosimetric data, such as MLC leaf offset and MLC transmission factor, have been compiled in a recent IROC-H survey that included over 2800 beam models from 642 institutions.<sup>65</sup> In this study we used these values from the community to determine the degree of change introduced into each parameter. In this way, this study investigated

the potential dosimetric impact of using beam modeling parameter values that are either erroneous or at least deviate from typical as established by the radiation oncology community. This data can inform the ways that errors can and do manifest for IMRT and VMAT treatments across the community at large. Understanding the expected error contributions of erroneous or atypical parameter values can also help explain and rectify the ongoing suboptimal IROC-H phantom performance rates by providing more in-depth guidance to the dose calculation variations that may exist.

## **5.2 Methods and materials**

### *5.2.1 TPS beam model creation and validation*

Beam models representing a 6 MV beam on a Varian Clinac-type machine with Millennium 120 multileaf collimator (MLC) were developed in Eclipse v13.5 with AAA algorithm (Varian Medical Systems, Palo Alto, CA) and RayStation 9A v8.99 with collapsed cone algorithm (RaySearch Laboratories, Stockholm, Sweden). Dosimetric characteristics were tailored to the linac-specific reference data from IROC-H's site visit program, which encompasses 23 output measurements (e.g. percent depth dose curves, output factors, and off-axis factors).<sup>55</sup> Non-dosimetric modeling parameters (e.g. source size, MLC leaf-tip offset, etc.) were defined to match median beam modeling parameters as reported by the radiation oncology community in an IROC-H survey (Glenn, et al.<sup>65</sup>). This was specific for the Varian Base class linac, defined in Kerns, et al.,<sup>55</sup> with standard 120-leaf MLC in this study. In this way, the most representative linac (of a widely-used model) was created.

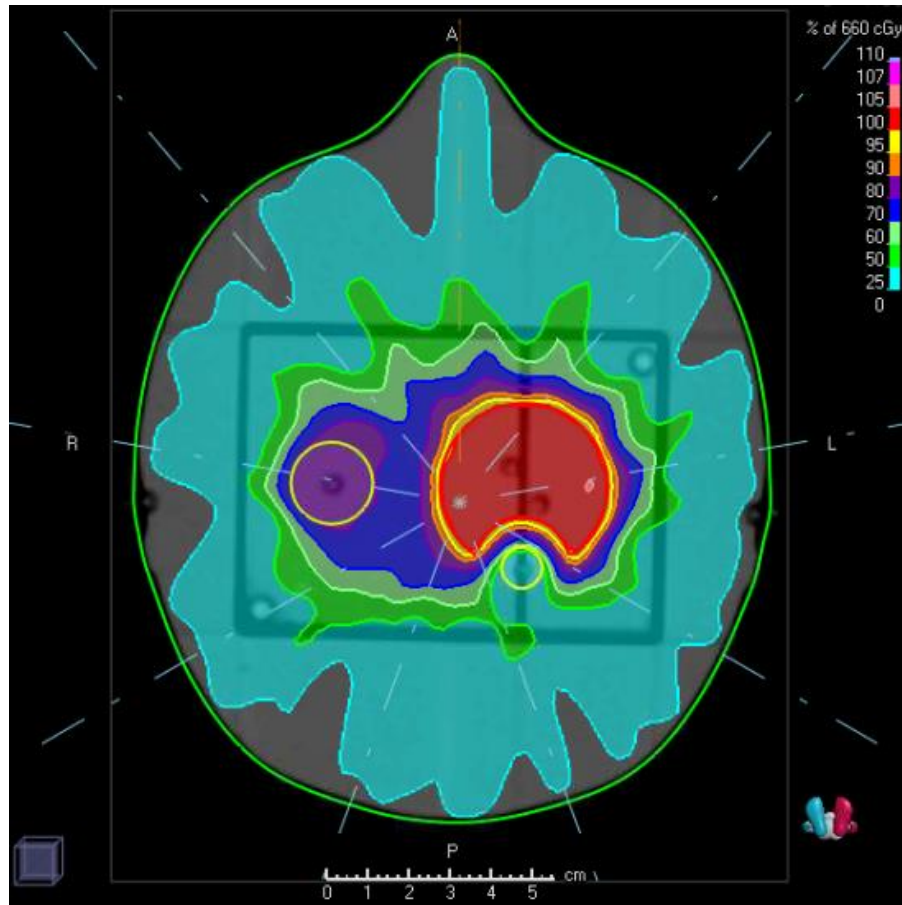
Baseline beam models for Eclipse and RayStation were then validated via two IROC-H head and neck (H&N) phantom irradiations on a clinical Varian Trilogy linac. The two plans, a



standard 9-field IMRT and 2-arc VMAT, were assessed for agreement between thermoluminescent dosimeter (TLD) dose and TPS-reported dose as calculated by both Eclipse and RayStation to ensure clinical applicability and reasonability of the baseline models prior to further manipulation and study.

### *5.2.2 Phantom plan development*

The IROC-H H&N phantom was scanned on a CT simulator following standard clinical workflow, and five IMRT plans (3 IMRT, 2 VMAT) were developed in Eclipse 13.5 and imported to RayStation for consistency across platforms. These plans were designed for dynamic IMRT delivery in a single fraction and follow dose prescription guidelines provided by IROC-H (Figure 5-1). Monitor units (MU) and general plan setup for each of the 5 plans are detailed in Table 5-1. All IMRT fields were planned about the center of the H&N phantom and spaced equidistantly, as is common in clinical practice. Likewise, the VMAT plans were developed for isocentric delivery with full 360° arcs. These plans were developed with a variety of complexities and beam angles to encompass a range of treatment strategies and plans as previously observed by IROC-H.<sup>66</sup>



**Figure 5-1.** Axial slice of the 9-field IMRT plan developed for the IROC-H H&N phantom. The phantom contains primary and secondary targets planned to be treated at 660 cGy and 540 cGy, respectively.

**Table 5-1.** Summary of H&N phantom plans developed for testing.

Plan Name	Delivery Type	Number of Beams	Total MU
IMRT5	IMRT	5	3741
IMRT7	IMRT	7	2470
IMRT9	IMRT	9	2729
VMAT1	VMAT	1 arc (360°)	1990
VMAT2	VMAT	2 arcs (each 360°)	2130

Note: Phantoms were prescribed 660 cGy to the primary target, delivered in a single fraction. All plans were developed for dynamic delivery.

### *5.2.3 Parameter manipulation (simulated beam model deviations) and evaluation*

The baseline models, in Eclipse and RayStation, had all parameters at the 50<sup>th</sup> percentile community value. In order to simulate the impact on dose agreement from variations that have been shown to exist in the radiotherapy community, permutations of the baseline beam models were created by individually manipulating parameters of interest within each TPS environment. The parameters of interest considered in this study, including both basic dosimetric characteristics and TPS-specific modeling parameters, are outlined in Table 5-2. When manipulations of a parameter were introduced, all other parameters were maintained at the baseline (50<sup>th</sup> percentile) value in order that effects may be isolated. Variations in beam modeling parameters were introduced as the 2.5<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 97.5<sup>th</sup> percentiles of beam modeling parameter survey responses.<sup>65</sup> In order to characterize variance in dosimetric characteristics, the baseline beam model was modified in RayStation by changing the photon spectrum, square-field output factor correction factors (jaw-defined fields), and off-axis factors. Changes in dosimetric parameters were made for the same percentiles (2.5<sup>th</sup> to 97.5<sup>th</sup>), based on reported measurement accuracy versus TPS calculation from IROC-H's site visit program.<sup>64</sup>

**Table 5-2.** Beam modeling parameters investigated for sensitivity

<b>Eclipse</b>	<b>RayStation</b>	<b>Dosimetric Characteristics</b>
Effective Target Spot Size	Primary Source Size	Percent Depth Dose (PDD)
MLC Transmission Factor	MLC Transmission	Small-field Output Factors*
Dosimetric Leaf Gap (DLG)	Leaf Tip Width	Off-Axis Factors
	Tongue and Groove	
	MLC Leaf Tip Offset	
	MLC Gain	
	MLC Curvature	

\* Jaw-defined small field output factors

Following beam model modifications, each of the five H&N phantom treatment plans was recalculated (plans were not reoptimized in this process) and compared to the baseline model and evaluated for changes in average dose across the 6 TLD locations in the phantom and average dose of the 2 TLDs in the organ-at-risk (OAR).

#### 5.2.4 Parameter interplay

To understand potential interdependencies in beam modeling parameters, we investigated the effect of changing multiple parameters simultaneously. We could then evaluate if, for example, two different parameters, both set simultaneously to the 97.5<sup>th</sup> percentile, had a different impact on the dose distribution than simply the sum of the two when set sequentially to the 97.5<sup>th</sup> percentile. Parameters that were investigated were only those that, by themselves, had a sizeable impact on the dose distribution; in particular, those parameters that introduced changes in average TLD dose greater than 1% across the 5 treatment plans were selected. These impactful parameters were then varied pair-wise against other impactful parameters in the same TPS. For this exercise, the most extreme values (2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles) were adopted for one parameter while the

other was varied across the distribution to assess the greatest extent to which interplay can occur. Interplay between parameters were assessed through regression analysis with interaction modeling. Additionally, the dosimetric effect of the combined scenario was visually compared to simply summing the average effects together (thus representing complete independence).

## **5.3 Results**

The baseline TPS models, used to irradiate the phantom for validation, showed agreement between measurement and calculation of within 5% across all locations for both planning systems and both plans. This level of agreement exceeds IROC-H's acceptability criterion and indicated that the baseline beam models are a good representation of a standard linac.

### *5.3.1 Eclipse*

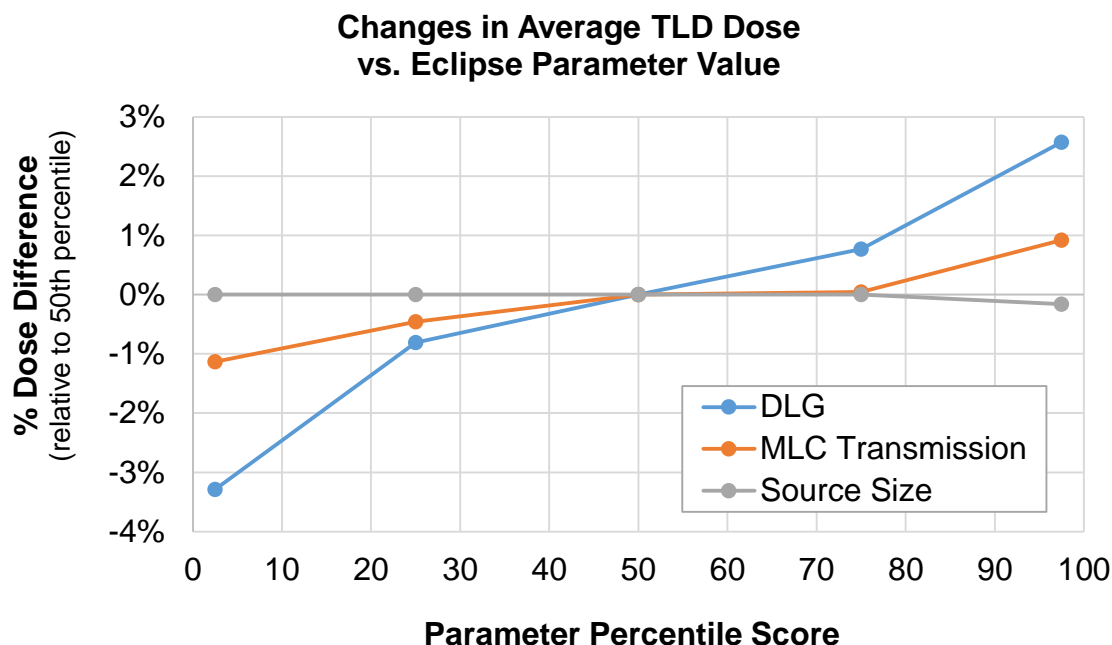
Table 5-3 describes the numeric parameter modifications implemented in Eclipse, based on the values from Glenn, et al.<sup>65</sup> Figure 5-2 shows the average percent difference in calculated TLD dose (over the 6 TLDs within the two targets, averaged across all five plans) between the baseline and modified beam models as a function of the percentile score for the parameters of interest as reported by the radiation oncology community. Based on the range of values from the community, the variations in DLG have the greatest impact on dosimetric accuracy. This was true for all endpoints examined: changes up to 6% were observed for target TLD dose, and changes up to 10% were observed in the OAR TLDs.

The dosimetric impact across the TLDs is plotted against actual DLG value in Figure 5-3. From this figure it can be seen that the average dose to the target changes approximately linearly with DLG value for all plans examined (Figure 5-3). Moreover, for all plans except IMRT5, the

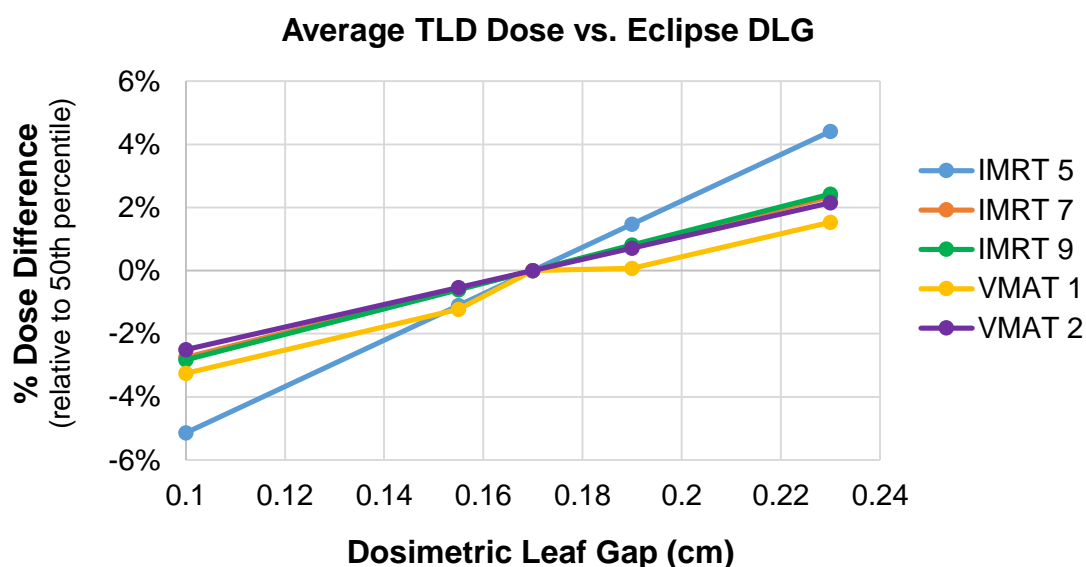
relative change in dose caused by DLG variations was consistent despite differences in treatment complexity. IMRT5, the plan that was very highly modulated, had a notably different slope than the others, and was more sensitive to changes in DLG. The other beam modeling parameters produced much smaller dose deviations; changes in dose caused by manipulation of the effective target spot size and MLC transmission factor were generally less than 1% in the target volume doses and <5% change in OAR dose.

**Table 5-3.** Parameter changes implemented for the 6MV Varian Clinac-type machine simulated in Eclipse

<b>Percentile</b>	<b>Effective Target Spot Size X [mm]</b>	<b>Effective Target Spot Size Y [mm]</b>	<b>Dosimetric Leaf Gap [cm]</b>	<b>MLC Transmission Factor</b>
2.5th	0.000	0.000	0.1000	0.0118
25th	0.000	0.000	0.1550	0.0145
50th	0.000	0.000	0.1700	0.0158
75th	0.000	0.000	0.1900	0.0165
97.5th	1.250	1.000	0.2300	0.0200



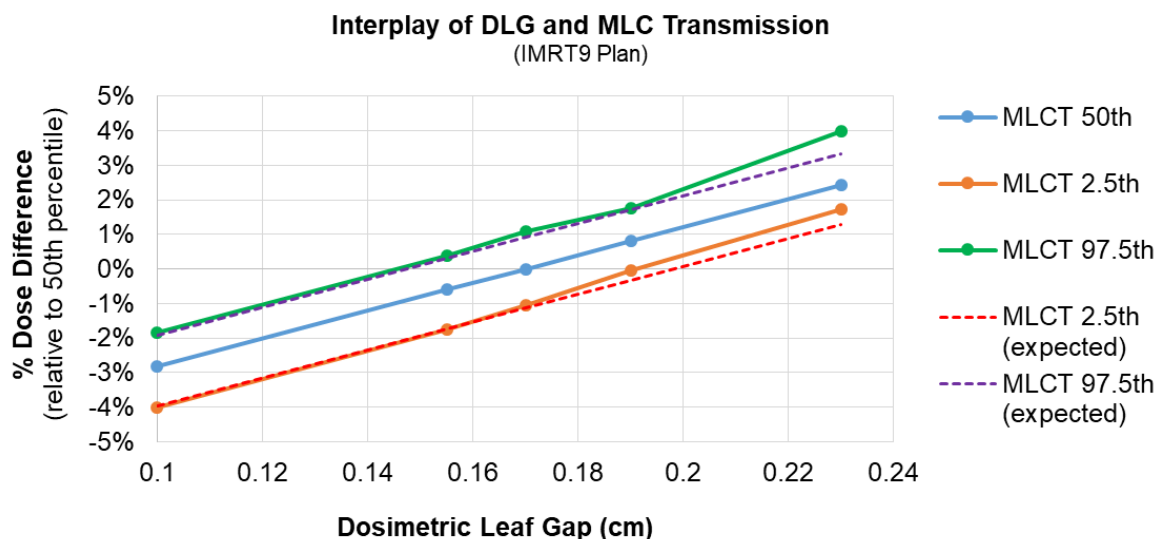
**Figure 5-2.** Average changes in dose calculated to target TLD structures for IROC-H head and neck phantom plans when parameters of interest are manipulated in Eclipse. Percentile score corresponds to community reported values.<sup>65</sup> DLG = dosimetric leaf gap.



**Figure 5-3.** Changes in average TLD dose calculated for each of the five IROC-H head and neck phantom plans following manipulation of the dosimetric leaf gap (DLG) in Eclipse. Points on the curve correspond to percentile scores from the radiation oncology community for a Varian Base Class machine with a Millennium 120 leaf MLC.

To further investigate these results, we assessed the interdependence of the DLG and MLC transmission factor, given that both produced differences in average TLD dose greater than 1%. Figure 5-4 depicts a comparison between the cases for which only the DLG was varied and the MLC transmission factor was left at the 50<sup>th</sup> percentile value (i.e. MLCT 50<sup>th</sup>) and when the DLG and MLC were varied together (with the MLC transmission factor defined at the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles). Regression modeling of the calculated curves determined that the two parameters behave linearly ( $p < 0.001$ ) and interaction was not significant. That is, there is no evidence to suggest that the effect of DLG on dose is different for different values of MLC transmission factor. Additionally, these were compared with the cases where average change in dose caused by variation in MLC transmission factor (here, -1.14% for the 2.5<sup>th</sup> percentile and +0.92% for the 97.5<sup>th</sup> percentile) was simply summed with the DLG-only case, thereby assuming the DLG and MLC transmission factor behaved independently (the “expected” cases). The calculated and expected dose difference curves are consistent, further emphasizing the linearity and predictability of dose response with respect to these parameters.





**Figure 5-4.** Interplay between dosimetric leaf gap (DLG) and MLC transmission (MLCT). “DLG only” varies DLG while the MLCT remains at the 50<sup>th</sup> percentile value. MLCT 2.5<sup>th</sup> and MLCT97.5<sup>th</sup> illustrate varying DLG while the MLC transmission is set to those percentile values. “Expected” cases describe the change in dose if individual, average parameter effects were simply summed together. Note that this comparison is for the IMRT9 treatment plan.

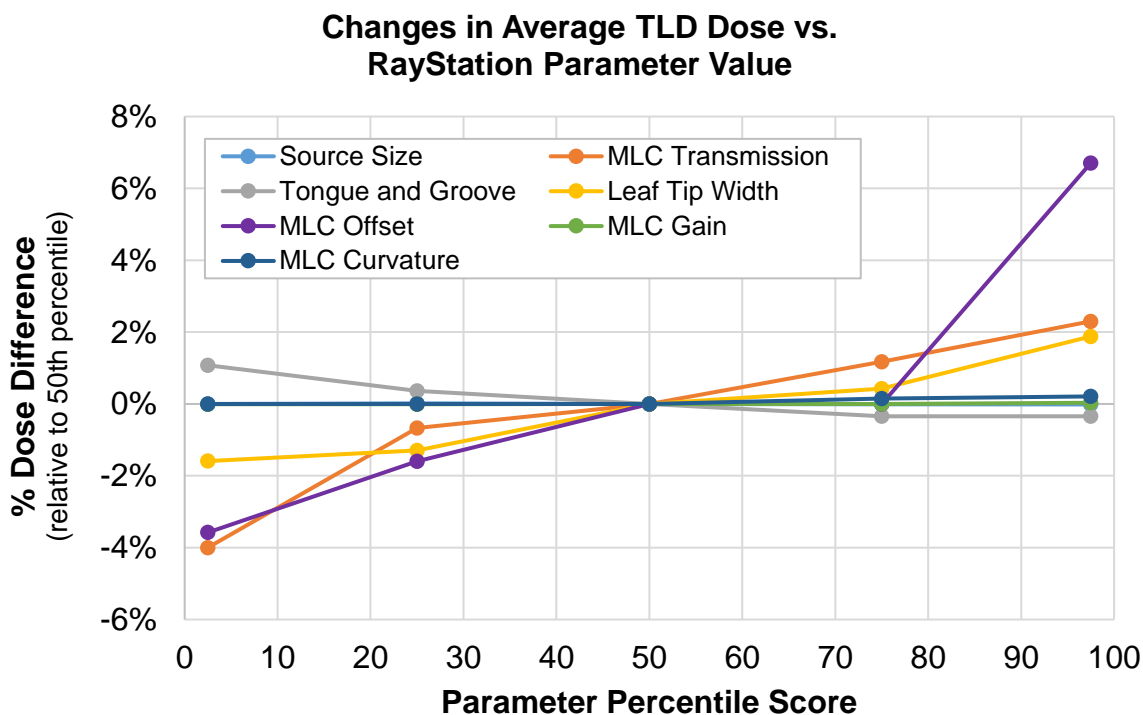
### 5.3.2 RayStation

Table 5-4 describes the numeric parameter modifications implemented in RayStation, based on the values from Glenn, et al.<sup>65</sup> Figure 5-5 shows the changes in average TLD dose in the target relative to the baseline (50<sup>th</sup> percentile values) as a function of percentile score. Analogous to the DLG in Eclipse, dose calculation accuracy had the strongest dependence on the MLC position offset: changes in MLC position offset produced as much as a 13% change in the target TLD dose and 25% in the OAR dose.

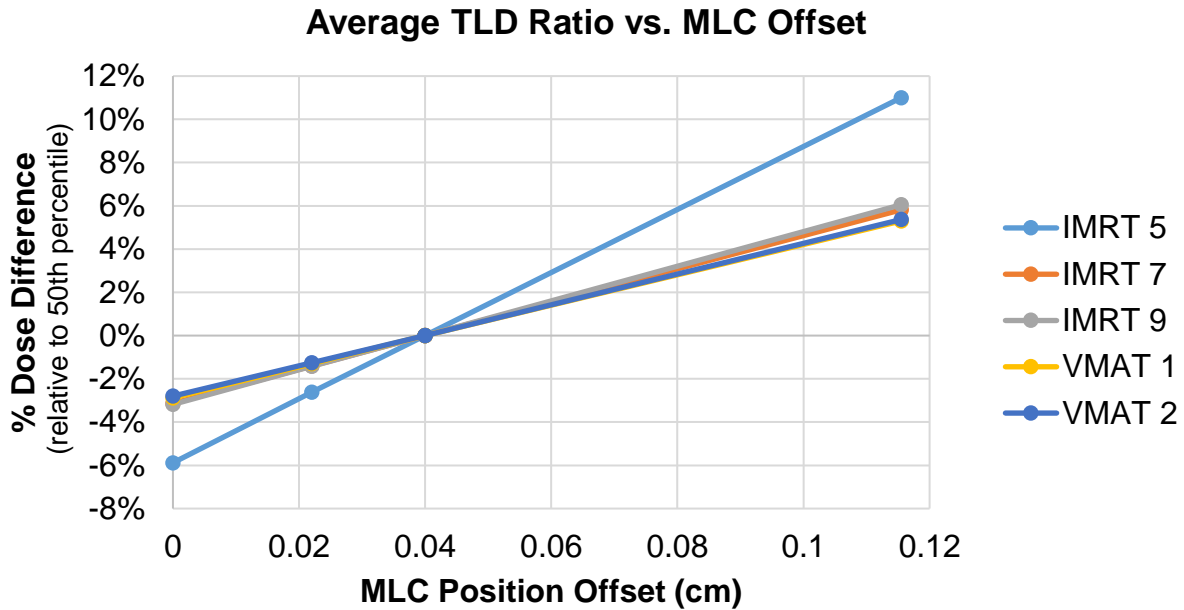
The dosimetric impact across the target TLDs is plotted against MLC Offset value in Figure 5-6. From this figure it can be seen that the average dose to the target changes linearly with MLC offset for all plans examined (Figure 5-6). As with DLG, the MLC offset was most sensitive to the IMRT5 plan, which was very highly modulated, and was uniformly sensitive to all of the other plans.

**Table 5-4.** Parameter changes implemented for the 6MV Varian Clinac-type machine simulated in RayStation.

Percentile	Primary Source Size X [mm]	Primary Source Size Y [mm]	Tongue and Groove [cm]	Leaf Tip Width [cm]	MLC Transmission	MLC Position Offset [cm]	MLC Position Gain	MLC Position Curvature [1/cm]
2.5th	0.04000	0.05000	0.01	0.177	0.007	0	0	0
25th	0.05350	0.05200	0.03	0.200	0.016	0.022	0	0
50th	0.05700	0.07000	0.04	0.320	0.018	0.040	0.0015	0
75th	0.10300	0.07250	0.05	0.360	0.022	0.040	0.0015	0.0008
97.5th	0.12345	0.10075	0.05	0.500	0.025	0.116	0.0150	0.0010



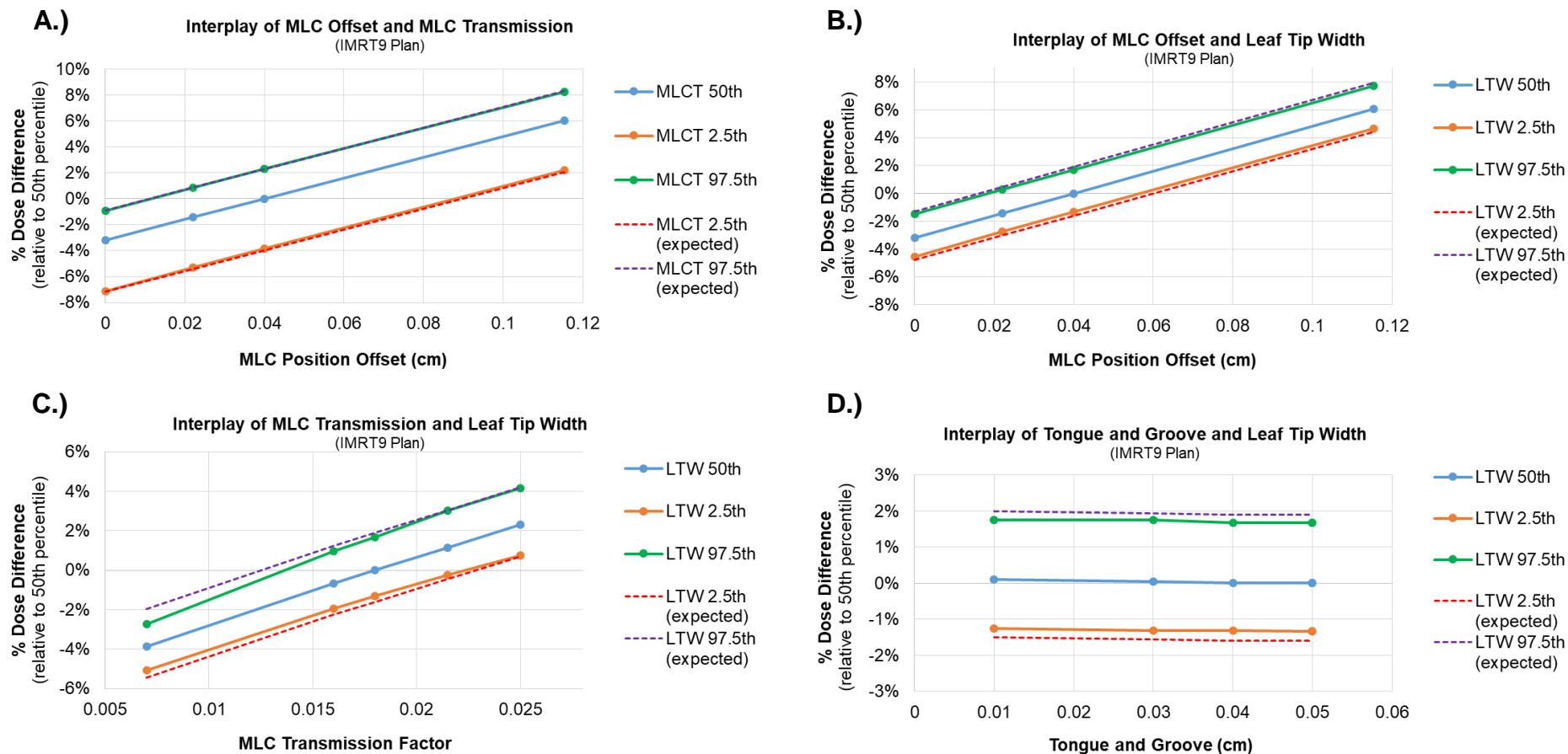
**Figure 5-5.** Average changes in dose calculated to TLD structures for IROC-H head and neck phantom plans when parameters of interest are manipulated in RayStation. Percentile score corresponds to community reported values.<sup>65</sup>



**Figure 5-6.** Changes in average TLD dose calculated for each of the five IROC-H head and neck phantom plans following manipulation of the MLC offset in RayStation. Points on the curve correspond to percentile scores from the radiation oncology community for a Varian Base Class machine with a Millennium 120 leaf MLC.

Other parameters related to the MLC, including the MLC transmission factor, tongue and groove, and leaf tip width, also greatly impacted the resultant dose recalculations. For MLC transmission, average changes in target TLD dose ranged from -4% to +2% with singular points up to 10% off in the OAR TLDs. Changing the leaf tip width generated more moderate changes in the target TLD dose (up to 3%) but produced changes as high as 15% in the OAR TLDs. Interestingly, the tongue and groove produced negligible changes in any of the IMRT plans, but introduced uniform and consistent changes in the target TLD doses up to 3% in both VMAT plans. In general, the differences in dose as a result of model manipulation were relatively consistent across the 5 plans examined, with exception to the tongue and groove, which had the greatest effect on VMAT plans. The other parameters examined, based on the distributions of community data, produced minimal changes in plan dose.

Interplay was examined for the MLC position offset, MLC transmission factor, leaf tip width, and tongue and groove using the IMRT9 plan. Figure 5-7 shows several tested relationships where the actual effect of combining parameter changes were compared against cases for which the average change in dose was simply summed based on the impact of each individual parameter (assuming they were independent). The results, shown in Figure 5-7, illustrate consistently that the two approaches (calculating interplay and simply adding effects assuming they're independent) yielded similar results. This is corroborated through linear regression and interaction modeling. Only the interaction between MLC transmission factor and leaf tip width (Figure 5-7C) was found to be significant ( $p < 0.001$ ); however, the adjusted coefficients of determination for the dependent and independent cases were very similar ( $R^2$  equal to 0.998 versus 0.994, respectively), indicating that this interaction does not contribute largely to the overall change in dose. All other examined parameter combinations likewise demonstrate linear effects on dose ( $p < 0.001$ ) with no significant interaction among impactful parameters. That is, there was little if any additional compounding or suppressing of changes in average dose relative to the expected cases, demonstrating that these parameters had minimal interdependence over the range used clinically.



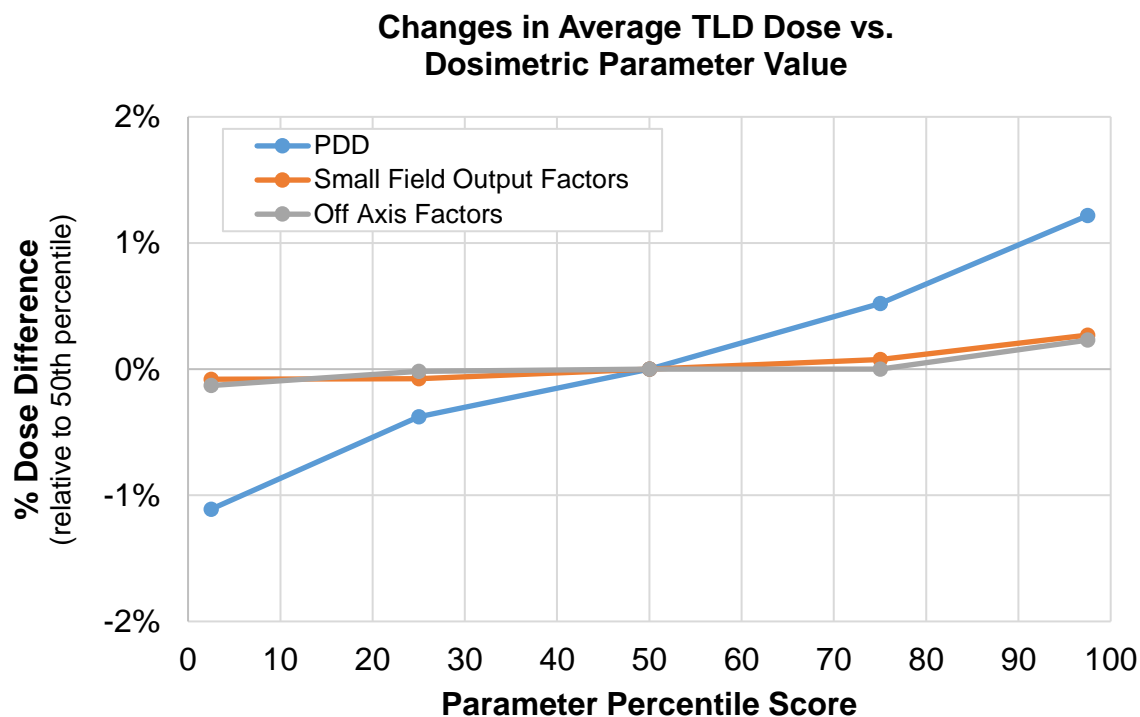
**Figure 5-7.** Select cases demonstrating the extent of interplay between (A) MLC offset (MLCO) and MLC transmission factor (MLCT), (B) MLC offset and leaf tip width (LTW), (C) MLC transmission factor and leaf tip width, and (D) tongue and groove (T&G) and leaf tip width in RayStation using the IMRT9 plan. Each value was evaluated at the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile and compared to the “expected” cases, which described the change in dose if individual parameter effects were simply added together.

### 5.3.3 Dosimetric Parameters

Table 5-5 describes the numeric dosimetric modifications implemented in the beam modeling to achieve the variance observed in the community between modeled and true values.<sup>55</sup> Figure 5-8 shows the relative changes in average TLD dose introduced by modifying the underlying beam model according to the percentile value for each parameter. The only parameter to introduce changes in the target TLD dose in excess of 1% was the PDD, which on average ranged from -1.1% to +1.2% with a maximum change of 2.5% in the OAR TLDs. Unsurprisingly, changes in target TLD dose were less than the PDD errors introduced because the depths of the targets were less than 10 cm. The small field output factors had a wide range of errors as measured by IROC;<sup>64</sup> however, these errors did not manifest as a variation in TLD dose in this study. This is because the output factors defined in Table 5-5 are defined by the jaw. The span across the primary and secondary targets is approximately 6-8 cm across, so no small jaw-defined fields are typically used in treating the phantom. Similarly, the off-axis factors have known errors as shown in Table 5-5, but these did not materialize in the phantom cases. This is because the targets are relatively small and close to central axis.

**Table 5-5.** Dosimetric parameter changes implemented for the 6MV Varian Clinac-type machine simulated in RayStation.

<b>Percentile</b>	<b>PDD</b> (20 cm depth)	<b>Small-Field Output Factors</b> (at 2x2 cm)	<b>Off-Axis Factor</b> (10 cm from central axis)
2.5th	-2.3%	-5.8%	-2.0%
25th	-0.7%	-0.5%	-0.4%
50th	0%	0%	0%
75th	+0.6%	+0.4%	+0.5%
97.5th	+2.5%	+2.3%	+1.6%



**Figure 5-8.** Average changes in dose calculated to TLD structures for IROC-H head and neck phantom plans when dosimetric characteristics are manipulated in RayStation according to the documented variations in measurement accuracy from IROC-H site visit data.<sup>64</sup>

Because PDD was the only factor to contribute plan changes greater than 1%, interplay effects among dosimetric characteristics were not assessed.

## 5.4 Discussion

This study explored the dosimetric effects of changing common beam modeling parameters on clinically-acceptable H&N phantom treatment plans to understand how these changes may contribute to dose calculation accuracy or inaccuracies in the context of IMRT and VMAT planning. Small variations in MLC offset, as described by either the DLG in Eclipse or MLC leaf-tip offset in RayStation, can have substantial impacts on the resultant plan performance, affecting

both the targets and OAR doses substantially. Other parameters modeling the MLC and radiation source characteristics, based on community-reported data, were impactful to a lesser degree.

The magnitudes of dose differences caused by variations in beam modeling are well-corroborated by previous studies in Eclipse and RayStation. Like that reported here, McVicker, et al.<sup>9</sup> demonstrated changes in IROC-H H&N phantom target TLD dose on the order of 5% or more for changes in DLG greater than 1mm while the MLC transmission factor more greatly affected OAR TLDs than those in the target structures. Additionally, Kielar, et al.<sup>67</sup> demonstrated the linear proportionality between error in defining the DLG and dose error, much like that shown in Figure 2. For RayStation, Koger, et al.<sup>62</sup> also demonstrated the linear response of MLC position offset and observed, despite using a different machine setup, that additional offsets of 1 mm can produce changes in TLD dose in excess of 10%, much like Figure 5-1. Likewise the leaf tip width was found to be a parameter of importance in accurate MLC modeling, generating TLD dose differences of up to 2% in the targets for the range of values examined.

At first glance, the results from this work may seem unsurprising; multiple previous studies have likewise noted a strong dependence of the dose calculation for modern treatments (IMRT and VMAT) with the MLC leaf position.<sup>11,63,68-72</sup> However, the current work is unique in that the dose differences calculated herein are based upon the actual variations in beam modeling adopted by the radiotherapy community, making the magnitude of the dose deviations particularly relevant to clinical practice. Moreover, these differences were evaluated in a reference geometry (the IROC-H H&N phantom) where the radiation oncology community is known to struggle. To maximize the breadth of impact, these evaluations were conducted using several common IMRT delivery methods. Additionally, this work examines interplay among parameters as a factor of IMRT performance, which has not been extensively investigated to date.



In Glenn, et al.,<sup>65</sup> the authors state that the beam modeling parameters that had the greatest spread among the community are those representing the MLC characteristics. Interestingly enough, these very same parameters, namely the MLC transmission factor, DLG in Eclipse, and the MLC offset in RayStation, generated the most considerable dose changes among the treatment plans studied herein. But more importantly, these very same factors were also ones that could be theoretically measured, thus underscoring the need for physicists to be extremely cautious when assigning these values in their clinical TPS. In fact, current guidance for TPS commissioning is limited to several tests to validate the overall TPS performance; recommendations for individual parameter assignment are limited to the TPS vendor. It is solely up to the physicist to understand the intended effects of parameter assignment and consequently understand what values will generate the most robust model. For example, both Kim, et al.<sup>63</sup> and Kielar, et al.<sup>67</sup> determined that it was necessary to adjust the measured physical DLG values to reduce dose calculation errors for their system. To compound this, measured DLG values can be different based on the measurement settings (e.g. field size, depth, and ion chamber).<sup>73</sup> This work can help physicists may better understand how each of these parameters generally contributes to IMRT plan accuracy and the interplay among major parameters in order to make informed decisions in beam model commissioning.

What is important to keep in context is that deviations in dose calculation accuracy are generally very difficult to detect using conventional QA methods. Studies in detectability, such as that of Koger, et al.,<sup>62</sup> McVicker, et al.,<sup>9</sup> Nelms, et al.,<sup>11</sup> and Kry, et al.<sup>10</sup> point to the gross lack of sensitivity for traditional IMRT QA methods in identifying beam modeling inaccuracies. Only external validation through an independent phantom, such as that from the IROC-H phantom program, can sufficiently capture the extent of delivery errors caused by poor dose calculations.

However, more work is needed to determine whether discrepancies in phantom dose distributions may be connected with atypical beam modeling parameter selection.

This work is limited in that it only focuses on one of the most common clinical systems: a Varian Clinic-type accelerator. It is possible that these results may differ from other clinical units with different physical configurations and geometries. Additionally, dose calculations were performed for the IROC-H H&N phantom, which is a simplified representation of human anatomy with little heterogeneity. Inherently, some dose effects may not have manifested through this choice of experimental setup. For instance, off-axis factor errors were likely not fully evident because the phantom is smaller than the dimension for which the greatest change in modeling was implemented (10 cm from central axis). It may be of value to investigate the potential effects of dose modeling variance on other phantom setups or geometries, where a variety treatment strategies may further highlight or unmask the true effects of modeling on clinical care.

It also is important to note that while this work intends to describe the range of dose calculation variations among the radiotherapy community, the results of this study do not imply that the use of atypical beam modeling parameters is totally inappropriate. This work is based upon average survey data and can only provide so much information regarding appropriateness for use. In fact, machine models that are intended for specific purposes (e.g. VMAT-dedicated units, radiosurgery units, etc.) may require use of parameter values that deviate from the norm.

## **5.5 Conclusions**

In this study, several beam modeling parameters encompassing both basic dosimetry data (i.e. PDDs, output factors, and off-axis factors) and other non-dosimetric modeling elements were assessed for their potential effects on IROC-H H&N phantom performance, based on the

radiotherapy community's range of measured and reported values. Of interest, the parameters related to the modeling of the MLC, specifically the DLG for Eclipse and the MLC Offset for RayStation, demonstrated substantial impact with regards to dose to the target, corroborating well with previous works. By applying the most extreme parameter values used clinically by the radiotherapy community, differences from the baseline, average-performance beam model produced clinically-compromising dose calculation errors. This result implies that these parameters can have a substantial clinical impact on the overall development and accuracy of IMRT plans and are of the utmost importance to commission correctly.

The quality and accuracy of the TPS radiation beam model is essential to providing high quality treatments. It is clear that, despite fundamental differences in TPS modeling formalisms, that parameters defining the MLC provide the greatest challenge to commission correctly. Understanding the ways in which these parameters influence the resultant dose calculation, both individually and collectively, can assist both IROC-H and the radiotherapy community at large to adopt the most appropriate values for modeling and improve IMRT performance.

## Chapter 6: Relationships between Beam Modeling and Phantom Performance

This chapter is based upon the following publication:

**M. Glenn**, F. Brooks, C. Peterson, R. Howell, D. Followill, J. Pollard-Larkin, S. Kry. “Photon beam modeling variations predict errors in IROC Houston phantom results.” To be submitted for publication, 2020.

Additional materials regarding this work can be found in Appendix E.

### 6.1 Introduction

The overall performance of radiation therapy is contingent upon the quality of the radiation beam modeling as this is the basis of treatment planning. Because modern treatment technologies like intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) require greater complexity in treatment delivery to achieve conformal dose distributions, it is critical that the underlying beam modeling be accurate and robust to a multitude of clinical scenarios.

The Imaging and Radiation Oncology Core Houston Quality Assurance Center (IROC Houston) offers end-to-end quality assurance through its anthropomorphic phantom program to evaluate that what is planned in the treatment planning system agrees with the dose that is delivered. In this program, an institution irradiates an IROC Houston phantom containing thermoluminescent dosimeters (TLDs) and radiochromic film, whose dose measurements are then compared against the institution’s calculated dose.<sup>5–8</sup> The irradiation either passes or fails to meet the acceptance requirements depending on the level of agreement between the measured and

calculated doses, required to be within 5-7% in absolute dose and with 85% of pixels passing 5%/3mm to 7%/4mm gamma criteria, depending on the phantom used.

Currently, 8-15% of recent phantom irradiations fail to meet these acceptance criteria.<sup>4,18,74</sup> Patterns have been identified among failing phantoms. For the head and neck (H&N) phantom, failures are mostly caused by systematic errors in dose; that is, the dose is administered in the correct location, but with the wrong magnitude.<sup>4</sup>

The cause of these dosimetric errors is of great interest and concern given the frequency and potential effects on patient care. Our recent evaluation identified that many poorly-performed irradiations are associated directly with errors in the TPS dose calculation.<sup>16</sup> However, substantial questions remain about what aspect of dose calculation are incorrect.

Recently, IROC Houston identified substantial variations in how radiotherapy institutions model their clinical beams for linear accelerators (linacs) of the same make, model, beam energy, and configuration, particularly with respect to parameters describing multileaf collimator (MLC) characteristics.<sup>65</sup> These vast disparities are even more suspect given that, dosimetrically, many of today's Linacs perform similarly dosimetrically.<sup>55,56</sup> Previous studies have underscored the sensitivity of certain parameters to variance, including the dosimetric leaf gap in Eclipse or MLC offset in RayStation.<sup>62,63,67,70,72</sup> Thus, there is interest in examining phantom performance in the context of an institution's beam modeling parameter choices.

The goal of this study was to investigate the potential relationships between an institution's beam modeling parameters and the accuracy of their phantom delivery. Understanding this link between phantom performance and specific choices of beam model parameter values can help to elucidate the cause of errors seen in IROC phantoms. Characterizing TPS-related errors on a multi-

institutional scale will lead to improvement in treatment accuracy and patient outcomes at radiotherapy facilities exhibiting those same TPS-related errors in dose calculation.

## **6.2 Methods**

We conducted a retrospective review of 337 phantom irradiations performed as part of IROC's regular activities from August 2017 to November 2019, along with concurrent reported TPS beam modeling parameters associated with each irradiation. Because of the contemporary nature of this data, the results of this work are up-to-date and reflect current treatment equipment, TPS model commissioning, and treatment strategies. Phantoms examined in this work were the H&N, prostate, and spine phantoms. Each of the phantoms contain unique challenges for planning and delivery, including multiple target volumes (for the H&N phantom) and avoidance structures for which IMRT or VMAT are necessary to treat appropriately. These phantoms are designed for static delivery and have previously been identified by IROC Houston as having major error modes related to systematically under- or overdosing the targets.<sup>4</sup>

### *6.2.1 Relationships between TPS beam modeling parameters and phantom irradiation performance*

We first examined if use of atypical beam modeling parameter values was associated with poor phantom results. Atypical beam modeling parameter values were assessed by comparison to the distribution of community values based on an IROC Houston administered TPS beam modeling parameter survey.<sup>65</sup> Table 6-1 describes some of the parameters collected in this survey, as well as the range of dose effects each parameter exhibits, based on reported variations (see Chapter 5). Phantom irradiations examined herein were grouped based on the presence of atypical

parameter values (here defined as being either <10<sup>th</sup> percentile or >90<sup>th</sup> percentile compared to the community values for the same linac type, beam energy, and MLC make and model). All TPS parameters of interest were examined individually. These groups were compared using Fisher's exact test to determine if the proportions of institutions adopting atypical values were different among passing and failing phantom irradiations (those having dose errors greater than 7%). Likewise, this analysis was performed for each parameter by comparing well-performed and poorly-performed irradiations; poorly performing phantoms may still be within tolerance, but have at least one TLD measurement that differs by more than 5% from the TPS calculated dose. Finally, we looked at relationships between individual parameter values and phantom performance accuracy.

**Table 6-1.** Treatment planning system beam modeling parameters requested via IROC Houston surveys and their range of dose effects, as previously determined by phantom dose recalculations using a common linac model (see Chapter 5).

TPS Parameter	Estimated Dose Effects	
	2.5 <sup>th</sup> Percentile	97.5 <sup>th</sup> Percentile
<b>Eclipse (AAA and AcurosXB)</b>		
Effective Target Spot Size X and Y [mm]	0%	0%
MLC Transmission Factor	-1%	+1%
Dosimetric Leaf Gap [cm]	-3%	+3%
<b>RayStation</b>		
Primary Source X Width and Y Width [cm]	0%	0%
MLC Transmission	-4%	+2%
Tongue and Groove [cm]	+1%	-1%
Leaf Tip Width [cm]	-2%	+2%
MLC Position Offset [cm]	-4%	+7%
MLC Position Gain	0%	0%
MLC Position Curvature [1/cm]	0%	0%

### 6.2.2 Impact of beam modeling variations on dose distributions

We evaluated the impact of atypical beam modeling parameter values on the dose distribution in the IROC phantom. These results could then be qualitatively compared to actual errors observed in the community. To this end, we designed beam models in Eclipse (Varian Medical Systems, Palo Alto, CA) with the AAA algorithm and in RayStation (RaySearch Laboratories, Stockholm, Sweden) to represent a standard Varian Base class<sup>55</sup> accelerator (e.g. Trilogy, 2100iX, etc.) with average modeling characteristics.<sup>65</sup> We then calculated dose on clinically-acceptable IMRT head and neck phantom treatment plans after individually modifying the parameters in Table 6-1 according to their reported distributions<sup>65</sup> to understand composite



plan changes based on the variance of modeling choices as seen in the radiotherapy community. These dose distributions were compared with IROC Houston phantom irradiation results to similar cases (in which one or more parameters was atypical) to assess qualitative similarities.

### *6.2.3 Estimating TPS dose calculation errors based on community data*

We previously showed that as TPS parameter values in Eclipse and RayStation were varied, the dose distribution varied correspondingly in a linear manner with parameter value, and this dose variation was independent of plan type (VMAT or IMRT), as described in Chapter 5. Moreover, each parameter was independent and not influenced by the values used for other parameters within the TPS. A range of estimated dose contributions, based upon community reported data, is included in Table 6-1, using median parameter values as a baseline. Based on these characteristics, we predicted the expected dosimetric deviations caused by beam modeling variations for the IROC Houston phantom irradiations performed with a Varian Base class accelerator. For example, if the DLG value used by the institution was at the 90<sup>th</sup> percentile, we expect this to cause an overestimation (and thereby underdose) of 3%. If their MLC transmission was at the 80<sup>th</sup> percentile, we expect this to cause an overestimation of 1%. Because these are independent, we can estimate a TPS error of 4% in this phantom irradiation. This predicted TPS error was then compared with the actual phantom error that was measured for that institution. These results were evaluated using the Pearson correlation between estimated and true dose errors observed. In order to isolate cases for which dose errors were the main contributor to phantom performance, and thus best test whether dose error contributions could be estimated, we excluded 11 phantom cases that exhibited localization errors of greater than 3 mm.

### 6.3 Results

337 previous static phantom irradiations investigated in this work composed of records with both complete and incomplete survey responses. Table 6-2 summarizes the breakdown of phantom irradiations examined. The majority of irradiations were performed on Varian machines (86%) using Eclipse as the primary TPS (78%). Most irradiations (77%) were performed on the IROC head and neck phantom. 31 phantoms (9.2%) failed to meet credentialing criteria ( $\pm 7\%$  TLD dose and 7%/4 mm gamma criteria for film) and 57 (16.9%) exhibited poor performance (at least one measurement outside  $\pm 5\%$ ).

**Table 6-2.** Summary of phantom irradiations.

Category	N	(%)
<b>TPS</b>		
Eclipse (AAA)	226	67.1%
Eclipse (AcurosXB)	38	11.3%
Pinnacle	40	11.9%
RayStation	33	9.8%
<b>Linac Types</b>		
Varian Base Class (Clinac series)	116	34.4%
Varian TrueBeam	165	49.0%
Elekta Agility (VersaHD, etc.)	40	11.9%
Other	16	4.7%
<b>Beam Energy</b>		
6 MV	287	85.2%
6 FFF	22	6.5%
10 MV	21	6.2%
10 FFF	6	1.8%
15 MV	1	0.3%
<b>Phantom</b>		
Head and neck	258	76.6%
Spine	34	10.1%
Prostate	45	13.4%

### 6.3.1 Relationship between TPS parameters and phantom performance

Atypical parameter values were identified in 19 failing cases (61.3% of all failures) and 30 poorly performed irradiations (52.6% of all poor cases). In contrast, atypical parameter values were present in 113 (41.7%) of well-performed irradiations. These proportions were found statistically different, with atypical parameters consistently present more often in failing irradiations (Fisher's Exact,  $p=0.01$ ). Additionally, when examining only parameters that caused potential changes in dose greater than 1% (Table 6-1), 51.6% of failing irradiations and 42.1% of poorly performed cases reported atypical parameters, while 31.7% of well-performed irradiations reported atypical values. Impactful atypical values were more associated with failing irradiations (Fisher's Exact,  $p=0.008$ ).

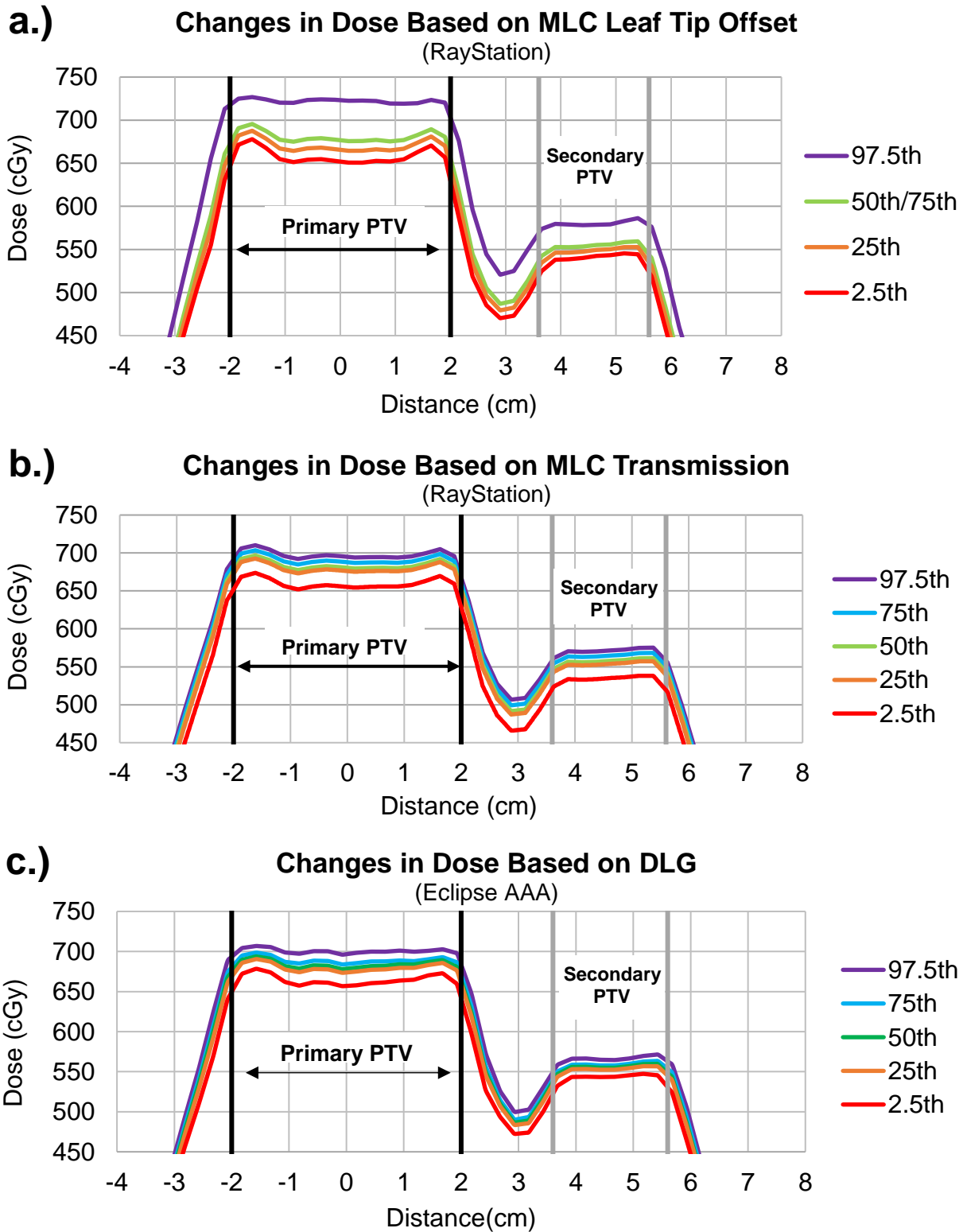
When evaluating specific TPS parameter values, atypical values of dosimetric leaf gap (DLG) in Eclipse AAA were related to poor performance ( $p=0.048$ , Fisher's Exact test) and failing irradiations ( $p=0.014$ ), consistently occurring more often with each negative outcome than those using more typical DLG values. A fair Pearson's correlation coefficient was identified between Eclipse DLG percentile score and average TLD dose calculation error for all applicable phantoms examined ( $r=0.293$ ,  $p<0.001$ ); that is, using a larger DLG value, as compared to the community consensus, was associated with overestimating the dose to the phantom. Interestingly enough, for models composed in Eclipse AAA, values used for DLG and MLC transmission factor were positively correlated ( $r=0.615$ ,  $p<0.001$ ), meaning users assigning a higher than typical DLG were more likely to use a greater value for the MLC transmission factor as well.

The only other relationship observed among specific parameters was that the primary source X width in RayStation was related to poorly performed ( $p=0.007$ ) and failing irradiations ( $p=0.042$ ). However, this result was based on dramatically fewer cases than observed with Eclipse users, and is inconsistent with our previous findings (see Chapter 5) that showed this parameter does not substantially affect the dose calculation (Table 6-1); consequently, this relationship requires further investigation.

### *6.3.2 Impact of beam modeling parameters on dose distribution*

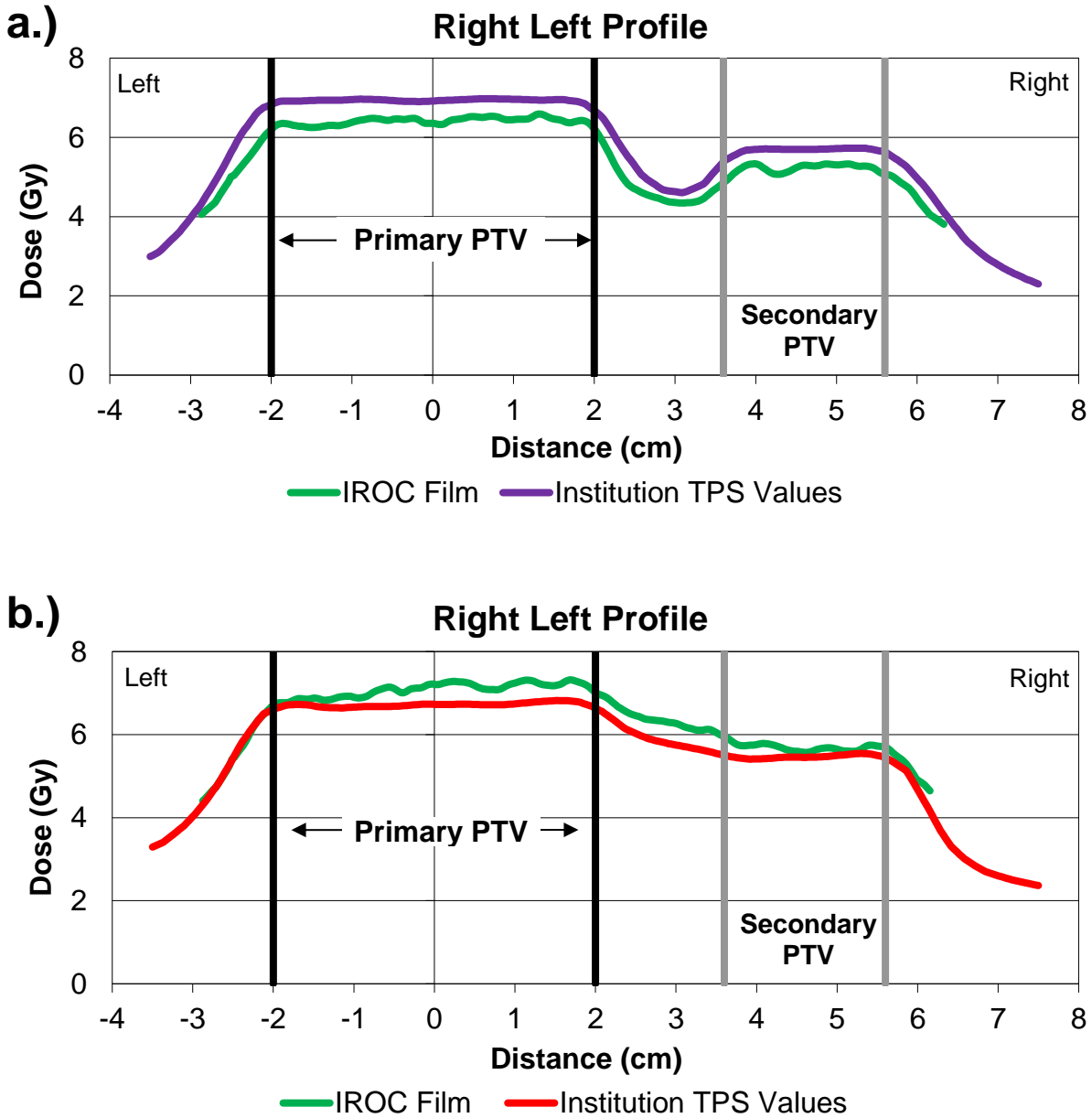
Recalculating clinically acceptable H&N phantom treatment plans with different values of beam modeling parameters generated systematic changes in dose for the range of parameter values examined (Table 6-1). Figure 6-1 visualizes such dose changes for some parameters of interest for a 9-field IMRT plan, showing the same dose distribution but with a different dose. This pattern remained true for all parameters tested. Like the ranges of potential dose contributions in Table

6-1, changing parameters representing the MLC leaf-tip offset (i.e. dosimetric leaf gap in Eclipse and the MLC position offset in RayStation) produced the greatest magnitude of dose changes, often surpassing 5% for atypical values used in the community. Parameters representing the source size, MLC gain, and MLC curvature produced no changes in the resultant dose distributions, consistent with previous observations shown in Chapter 5.



**Figure 6-1.** Average changes in plan dose calculated to the primary and secondary PTVs following manipulation of (a) RayStation MLC leaf tip offset, (b) RayStation MLC transmission factor, and (c) Eclipse dosimetric leaf gap (DLG) for an IMRT H&N phantom plan based on 2.5<sup>th</sup> to 97.5<sup>th</sup> percentiles of beam modeling survey results.

These systematic differences in dose calculations shown in Figure 6-1 qualitatively match the most common form of phantom errors that are observed: systematic dose errors.<sup>4</sup> Moreover, evaluation of previous phantom irradiations revealed cases that display these systematic dose deviations (Figure 6-1) when atypical values were used. Figure 6-2 shows right-left film plane measurements of two previous H&N phantom irradiations compared to TPS-calculated dose profiles for the same regions. In Fig. 6-2.a, the institution's beam model used an atypical value for DLG (91st percentile) and systematically overestimated the dose by approximately 7%. To the opposite effect, Fig. 6-2.b depicts dose profiles for an institution whose beam model used extremely low values for DLG and MLC transmission (1st and 2.5th percentile, respectively) that subsequently underestimated the delivered dose by 5% on average. An additional detail is that these two irradiations were accompanied by documentation stating that the treatment plans passed institutional IMRT QA standards, reporting at least 96.3% and 99.1% pixels passing, respectively, for all treatment fields. For these institutions, IMRT QA was insufficient in identifying problematic plans that showed systematic and understood dose errors. This is consistent with previous general IROC Houston observations.<sup>10</sup>

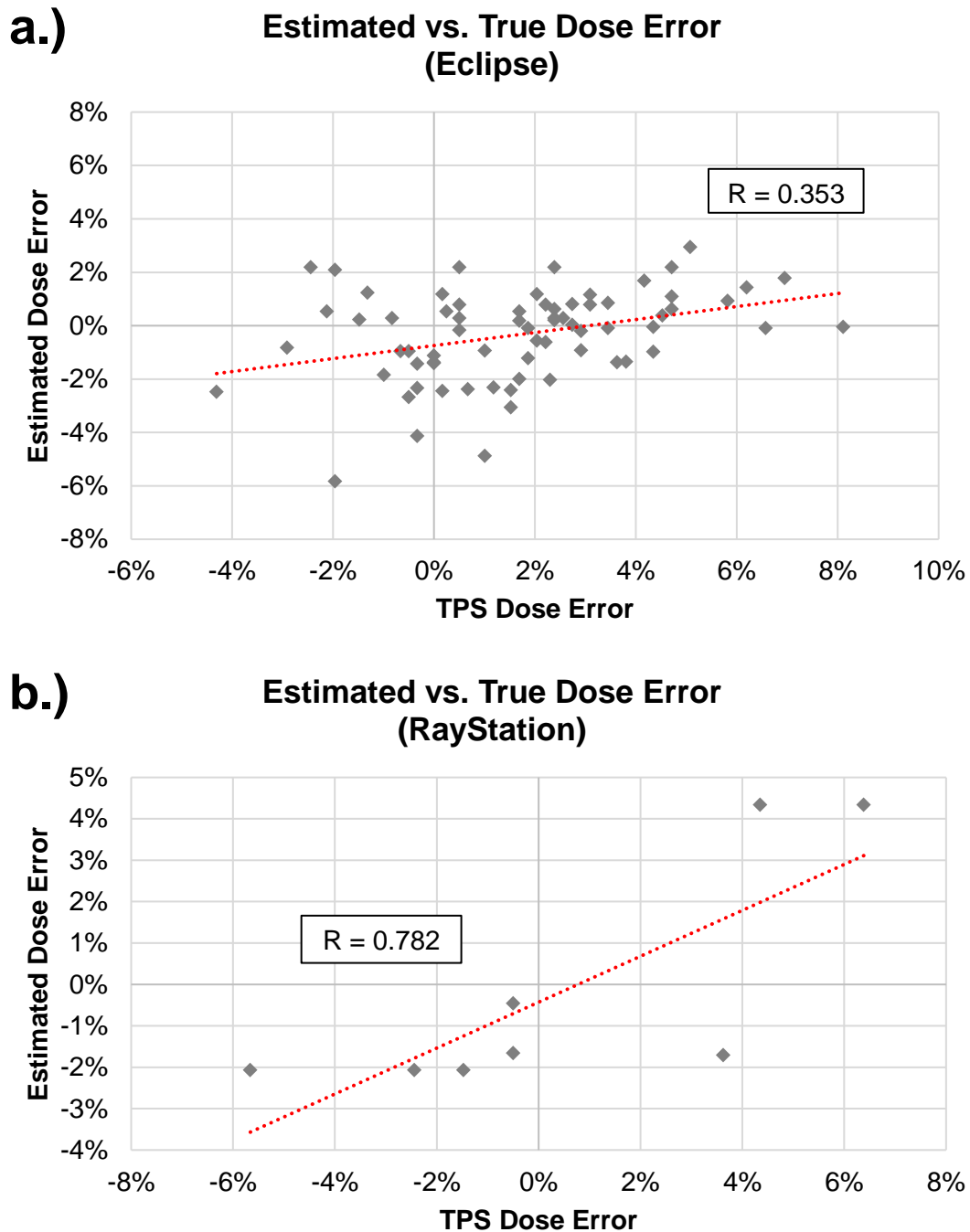


**Figure 6-2.** Phantom cases of interest. (a) Irradiation using Eclipse AAA with high dosimetric leaf gap (91<sup>st</sup> percentile) that overestimated the dose delivered (purple) compared to film measurement (green) by ~7%. The institution overestimated the dose on a second attempt three months later using the same beam model. (b) Irradiation using Eclipse AAA with very low DLG (1<sup>st</sup> percentile) and MLC transmission (2.5<sup>th</sup> percentile), that underestimated the dose delivered (red) compared to film measurement (green) by 5% on average.



### 6.3.3 *Estimating dose error*

Phantom irradiations performed with a Varian Base class Linac and having no identifiable localization issues were selected to test the potential for estimating dosimetric errors caused by beam modeling parameter variations. Total estimated dose errors, defined as the sum of individual dose error contributions from each investigated parameter, were plotted against phantom dose inaccuracy for both Eclipse ( $r=0.353$ ,  $p=0.003$ ) and RayStation ( $r=0.782$ ,  $p=0.022$ ). It is clear from Figure 6-3 that the predicted dose error does not capture the entire difference between measurement and TPS calculation (as the correlation coefficients are not approximately equal to 1); this result is unsurprising given that there exist other potential error modes that are not accounted for in this estimation. However, these estimations typically accounted for a large portion of observable error (for example, estimating +4% error when the true TLD dose error was +6%). This indicates that dose calculation errors are identifiable and contribute substantially to errors occurring in static IROC phantom irradiations.



**Figure 6-3.** Comparison of estimated versus true dosimetric error for phantom irradiations performed using a Varian Base class Linac<sup>55</sup> for (a) Eclipse and (b) RayStation. Phantom irradiations with setup errors greater than 3 mm were excluded.

## 6.4 Discussion

Our study highlights that TPS modeling parameters are involved in failing IROC phantom irradiations. First, there exist clear relationships between beam modeling parameter value and phantom outcome. The DLG was an important parameter in the dose calculation; while this finding has been known,<sup>63,67,70</sup> in this work we show that atypical modeling of this parameter is directly associated with poor performance on IROC phantom. The effects of the DLG were observable despite many other potential compounding issues that could result in a poor dose calculation, such as Hounsfield unit-to-density curve errors, input dosimetric beam data errors, machine delivery errors, among others.

Second, when manipulating beam models, this study notes the variation of any single parameter produced systematic changes in dose across all regions in the phantom. These results directly parallel the observations of Carson, et al.<sup>4</sup> and Edward, et al.,<sup>74</sup> which identified systematic dose discrepancies as the predominant cause of poor phantom performance among static phantoms. The examples in Figure 6-2 demonstrated the same pattern.

An unsettling tangential observation also arose from this work: all 337 plans observed in this cohort reported passing IMRT QA. For the interesting cases shown in Figure 6-2, which showed systematic dose differences of 7% and 5% respectively, the institutions reported 96% or more pixels passing using 3%/3 mm absolute gamma analysis for all fields evaluated, as measured by portal dosimetry. These results substantiate other efforts that have shown that IMRT QA is insensitive at identifying unacceptable plans.<sup>10,11</sup> Likewise, beam modeling inaccuracies are difficult to detect using conventional IMRT QA methods,<sup>9-12,25,44</sup> but can have clinical consequences on dose accuracy.<sup>9,62</sup>

Limiting factors of this work are the incompleteness of data reported by institutions, limiting our analyses to one common linac class (Varian Base), and the narrow scope of our dose error estimations. While most provided all parameters requested in the IROC Houston beam modeling survey, several did not. Without all data provided, a number of dose error estimations we drew for a subset of phantom irradiations are not entirely representative of the full potential for TPS-related dose errors. Despite this, we were still able to capture plenty of information in regards to phantom performance, as the correlations of Figure 6-3 demonstrate. This analysis also demonstrates that, for the most part, using parameter values that are not extreme can lead to more accurate representation of linac performance. However, this work is provided only in the context of a single linac class, so dose error estimations of other popular linacs could not be assessed or interpreted. Finally, these dose error estimations only consider beam modeling as a contributing factor to phantom performance, which cannot fully represent all errors observed in static phantoms. With further characterization of other error modes, these estimations could be improved.

Though we identified the trend that atypical parameter values were associated with poorer phantom performance, the data do not advocate for using a universal standard beam model. For those failing irradiations that reported using atypical parameter values, especially for the DLG or MLC offset, it is extremely likely that changing these parameters to more typical values could have resulted in passing the phantom. However, there were also cases observed where an atypical TPS parameter value resulted in a highly accurate delivered dose. Because our dose error estimations were based on a reference linac model, such well-performing models describe non-reference machines, so this data is less applicable. In short, while the median value is likely to be a good value, it is essential to use values that are appropriate for the system and its intended use. Values implemented into a beam model need to be scrutinized and tested for robustness under a

variety of clinical scenarios before determining their suitability. Given the performance of current IMRT QA tools, and consistent with AAPM MPPG5,<sup>49</sup> this should include independent evaluation.

## **6.5 Conclusions**

This study examined the relationships between beam modeling parameter values and phantom irradiation performance. Notable correlations were identified between atypical DLG values implemented in Eclipse and overall phantom performance. In general, atypical TPS parameter values were directly correlated with actual delivery errors in the IROC phantoms, but this could not describe all cases.

Specific beam model parameters, especially those that represent the MLC behavior and characteristics, were found to be substantially involved in failing IROC phantom results. These results provides direct guidance to physicists who receive suboptimal results on a phantom in terms of how to improve the quality of their radiotherapy. As such, this has the potential to improve the quality of radiotherapy and substantially reduce the frequency of failing IROC phantoms in the community.

## Chapter 7: Discussion & Conclusions

### 7.1 General Summary and Conclusions

The main purpose of this work was to better characterize an identified subset of phantom performance errors, TPS beam modeling errors, and understand their prevalence and manifestation in the IROC Houston phantom credentialing program. This study examined multiple factors as they relate to phantom plan performance in order to better characterize treatment planning system modeling variations. The first aim was designed to understand observational trends in phantom performance as a whole to determine how treatment errors manifest in a common IMRT phantom, as well as determine if the severity of the dose error was related to the complexity of the treatment. The second aim was developed to describe how the radiotherapy community develops its clinical beam models (i.e. whether consensus exists among similar machine configurations) and how variation in beam model parameter selection may impact overall dose effects in a common phantom analog. Finally, the third aim utilized information found from the second aim to see whether relationships exist between choice of beam modeling parameters and overall phantom performance, as well as to illustrate the comprehensive effects of beam modeling variations on IMRT treatment plans.

#### *7.1.1 Specific Aim 1: Identify common observational characteristics of poor phantom performance*

The purpose of the first aim was to best understand how errors in phantom irradiations manifest and determine whether there were conditions for which such errors are more likely to manifest. Our working hypotheses for this aim were that dosimetric characteristics would be a

predominant factor in phantom performance, and complexity would be correlated with the degree of dosimetric error exhibited among phantom irradiations.

For the first objective of this aim, we performed a retrospective analysis of phantom irradiation results to understand common observational patterns of failure among head and neck phantom irradiations, including localization, systematic dose, as well as global and local dose errors. While phantom errors manifest in multiple ways, most often, these errors appeared systematic (nearly 70% of failures observed); that is, the dose that was delivered had the right distribution and location, but was simply the incorrect magnitude when compared to the planned dose. This proved the first part of our working hypothesis true. In addition to understanding observational failure modes, we also determined that it was technically feasible to tighten IROC Houston criteria from  $\pm 7\%/4$  mm to  $5\%/4$  mm. This new criteria could help identify errors at reasonable, clinical action level while not decreasing the effective pass rate for IROC phantoms too greatly.

For the second objective of this aim, we evaluated complexity as a factor of phantom performance using 16 different metrics that describe different aspects of machine motion and treatment apertures defined by the MLC. Complexity was universally not associated with plan performance, thus proving the second part of our working hypothesis false. However, this also indicates that a robust and well-tuned beam model should perform well (that is, calculate dose with little uncertainty) under an extensive range of clinical scenarios. Additionally, we determined that nearly all metrics were correlated with one another, meaning that they provide very similar information in terms of how complex an IMRT or VMAT plan is. This understanding could help simplify future works in complexity by evaluating more universal measures, such as the monitor units (MU) of a treatment plan.

### *7.1.2 Specific Aim 2: Determine critical beam modeling parameters for accurate dose calculation*

The purpose of Specific Aim 2 was to characterize how the radiotherapy community develops its clinical beam models (i.e. how it assigns specific parameter values for similar Linacs) and determine the realistic extent of dose calculation variations, based upon the community-reported data. Our working hypothesis was that modeling parameters associated with the multileaf collimators (e.g. dosimetric leaf gap) will be the most sensitive to changes and thereby present the greatest potential for changes in resultant dose calculations.

For the first objective of this aim, we instituted a TPS beam modeling parameter survey to coincide with the IROC Houston phantom credentialing program. Institutions submitted their TPS data alongside phantom irradiations either electronically or via hand-written documentation. From this data we developed a reference dataset that may be used by the radiotherapy community at large as a second check in the TPS commissioning and QA process. We also determined that for some parameters, especially those associated with the MLC, extensive variations exist in the radiotherapy community for institutions that treat with the same models of Linacs.

For the second objective, we designed five clinically acceptable head and neck phantom plans, created baseline, average beam models in both RayStation and Eclipse, and performed a sensitivity analysis on the beam modeling parameters of interest, as defined from the initial TPS beam model parameter survey. We evaluated the individual parameter effects independently, as well as the potential for interplay among parameters, by manipulating numeric values from the 2.5<sup>th</sup> to the 97.5<sup>th</sup> percentile. This was also performed for dosimetric characteristics, including jaw-defined small field output factors, PDDs, and off-axis factors. Those parameters related to the modeling of the MLC, especially the DLG in Eclipse and the MLC offset in RayStation, were the



most susceptible to variations, as well as indicated the greatest changes in dose calculations, thus rendering our working hypothesis for this aim true. Dosimetric characteristics, while previously having been known to be issues, were not well evidenced in this work, suggesting that different models may be more applicable for observing these phenomena.

### *7.1.3 Specific Aim 3: Characterize relationships between beam modeling choices and phantom performance*

The purpose of Specific Aim 3 was to corroborate the results of Specific Aim 2 with previous phantom irradiation results to create a more complete narrative of TPS errors in IROC phantom performance. Using 337 phantom irradiations performed from August 2017 through November 2019, we quantified the relationships between beam model parameter selection and overall phantom performance and evaluated the hypothesis of the work (see Section 7.2). Moreover, we determined if relationships existed among parameters (i.e. whether using one value of a parameter made it more likely for users to choose a given value for a different parameter). We also determined how changing parameters (much like what was done in Specific Aim 2) affects the entire dose distribution, not just singular points like the TLD, and attempted to estimate the potential contributions to dosimetric errors based on dose changes observed from our sensitivity analyses.

In this work we determined that the DLG, a parameter previously found important in dose calculation accuracy, was correlated with TPS accuracy. The estimated error due to beam modeling variations proved informative as well; for the subset of irradiations tested, correlations between estimated and true dose errors were found despite other confounding factors that could contribute

to dose delivery inaccuracies. Together, these points further emphasize the overwhelming effects of beam modeling deficiencies on plan performance.

## **7.2 Evaluation of the Hypothesis**

The hypothesis for this project was that inaccurate TPS commissioning (i.e. beam model parameter selection) constitutes the primary phenomena related to phantom dosimetric errors, with 50% of irradiations identified as having large dose disagreements reporting atypical values in one or more of the investigated parameters. Using the results from Specific Aim 3, we did in fact find conclusive relationships between use of atypical beam modeling parameters and poor phantom performance (see Section 6.3.1). We observed that 57 (16.9%) phantom irradiations had dosimetric errors of greater than 5%, our established criterion for poor performance. Of those with poor performance, 30 (52.6%) phantom irradiations had concurrent beam modeling surveys reporting atypical parameter values, with 24 (42.1%) having an atypical parameter that were previously identified as important in dose calculations (see Chapter 5). For the failing cases (having >7% errors), 19 (61.3%) demonstrated atypical parameters, of which 16 (51.6%) had atypical parameters that were previously identified as important for dose calculation accuracy. Given these results, we can report that our hypothesis was met.

Atypical parameter values were not universally predictive of a dose error, likely due to the multiple modes of error that can occur in the treatment process, limited statistics for certain Linac representations, as well as limited parameter reporting. There exist many confounding factors that are difficult to decouple, thus introducing non-negligible noise into our analyses (e.g. Was the phantom setup accurately? Was the Linac calibrated appropriately?). Despite these limitations, there were still multiple cases for which this work highlights how detrimental erroneous modeling

can be. Statistically, we determined that atypical values of the dosimetric leaf gap in Eclipse (those defined as either <10<sup>th</sup> percentile or >90<sup>th</sup> percentile based on community reported data) occurred more frequently in both poor-performing and failing phantom irradiations. No other parameters had either a great enough effect size or enough samples to identify statistical significance. Even so, we were able to also correlate expected and true error contributions from parameter deviations.

While the relationships we have found are not all-encompassing, we have focused only on RayStation and Eclipse, and excluded other TPSs. Nevertheless, what we have identified is applicable to over 80% of the radiotherapy community, as most currently use Eclipse or RayStation as their primary TPS. Therefore, this work has extensive value for those who commission and maintain TPS photon beam models.

### **7.3 Future Research and Applications**

There exist several avenues through which this study can be continued in future work. First, the sensitivity of beam modeling parameters were only investigated for one current and popular Linac configuration, a Varian Base class accelerator with standard 120-leaf MLC, and two common TPS: Eclipse and RayStation. Given the continued interest in stereotactic treatments and newer accelerators, like the Varian TrueBeam, it may be of particular interest to understand the modeling variance and associated changes in dose calculation accuracy of these newer systems, both with and without high definition MLC. Additionally, Elekta units with Agility 160-leaf heads may be of potential interest characterize for a more well-rounded understanding of where beam modeling errors most often occur. While less widely used, expanding this work to encompass the variations exhibited in Pinnacle may better help address the current variations and understand how the TPS parameters function.

Second, this work may be better informed through the testing of patient plans. One major limitation of the work presented here is that it is relevant to static IROC phantoms, which are standardized, mostly homogeneous, and simplistic representations of human anatomy. By adapting this work for patient plans, one may be able to use vastly different geometries to simulate and identify error modes that cannot be effectively evaluated in this work (namely off-axis errors, small jaw-defined fields, and PDD errors). It is also well-known that heterogeneities in human anatomy, particularly where there are interfaces between high- and low- density tissues, present extensive challenges for accurate dose calculation, thus it is possible that errors observed here may occur to either greater or lesser extents, depending on the treatment objectives and patient anatomy.

Finally, the information found herein may be implemented as a means for better diagnosing when unexpected things occur in the IROC phantom program. Continued collection of modeling parameters will refine community statistics and improve IROC Houston's ability to provide comprehensive feedback. With the improvement of these processes, IROC Houston can continue to ensure that institutions provide the utmost quality in their radiation therapy programs, and clinical trials using radiation therapy may minimize the uncertainty in their work.

## **7.4 Takeaway Messages**

Based on the works performed herein, it is clear that there are still avenues through which the radiotherapy community can improve radiation therapy performance, particularly with regard to beam model commissioning. Issues in beam modeling are commonplace, affect the whole plan systematically, and can have serious clinical implications should they persist. While there will always be uncertainties regarding radiation dose accuracy, the data provided by this work may be used for reference to better inform how future errors may be prevented. Through continued

surveillance and auditing (aided by this work), IROC Houston may provide enhanced recommendations through its programs to ensure radiotherapy clinics provide the highest quality care to their patients.

## Appendix A: Supplement to Chapter 2

This appendix serves as the supplement to  
Chapter 2: Sources of Error in IROC Phantom Irradiations

This appendix contains additional dose profiles and computational work used to support the conclusions of Chapter 2. A record of individual phantom gamma index recalculations, other calculation spreadsheets, and examples of phantom irradiations with obvious errors may be found on the IROC Houston network drive: J:\Everyone\Mallory\Dissertation\Appendix A.

The probability of failing the phantom test based on TLD uncertainty was assessed using various standard deviations in TLD performance reported by the community.<sup>22,75,76</sup> As shown in

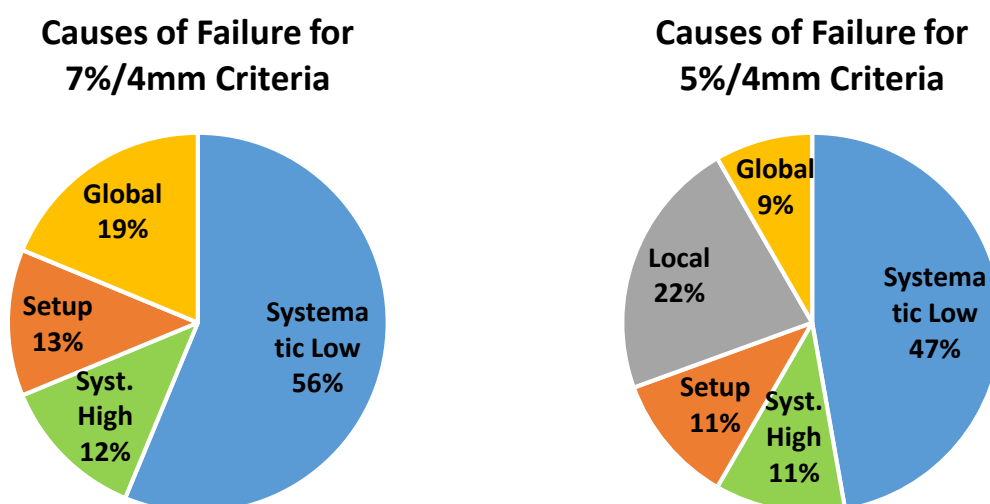
Table A-1, for a well-performed irradiation, the uncertainty associated with any individual TLD measurement is minimal and should not contribute greatly given current IROC criteria for acceptance. However, as the criteria are tightened, there exist greater possibilities for noise-induced phantom failures. This work supports the idea that phantom criteria can be tightened to 5% with minimal complications due to TLD uncertainty, and may improve the identification of sub-par irradiations.

**Table A-1.** Probabilities of phantom failure based on different reported TLD uncertainties and action criteria for a well-performed irradiation.

Source	St. Dev.	7% Criteria		5% Criteria		3% Criteria	
		P(1 fail)	P(x>=1)	P(1 fail)	P(x>=1)	P(1 fail)	P(x>=1)
Average st. dev. for all TLD	0.011	9E-12	5E-11	6E-07	3E-06	0.0015	0.0087
Kirby, et al. 1992	0.023	0.0011	0.0066	0.0168	0.0966	0.1281	0.5606
Double-loaded TLD	0.016	4E-06	2E-05	0.0007	0.0043	0.0314	0.1742
Incl. energy variance	0.021	0.0004	0.0021	0.0088	0.0518	0.0956	0.4527
Izewkska, et al. 2008	0.016	3E-06	2E-05	0.0006	0.0035	0.0287	0.1603

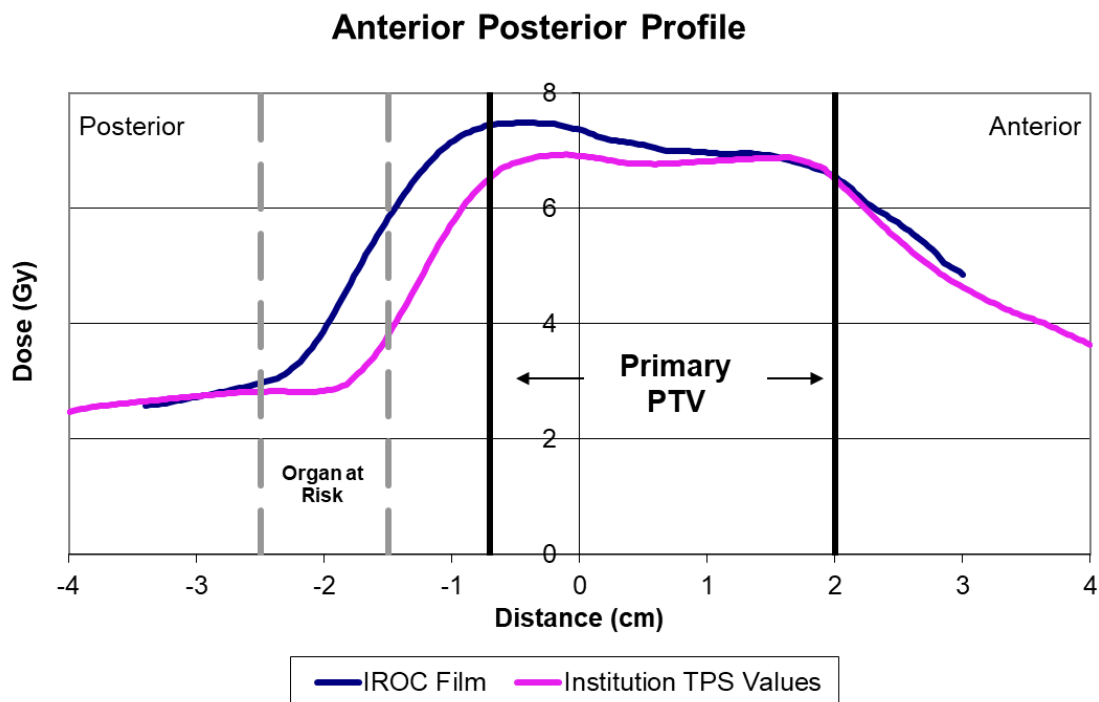
Jangda, et al. 2012	0.025	0.0027	0.0161	0.0278	0.1557	0.1615	0.6525
---------------------	-------	--------	--------	--------	--------	--------	--------

Figure A-1 depicts the proportions of common characteristics phantom failure for two different criteria, the current IROC standard of 7%/4 mm and the proposed criteria 5%/4 mm. In general, the proportions of failing phantoms are overwhelmingly systematic. Here it is also interesting to note that, unlike the current criteria, the 5%/4mm criteria could distinguish phantoms that failed due to localized causes (inaccurate dose in one region in the phantom).

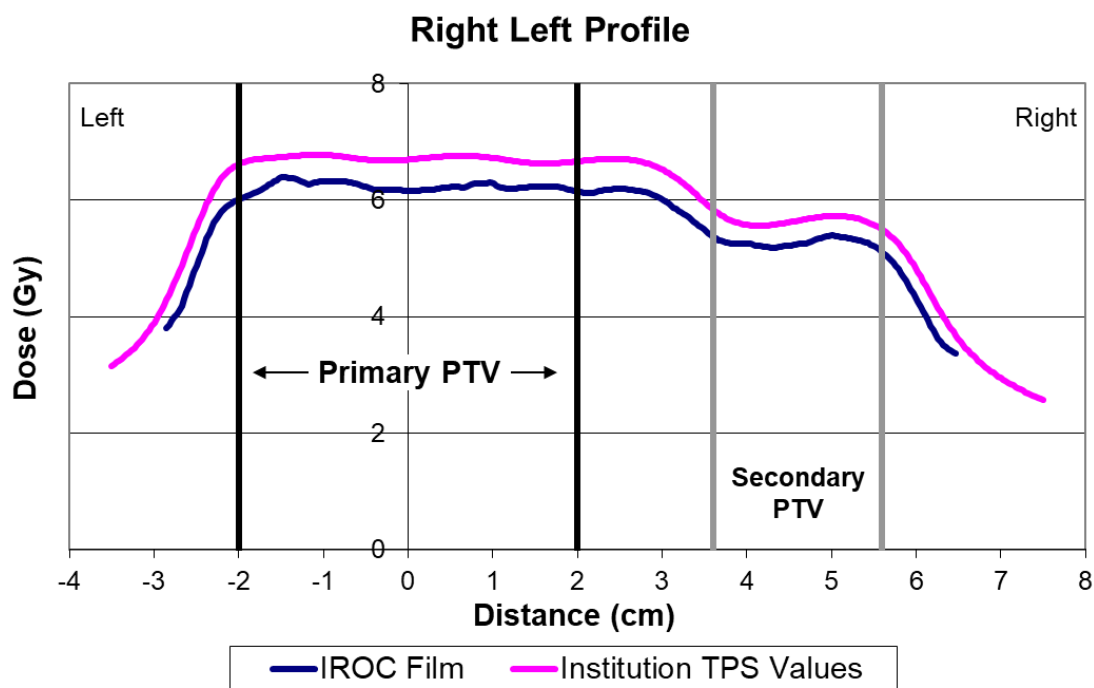


**Figure A-1.** Proportions of common characteristics for failing phantoms at 7%/4 mm (left) and 5%/4 mm criteria (right).

Figures A-2 through A-5 depict examples of common error modes described in Chapter 2.

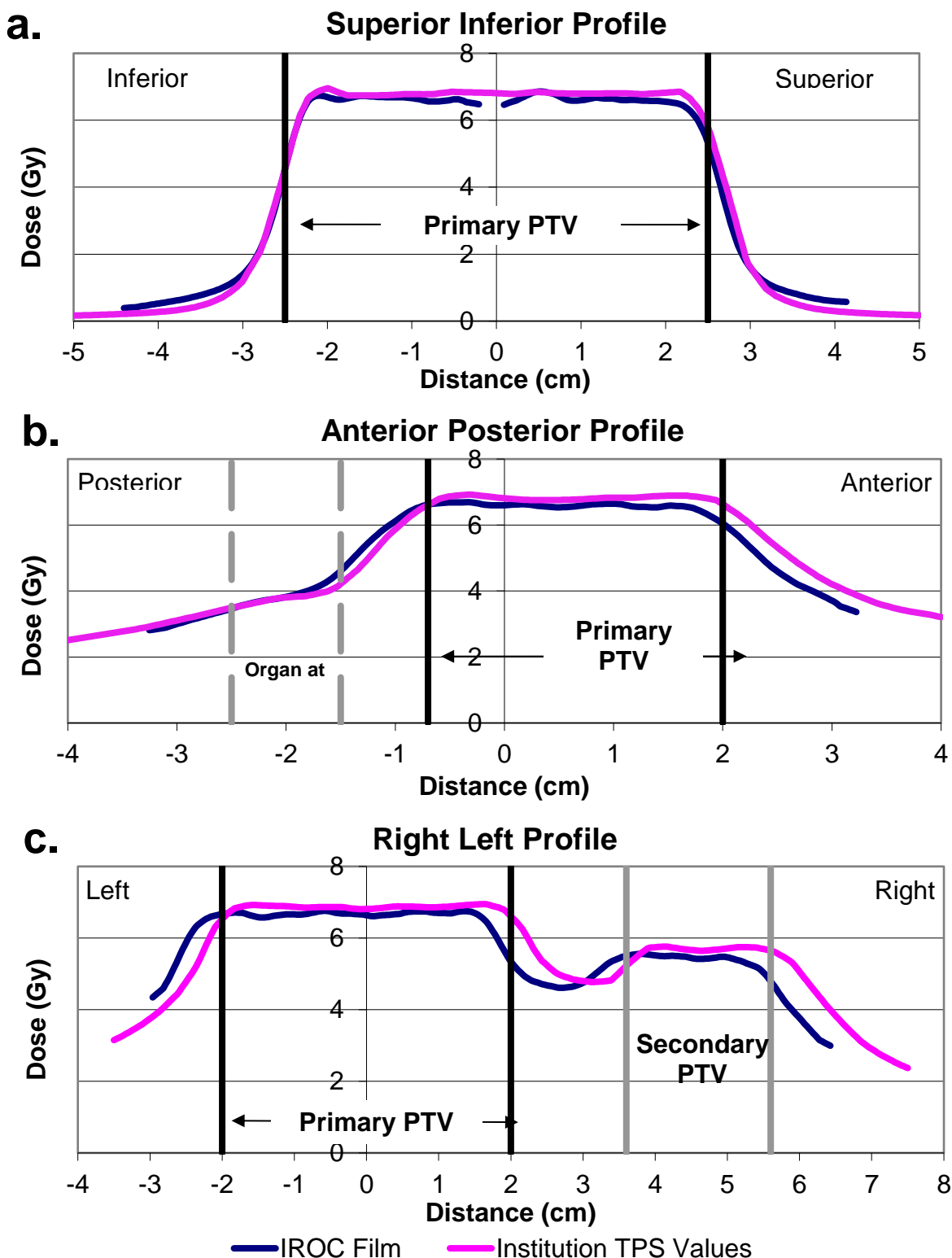


**Figure A-2.** Superior-inferior profile of a phantom irradiation exhibiting a global error (see Figure 2-2 for other film planes).

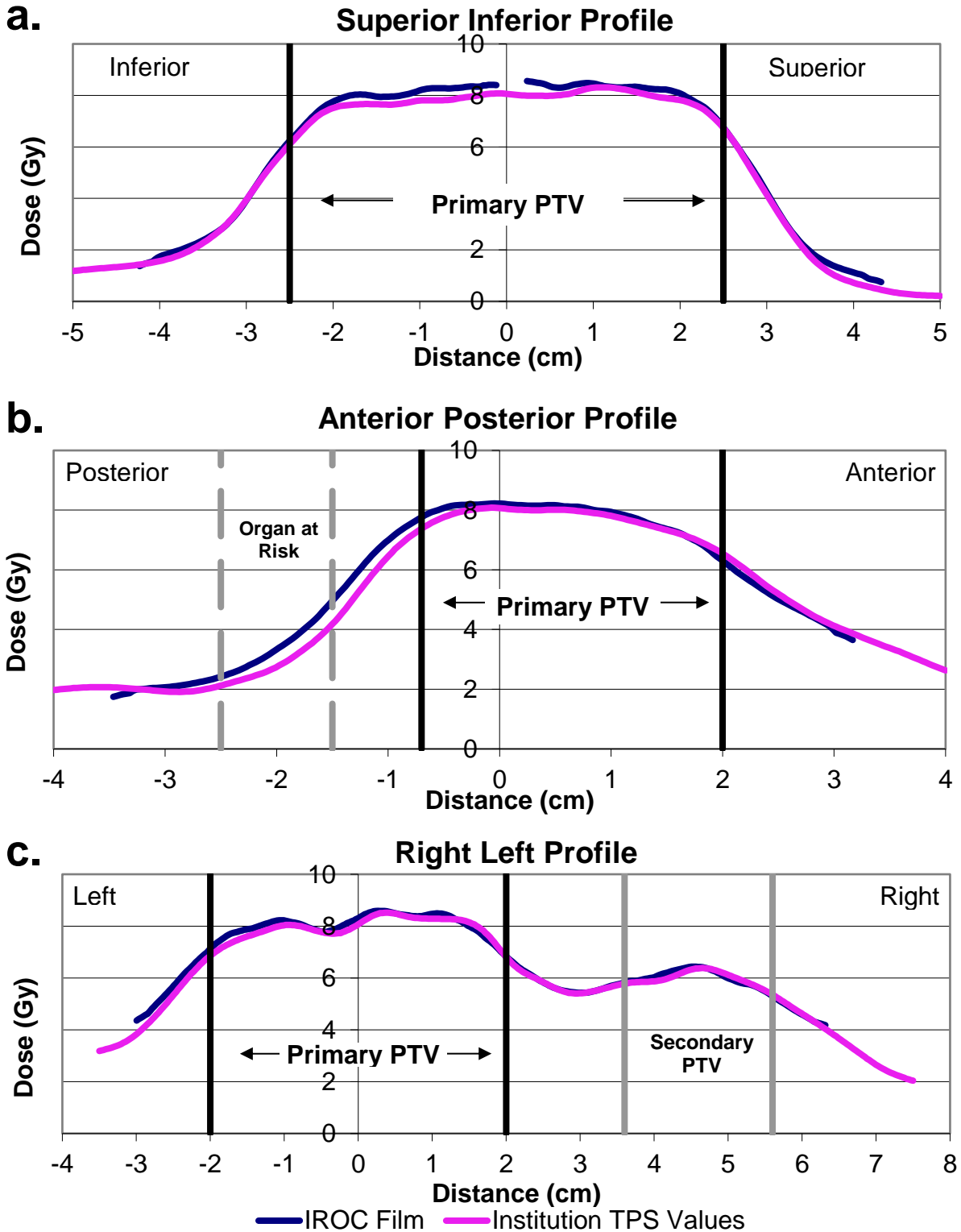


**Figure A-3.** Right-left profile of a phantom irradiation exhibiting a systematic error (see Figure 2-3 for other film planes).





**Figure A-4.** Sample setup/localization error. These plots of (a) a superior–inferior profile taken from a sagittal film, (b) an anterior–posterior profile and (c) a right–left profile from an axial film are from an institution trial exhibiting a localization error, as there is a right–left shift of 4 mm between the planned and delivered doses.



**Figure A-5.** Sample local error. These plots of (a) a superior–inferior profile taken from a sagittal film, (b) an anterior–posterior profile and (c) a right–left profile from an axial film are from an institution trial exhibiting a local error, as there is a distinct error in the region between the primary PTV organ at risk (see 5.b) which is not seen elsewhere in the profiles.

## **Appendix B: Supplement to Chapter 3**

This appendix serves as the supplement to  
Chapter 3: Complexity as a Predictor of Phantom Performance

This appendix contains information related to Chapter 3, including reference for the PlanAnalyzer code provided by Victor Hernandez and Jordi Saez. Raw data and the PlanAnalyzer MATLAB code are archived on the IROC Houston network drive at:

J:\Everyone\Mallory\Dissertation\Appendix B.

The reference folder contains the ZIP file “PlanAnalyzer\_v6\_Mallory.zip” which contains the raw code with modules to assess different complexity metrics, as well as a reference document, “DOC\_ComplexityAnalyzer.doc,” prepared by Victor Hernandez, which contains program information and describes each of the complexity metrics in greater detail.

Table B-1 demonstrates the clear correlations between complexity metrics for all plans examined (see highlighted values with asterisks). That is, the information captured by these metrics is not totally unique. This pattern is upheld for subsets of information, including IMRT plans (Tables B-2 and B-3), VMAT (Table B-4), and the Eclipse and Pinnacle TPS (Tables B-5 and B-6). Note that not all metrics are calculated for all plans, given that some metrics are specific to VMAT delivery (specifically for continuous gantry motion). Figures B-1 through B-5 depict scatterplots of interest not shown in Chapter 3.

**Table B-1.** Correlations between complexity metrics for all plans examined.

		MU (Corrected)	MCS (McNiven 2010)	EdgeMetric (Younge 2012)	PI - Plan Irregularity (Du 2014)	PM - Plan Modulation (Du 2014)
<b>Plan Error (Absolute Average % Difference between TPS and TLD)</b>	Correlation Coefficient	-.025	.031	<b>-.119*</b>	-.082	-.018
	Sig. (2-tailed)	.645	.570	.027	.130	.742
	N	343	343	343	343	343
MU (Corrected)	Correlation Coefficient	1.000	.019	-.009	-.047	<b>.577**</b>
	Sig. (2-tailed)		.729	.863	.388	.000
	N	343	343	343	343	343
MCS (McNiven 2010)	Correlation Coefficient	.019	1.000	<b>-.625**</b>	<b>-.569**</b>	<b>-.275**</b>
	Sig. (2-tailed)	.729		.000	.000	.000
	N	343	343	343	343	343
EdgeMetric (Younge 2012)	Correlation Coefficient	-.009	-.625**	1.000	<b>.954**</b>	<b>.199**</b>
	Sig. (2-tailed)	.863	.000		.000	.000
	N	343	343	343	343	343
PI - Plan Irregularity (Du 2014)	Correlation Coefficient	-.047	-.569**	.954**	1.000	<b>.200**</b>
	Sig. (2-tailed)	.388	.000	.000		.000
	N	343	343	343	343	343
PM - Plan Modulation (Du 2014)	Correlation Coefficient	<b>.577**</b>	<b>-.275**</b>	.199**	.200**	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	
	N	343	343	343	343	343

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

**Table B-2.** Correlations between complexity metrics for dynamic IMRT plans.

		Corrected MU	MCS (McNiven 2010)	EdgeMetric (Younge 2012)	PI - Plan Irregularity (Du 2014)	PM - Plan Modulation (Du 2014)
<b>Plan Error (Absolute Average % Difference between TPS and TLD)</b>	Correlation Coefficient	<b>-.218*</b>	<b>.240*</b>	-.188	-.079	-.088
	Sig. (2-tailed)	.036	.021	.071	.450	.402
	N	93	93	93	93	93
Corrected MU	Correlation Coefficient	1.000	<b>-.411**</b>	<b>.451**</b>	<b>.585**</b>	<b>.667**</b>
	Sig. (2-tailed)		.000	.000	.000	.000
	N	93	93	93	93	93
MCS (McNiven 2010)	Correlation Coefficient	<b>-.411**</b>	1.000	<b>-.766**</b>	<b>-.677**</b>	<b>-.512**</b>
	Sig. (2-tailed)	.000		.000	.000	.000
	N	93	93	93	93	93
EdgeMetric (Younge 2012)	Correlation Coefficient	<b>.451**</b>	<b>-.766**</b>	1.000	<b>.810**</b>	<b>.553**</b>
	Sig. (2-tailed)	.000	.000		.000	.000
	N	93	93	93	93	93
PI - Plan Irregularity (Du 2014)	Correlation Coefficient	<b>.585**</b>	<b>-.677**</b>	<b>.810**</b>	1.000	<b>.646**</b>
	Sig. (2-tailed)	.000	.000	.000		.000
	N	93	93	93	93	93
PM - Plan Modulation (Du 2014)	Correlation Coefficient	<b>.667**</b>	<b>-.512**</b>	<b>.553**</b>	<b>.646**</b>	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	
	N	93	93	93	93	93

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

**Table B-3.** Correlations between complexity metrics for static IMRT plans.

		Corrected MU	MCS (McNiven 2010)	EdgeMetric (Younge 2012)	PI - Plan Irregularity (Du 2014)	PM - Plan Modulation (Du 2014)
<b>Plan Error (Absolute Average % Difference between TPS and TLD)</b>	Correlation Coefficient	.212	.057	-.003	.119	.097
	Sig. (2-tailed)	.172	.719	.987	.446	.535
	N	43	43	43	43	43
Corrected MU	Correlation Coefficient	1.000	.025	.094	<b>.397**</b>	<b>.500**</b>
	Sig. (2-tailed)		.872	.548	.008	.001
	N	43	43	43	43	43
MCS (McNiven 2010)	Correlation Coefficient	.025	1.000	<b>-.470**</b>	<b>-.384*</b>	.012
	Sig. (2-tailed)	.872		.001	.011	.937
	N	43	43	43	43	43
EdgeMetric (Younge 2012)	Correlation Coefficient	.094	<b>-.470**</b>	1.000	<b>.694**</b>	.034
	Sig. (2-tailed)	.548	.001		.000	.831
	N	43	43	43	43	43
PI - Plan Irregularity (Du 2014)	Correlation Coefficient	<b>.397**</b>	<b>-.384*</b>	<b>.694**</b>	1.000	.243
	Sig. (2-tailed)	.008	.011	.000		.117
	N	43	43	43	43	43
PM - Plan Modulation (Du 2014)	Correlation Coefficient	<b>.500**</b>	.012	.034	.243	1.000
	Sig. (2-tailed)	.001	.937	.831	.117	
	N	43	43	43	43	43

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

**Table B-4.** Correlations between complexity metrics for VMAT plans.

		Corrected MU	MCS	EdgeMetric	PI - Plan Irregularity	PM - Plan Modulation	LT/AL	MISpeed	MIAccel	MITotal
<b>Plan Error (Absolute Average % Difference between TPS and TLD)</b>	Correlation	.005	.021	-.077	-.008	.040	-.006	.100	.089	<b>.148*</b>
	$\sigma$ (2-tailed)	.945	.765	.272	.903	.571	.931	.151	.200	.033
	N	207	207	207	207	207	207	207	207	207
Corrected MU	Correlation	1.000	<b>-.310**</b>	<b>.418**</b>	<b>.297**</b>	<b>.533**</b>	<b>.262**</b>	.067	<b>.200**</b>	.091
	$\sigma$ (2-tailed)		.000	.000	.000	.000	.000	.335	.004	.191
MCS (McNiven 2010)	Correlation	<b>-.310**</b>	1.000	<b>-.592**</b>	<b>-.618**</b>	<b>-.626**</b>	.114	<b>.404**</b>	<b>.148*</b>	<b>.489**</b>
	$\sigma$ (2-tailed)	.000		.000	.000	.000	.103	.000	.033	.000
EdgeMetric (Younge 2012)	Correlation	<b>.418**</b>	<b>-.592**</b>	1.000	<b>.921**</b>	<b>.390**</b>	.113	-.110	.131	<b>-.238**</b>
	$\sigma$ (2-tailed)	.000	.000		.000	.000	.107	.115	.060	.001
PI - Plan Irregularity (Du 2014)	Correlation	<b>.297**</b>	<b>-.618**</b>	<b>.921**</b>	1.000	<b>.383**</b>	-.027	<b>-.149*</b>	.073	<b>-.275**</b>
	$\sigma$ (2-tailed)	.000	.000	.000		.000	.701	.033	.296	.000
PM - Plan Modulation (Du 2014)	Correlation	<b>.533**</b>	<b>-.626**</b>	<b>.390**</b>	<b>.383**</b>	1.000	.019	<b>-.257**</b>	-.033	<b>-.143*</b>
	$\sigma$ (2-tailed)	.000	.000	.000	.000		.787	.000	.641	.039
LT/AL - Leaf Travel (Masi 2013)	Correlation	<b>.262**</b>	.114	.113	-.027	.019	1.000	<b>.611**</b>	<b>.676**</b>	<b>.598**</b>
	$\sigma$ (2-tailed)	.000	.103	.107	.701	.787		.000	.000	.000
MISpeed (Park 2014)	Correlation	.067	<b>.404**</b>	-.110	-.149*	<b>-.257**</b>	<b>.611**</b>	1.000	<b>.893**</b>	<b>.879**</b>
	$\sigma$ (2-tailed)	.335	.000	.115	.033	.000	.000		.000	.000
MIAccel	Correlation	<b>.200**</b>	<b>.148*</b>	.131	.073	-.033	<b>.676**</b>	<b>.893**</b>	1.000	<b>.833**</b>
	$\sigma$ (2-tailed)	.004	.033	.060	.296	.641	.000	.000		.000
MITotal	Correlation	.091	<b>.489**</b>	<b>-.238**</b>	<b>-.275**</b>	<b>-.143*</b>	<b>.598**</b>	<b>.879**</b>	<b>.833**</b>	1.000
	$\sigma$ (2-tailed)	.191	.000	.001	.000	.039	.000	.000	.000	

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

**Table B-5.** Correlations between complexity metrics for Eclipse plans.

		Corrected MU	MCS (McNiven 2010)	EdgeMetric (Younge 2012)	PI - Plan Irregularity (Du 2014)	PM - Plan Modulation (Du 2014)	MISpeed (Park 2014)	MIAccel	MITotal	LT/AL - Leaf Travel (Masi 2013)
<b>Plan Error (Absolute Average % Difference between TPS and TLD)</b>	Correlation	.025	.017	-.093	-.058	.055	-.012	.029	.017	-.059
	$\sigma$ (2-tailed)	.697	.785	.142	.364	.390	.886	.726	.833	.467
	N	249	249	249	249	249	152	152	152	152
Corrected MU	Correlation	1.000	<b>.130*</b>	<b>-.130*</b>	<b>-.182**</b>	<b>.549**</b>	<b>.276**</b>	<b>.406**</b>	<b>.373**</b>	<b>.431**</b>
	$\sigma$ (2-tailed)		.040	.041	.004	.000	.001	.000	.000	.000
MCS (McNiven 2010)	Correlation	.130*	1.000	<b>-.710**</b>	<b>-.725**</b>	<b>-.293**</b>	<b>.204*</b>	.042	.065	-.018
	$\sigma$ (2-tailed)	.040		.000	.000	.000	.012	.606	.424	.829
EdgeMetric (Younge 2012)	Correlation	-.130*	-.710**	1.000	<b>.952**</b>	.056	<b>.366**</b>	<b>.466**</b>	<b>.465**</b>	<b>.424**</b>
	$\sigma$ (2-tailed)	.041	.000		.000	.375	.000	.000	.000	.000
PI - Plan Irregularity (Du 2014)	Correlation	-.182**	-.725**	.952**	1.000	.058	<b>.195*</b>	<b>.280**</b>	<b>.277**</b>	.139
	$\sigma$ (2-tailed)	.004	.000	.000		.360	.016	.000	.001	.087
PM - Plan Modulation (Du 2014)	Correlation	.549**	-.293**	.056	.058	1.000	<b>-.218**</b>	.036	.010	.139
	$\sigma$ (2-tailed)	.000	.000	.375	.360		.007	.661	.907	.089
MISpeed (Park 2014)	Correlation	.276**	.204*	.366**	.195*	-.218**	1.000	<b>.903**</b>	<b>.903**</b>	<b>.747**</b>
	$\sigma$ (2-tailed)	.001	.012	.000	.016	.007		.000	.000	.000
MIAccel	Correlation	.406**	.042	.466**	.280**	.036	.903**	1.000	<b>.993**</b>	<b>.849**</b>
	$\sigma$ (2-tailed)	.000	.606	.000	.000	.661	.000		.000	.000
MITotal	Correlation	.373**	.065	.465**	.277**	.010	.903**	.993**	1.000	<b>.843**</b>
	$\sigma$ (2-tailed)	.000	.424	.000	.001	.907	.000	.000		.000
LT/AL - Leaf Travel (Masi 2013)	Correlation	.431**	-.018	.424**	.139	.139	.747**	.849**	.843**	1.000
	$\sigma$ (2-tailed)	.000	.829	.000	.087	.089	.000	.000	.000	

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).



**Table B-6.** Correlations between complexity metrics for Pinnacle plans.

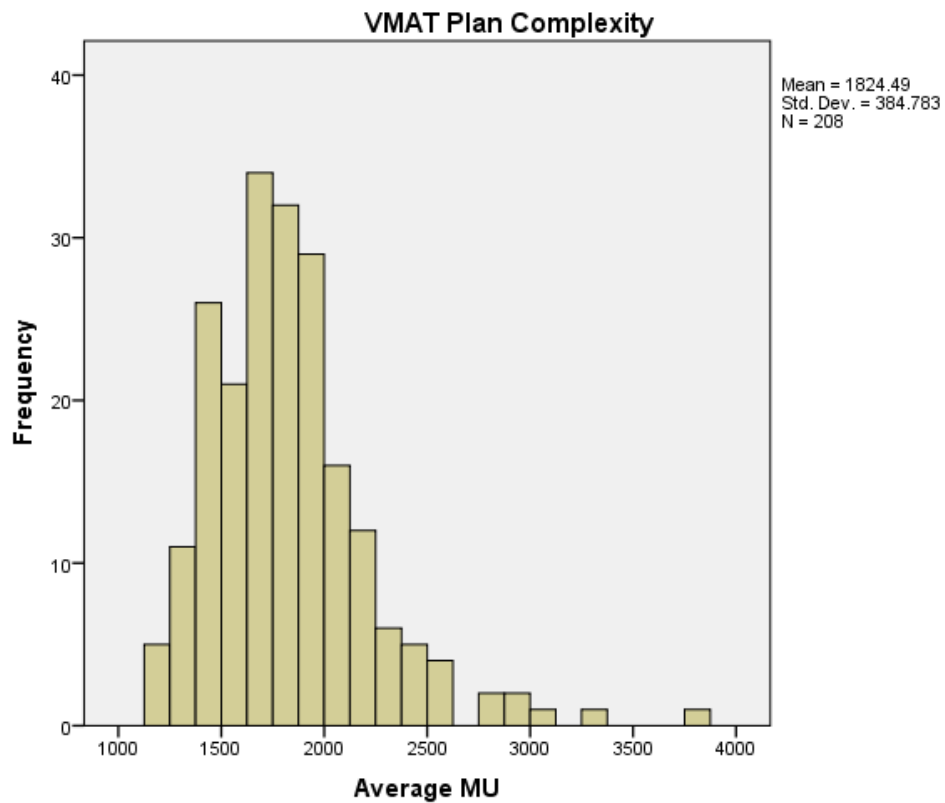
		Corrected MU	MCS (McNiven 2010)	EdgeMetric (Younge 2012)	PI - Plan Irregularity (Du 2014)	PM - Plan Modulation (Du 2014)	MISpeed (Park 2014)	MIAccel	MITotal	LT/AL - Leaf Travel (Masi 2013)
<b>Plan Error (Absolute Average % Difference between TPS and TLD)</b>	Correlation	.208	-.103	.233	.188	.158	.100	.113	.101	-.173
	$\sigma$ (2-tailed)	.087	.397	.054	.122	.196	.528	.474	.523	.274
	N	69	69	69	69	69	42	42	42	42
Corrected MU	Correlation	1.000	<b>-.319**</b>	.038	-.056	<b>.449**</b>	-.078	-.004	-.005	-.134
	$\sigma$ (2-tailed)		.008	.755	.648	.000	.623	.978	.973	.396
MCS (McNiven 2010)	Correlation	-.319**	1.000	.062	.225	<b>-.297*</b>	-.120	-.291	-.220	<b>.450**</b>
	$\sigma$ (2-tailed)	.008		.615	.063	.013	.449	.061	.161	.003
EdgeMetric (Younge 2012)	Correlation	.038	.062	1.000	.905**	<b>.238*</b>	.120	.233	.164	<b>-.355*</b>
	$\sigma$ (2-tailed)	.755	.615		.000	.049	.447	.138	.300	.021
PI - Plan Irregularity (Du 2014)	Correlation	-.056	.225	.905**	1.000	.140	.186	<b>.325*</b>	.248	<b>-.350*</b>
	$\sigma$ (2-tailed)	.648	.063	.000		.253	.238	.036	.113	.023
PM - Plan Modulation (Du 2014)	Correlation	.449**	-.297*	.238*	.140	1.000	-.109	.002	.039	<b>-.368*</b>
	$\sigma$ (2-tailed)	.000	.013	.049	.253		.494	.990	.807	.016
MISpeed (Park 2014)	Correlation	-.078	-.120	.120	.186	-.109	1.000	<b>.928**</b>	<b>.884**</b>	<b>.508**</b>
	$\sigma$ (2-tailed)	.623	.449	.447	.238	.494		.000	.000	.001
MIAccel	Correlation	-.004	-.291	.233	.325*	.002	.928**	1.000	<b>.897**</b>	<b>.378*</b>
	$\sigma$ (2-tailed)	.978	.061	.138	.036	.990	.000		.000	.014
MITotal	Correlation	-.005	-.220	.164	.248	.039	.884**	.897**	1.000	<b>.431**</b>
	$\sigma$ (2-tailed)	.973	.161	.300	.113	.807	.000	.000		.004
LT/AL - Leaf Travel (Masi 2013)	Correlation	-.134	.450**	-.355*	-.350*	-.368*	.508**	.378*	.431**	1.000
	$\sigma$ (2-tailed)	.396	.003	.021	.023	.016	.001	.014	.004	

\*\* . Correlation is significant at the 0.01 level (2-tailed).

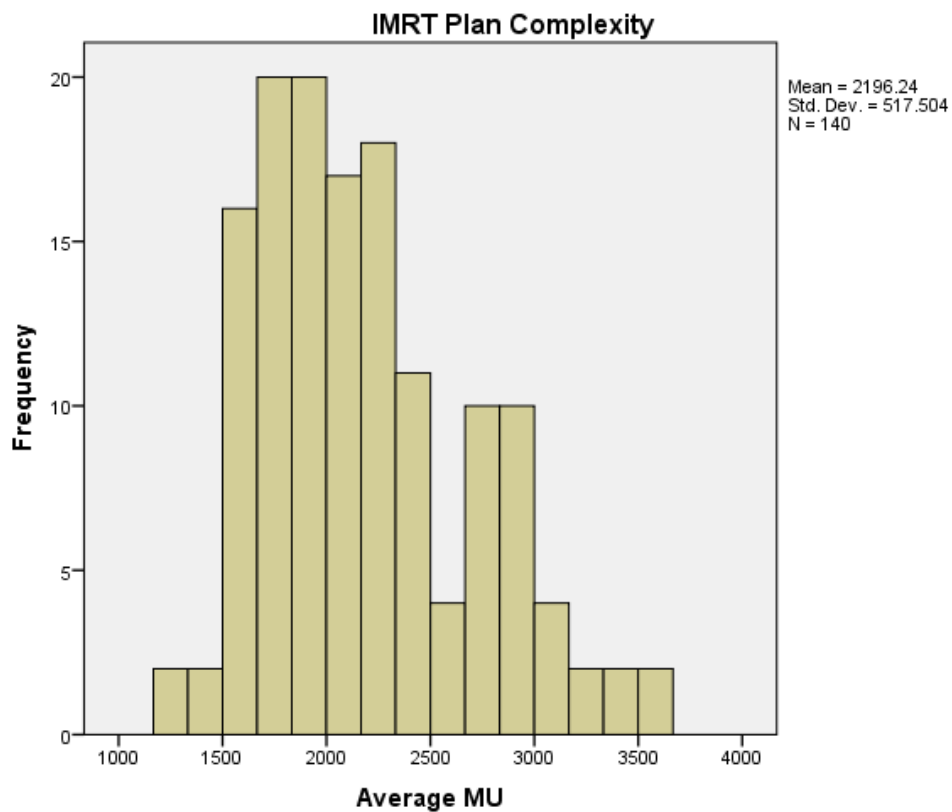
\* . Correlation is significant at the 0.05 level (2-tailed).

**Table B-7.** Distribution summary of complexity (as defined by MU) for VMAT and IMRT plans sampled in this work.

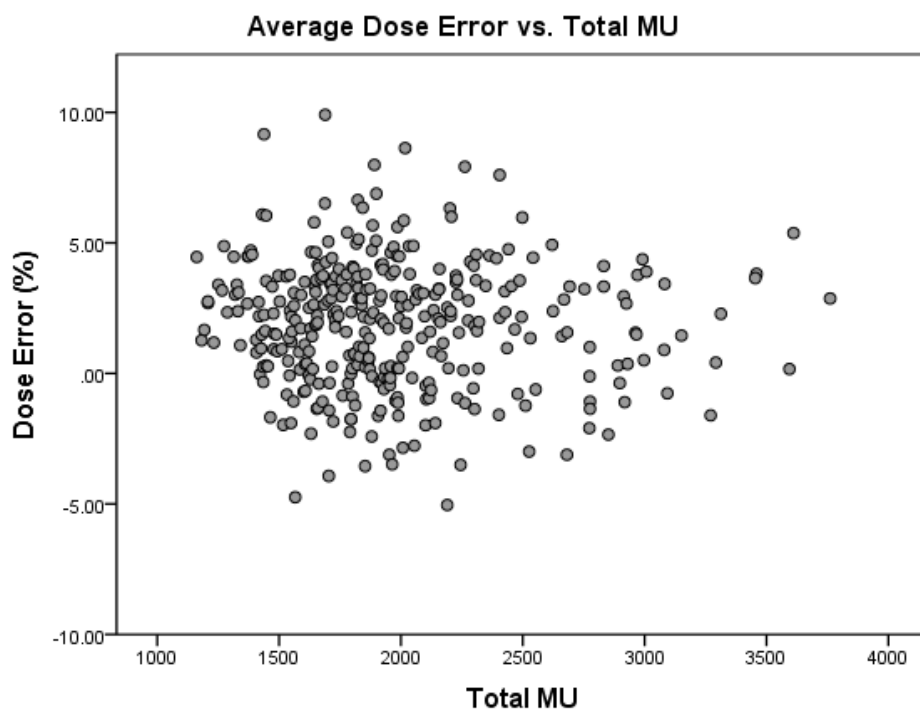
Average MU		VMAT Plans	IMRT Plans
N	Valid	208	140
	Missing	0	5
Mean		1824.49	2196.24
Median		1795.00	2099.00
Std. Deviation		384.783	517.504
Minimum		1162	1182
Maximum		3761	3611



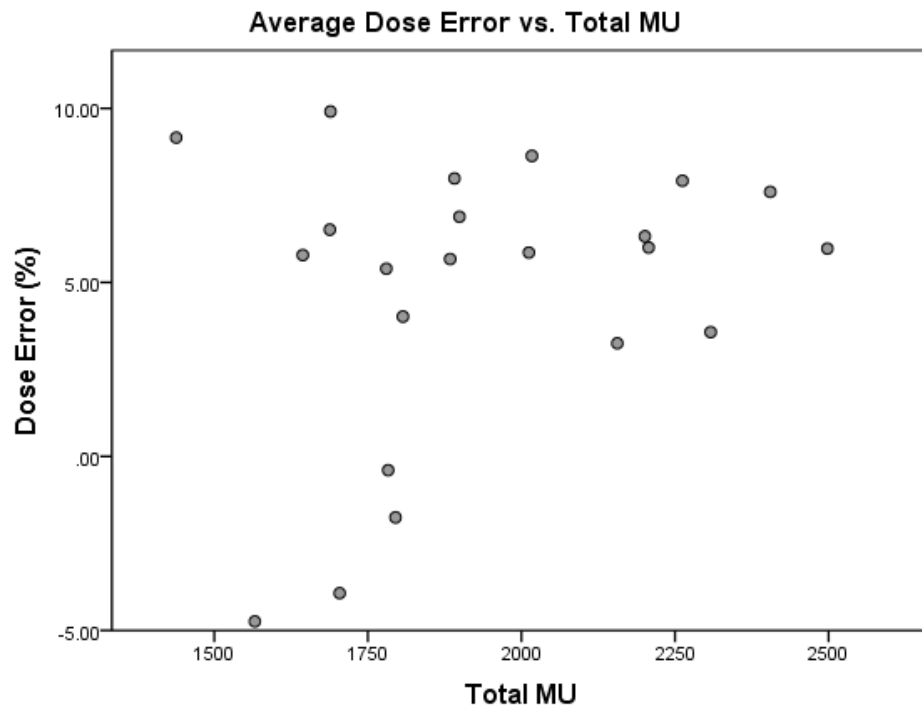
**Figure B-1.** Histogram of VMAT plan complexity, as defined by MU.



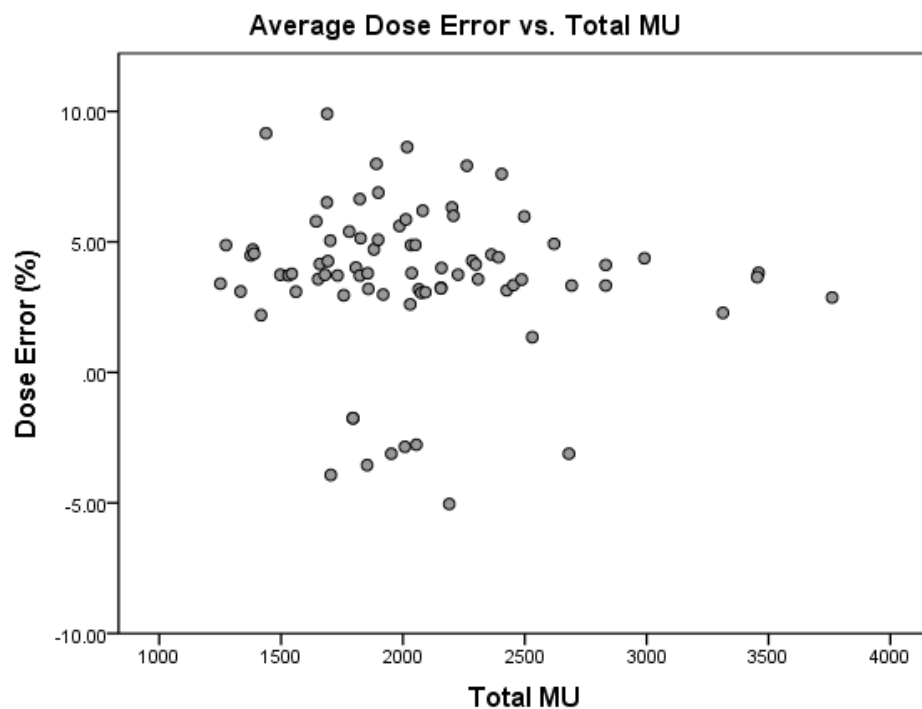
**Figure B-2.** Histogram of IMRT plan complexity, as defined by MU.



**Figure B-3.** Average dose error versus complexity (MU) for all machines (n=348).



**Figure B-4.** Average dose error versus complexity (MU) for failing phantom irradiations (n=22).



**Figure B-5.** Average dose error versus complexity (MU) for phantom irradiations diagnosed as having a TPS error through Mobius3D recalculation (n=81).

## **Appendix C: Supplement to Chapter 4**

This appendix serves as the supplement to  
Chapter 4: Survey of Photon Beam Modeling Parameters

This appendix contains the supplementary information referenced in the chapter text, namely the survey forms, instructions for how to identify the beam modeling parameters in each of the native TPS environments, and links to survey repository data. Additionally, this supplement includes other survey distributions of interest that are not included in Chapter 4 (see Figures C-1 and C-2). Summary statistics from the beam modeling parameter survey are included in Tables C-1 through C-24. Additionally, in-depth statistical information may be found on the IROC network drive at J:\Everyone\Mallory\Dissertation\Appendix C.

The compiled list of survey responses and summarized parameter value distributions are archived on Zenodo and the IROC Houston website at the following URLs:

- <https://doi.org/10.5281/zenodo.3357124>
- <http://rpc.mdanderson.org/RPC/IROCReferenceData.htm>

The un-masked dataset with RPC IDs may be found on the local IROC Houston network drive at J:\Everyone\Mallory\Dissertation\Appendix C. Because this dataset contains identifying information, it is not included in this text and is only available to IROC Houston personnel.

The following 7 pages include the beam modeling parameter survey form and visual instructions provided to radiotherapy institutions to aid in their response.

## Treatment Planning System Parameter Survey

Please provide the following information regarding your institution's treatment machine and treatment planning system (TPS) modeling parameters relevant to the photon energy used to irradiate the phantom. You may refer to the "How to Access Requested TPS Parameters" handout to determine where these parameters may be found in the TPS environment.

If your institution does not use Eclipse, Pinnacle<sup>3</sup>, or RayStation, please disregard this form.

MLC Model: \_\_\_\_\_

### **Eclipse Users:**

Effective Target Spot Size [mm] X: \_\_\_\_\_

Y: \_\_\_\_\_

MLC Transmission Factor: \_\_\_\_\_

Dosimetric Leaf Gap [cm]: \_\_\_\_\_

### **Pinnacle<sup>3</sup> Users:**

MLC has rounded leaves? Yes No

Rounded Leaf Tip Radius [cm]: \_\_\_\_\_

Tongue and Groove Width [cm]: \_\_\_\_\_

Additional Tongue and Groove Transmission: \_\_\_\_\_

Effective Source Size [cm] X (perpendicular): \_\_\_\_\_

Y (parallel): \_\_\_\_\_

Flattening Filter Gaussian Height: \_\_\_\_\_

Width: \_\_\_\_\_

MLC Transmission: \_\_\_\_\_

### **RayStation Users:**

Primary Source [cm] X Width: \_\_\_\_\_

Y Width: \_\_\_\_\_

MLC Transmission: \_\_\_\_\_

Tongue and Groove [cm]: \_\_\_\_\_

Leaf Tip Width [cm]: \_\_\_\_\_

MLC Position Offset [cm]: \_\_\_\_\_

MLC Position Gain: \_\_\_\_\_

MLC Position Curvature [1/cm]: \_\_\_\_\_

## How to Access Requested TPS Parameters: Eclipse Users

*Note: these instructions were developed using ARIA 13. If you are using an earlier version, parameters may be found under “Dosimetric Data” in the Beam Data workspace (see below).*

- Navigate to the “Beam Configuration” workspace.
  - This can be found under Quicklinks >> Treatment Planning >> Beam Configuration.
- Ensure that you are in the “Beam Data” workspace. This can be verified by selecting the workspace menu at the top to see your current selection.

### Version 11 or Earlier:

- On the left hand side, select the machine, beam energy, and algorithm used for IMRT.
- From the top menu select Beam Data >> Dosimetric Data. A pop-up window should appear.
  - Record values for **MLC transmission factor** and **Dosimetric Leaf Gap**.
- Continue with the instructions below.

### Version 13 or Higher:

- On the left hand side, select the machine, beam energy, and algorithm used for IMRT.
  - Under Algorithm >> Beam Data: MACHINE ID, select “Data: Parameters”.
  - Record values for **Effective target spot size (X and Y)**.

The screenshot shows the Eclipse Beam Configuration (Administrator) window. The left pane displays a tree view of beam data. The right pane displays a table of parameters.

**Left Pane Tree View:**

- External Beam: Eclipse CAP
  - Energy / Mode: 12E
  - Energy / Mode: 6X (selected with a red arrow and text: "Select appropriate beam energy and algorithm")
    - Algorithm: AAA\_CAP135MR1 (2/14 in Eclipse CAP,6X - 000)
    - Algorithm: AcurosXB\_CAP135MR1 (2/14 in Eclipse CAP,6X - 000)
    - Algorithm: CDC\_CAP135MR1 (Empty)
- Algorithm: AAA\_CAP135MR1 (2/14 in Eclipse CAP,6X - 000, \\Eclipse...)
  - Beam Data: Eclipse CAP,6X - 000
    - Data: General Parameters
    - Data: Model Parameters
    - Data: Parameters (highlighted with a red box)
    - Data: Calculated Parameters
  - Add-On: Open Field - 00
    - Data: Measured Diagonal Profiles
    - Data: Measured Depth Doses
    - Data: Measured Profiles
    - Data: Output Factors
    - Data: Mean Radial Energy

**Right Pane Table:**

Parameters - Parameter View	
Absolute dose reference field size [mm]	100.000000
Absolute dose calibration source-phantom distance [mm]	950.000000
Absolute dose calibration depth [mm]	50.000000
Reference dose at calibration depth [Gy]	1.000000
Reference MU at calibration depth [MU]	100.000000
Machine type	Varian Clinac 23EX
Effective target spot size in X-direction (IEC 61217) [mm]	0.000000
Effective target spot size in Y-direction (IEC 61217) [mm]	0.000000
Leaf transmission for Elekta Beam Modulator	0.000000
Open field profile measurement direction (IEC 61217)	Collimator X

- Navigate to RT Administration using the Quicklinks or Home menu.
  - Select “Radiation and Imaging Devices” at the top of the window.
  - In the “Overview” pane, select the MACHINE ID by clicking on its name.

RT Administration (Administrator)

No Patient Selected | No Current Activity | Worklist | Quicklinks | Logout

File Edit View Insert Workspace Tools Window

System and Facilities | Patients | **Radiation and Imaging Devices** | Clinical Data | Templates

External Beam: Eclipse CAP

Wedge | Tray | Blocks and Compensators | Imager

**Overview** | Operating Limits | Technique | Primary Fluence Mode | Energy Mode | Configured EMT | Slots | Applicator | MLC

New | Delete | Import

**Devices**

ID	Type	Model	Manufacturer	Scale	Object Status	Operating Status	Last Modified	Description
Def_CTScanner	CT Scanner	CTScanner	Varian Medical Systems		Active		1/15/2015 4:44:04 PM by SysAdmin	
<b>Eclipse CAP</b>	<b>External Beam</b>	Varian 2100C/D	Varian Medical Systems	IEC61217	Active	Ready	1/15/2015 4:21:13 PM by SysAdmin	
EclipseCAP_AGL	External Beam	Elekta Synergy Agility	Elekta	IEC61217	Active	Ready	1/15/2015 4:21:18 PM by SysAdmin	
EclipseCAP_ART	External Beam	Siemens Artiste	Siemens	IEC61217	Active	Ready	1/15/2015 4:21:16 PM by SysAdmin	
EclipseCAP_TB	External Beam	TDS	Varian Medical Systems	IEC61217	Active	Ready	1/15/2015 4:21:14 PM by SysAdmin	
EclipseCAPElekta	External Beam	Elekta SL 20	Elekta	IEC61217	Active	Ready	1/15/2015 4:21:15 PM by SysAdmin	
EclipseCAPSiem	External Beam	Siemens Generic	Siemens	IEC61217	Active	Ready	1/15/2015 4:21:15 PM by SysAdmin	
RPC New	Brachy Unit	Remote Afterloading	RPC		Active	Ready	8/10/2016 8:56:09 AM by Physicist	

- Select the “MLC” tab.
  - Record values for **MLC Transmission Factor** and **Dosimetric Leaf Gap** for the applicable energy used for IMRT.

RT Administration (Administrator)

No Patient Selected | No Current Activity | Worklist | Quicklinks | Logout

File Edit View Insert Workspace Tools Window

System and Facilities | Patients | Radiation and Imaging Devices | Clinical Data | Templates

External Beam: Eclipse CAP

Wedge | Tray | Blocks and Compensators | Imager

Overview | Operating Limits | Technique | Primary Fluence Mode | Energy Mode | Configured EMT | Slots | Applicator | **MLC**

New MLC... | Delete MLC | Import MLC...

**MLC**

ID	MLC120
Name	
Status	Active
Add-On Material	CAP_MLC
Select Add-On Material Energy Mod...	Select
Add-On Material Energy Modes	6X 6X-SRS 12E 12E-SRS
Internal Code	
Model	Millennium 120
Manufacturer	Varian Medical Systems
Serial Number	

**Dosimetric Properties for Material 'CAP\_MLC'**

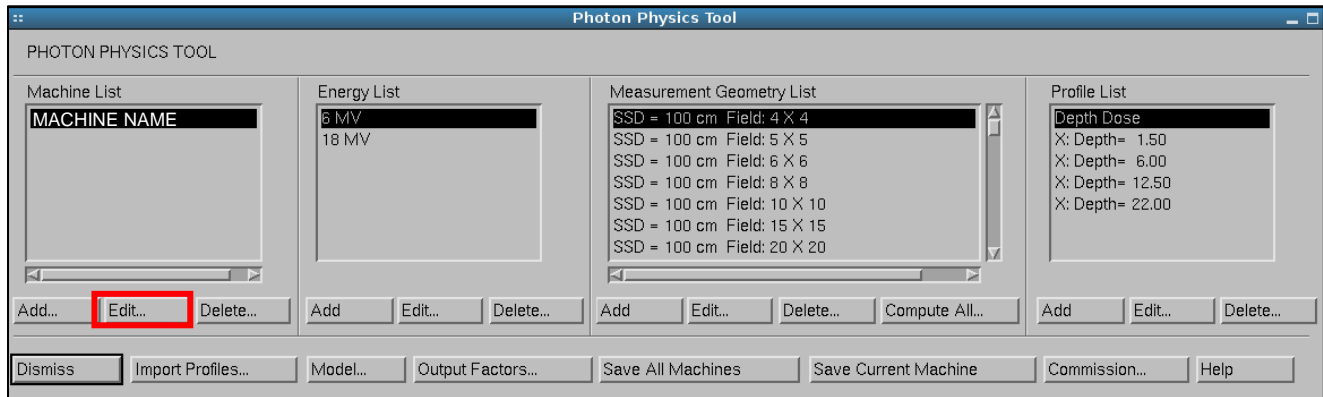
Energy	Transmission Factor	Dosimetric Leaf Gap [cm]	Eq. Machines Group ID	Apply Eq. Ma
<b>6X</b>	0.0200	0.2000		
6X-SRS	0.0200	0.2000		
12E	0.0200			
12E-SRS	0.0200			



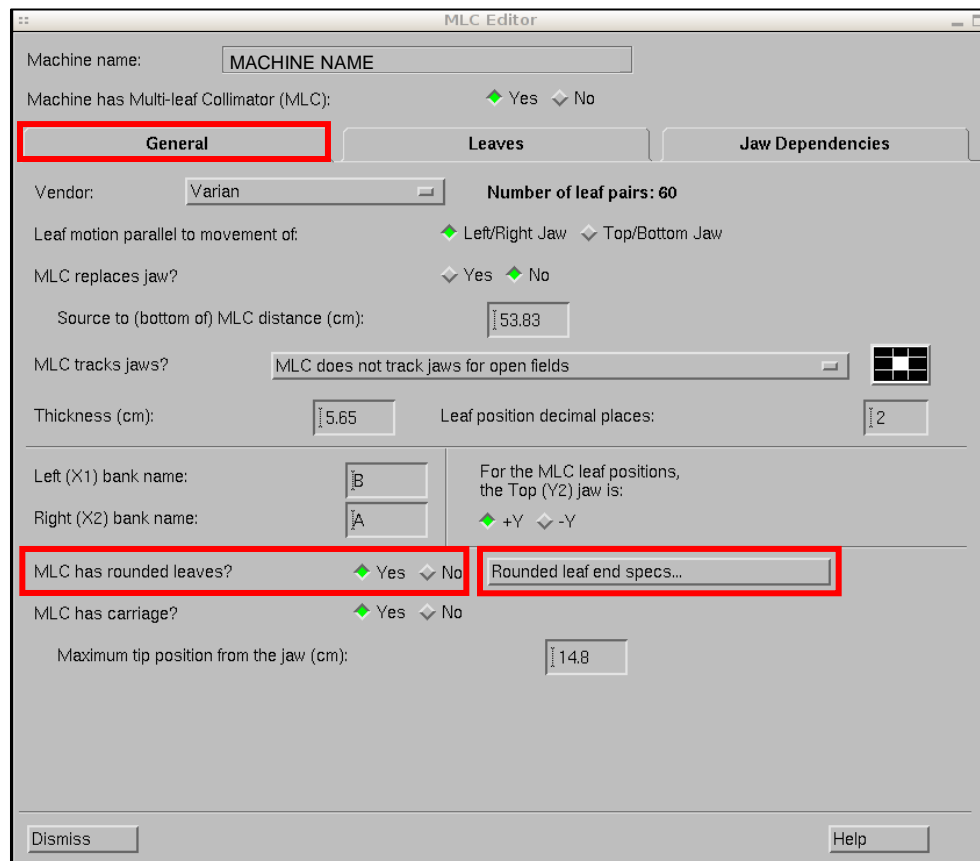
## How to Access Requested TPS Parameters: Pinnacle<sup>3</sup> Users

*Note: These instructions were developed using Pinnacle<sup>3</sup> version 9.10.*

- Open Physics Tools, and select “Photon Physics Tool...”
- In the Photon Physics Tool window, select the machine and energy used for IMRT from the machine list and select “Edit...”

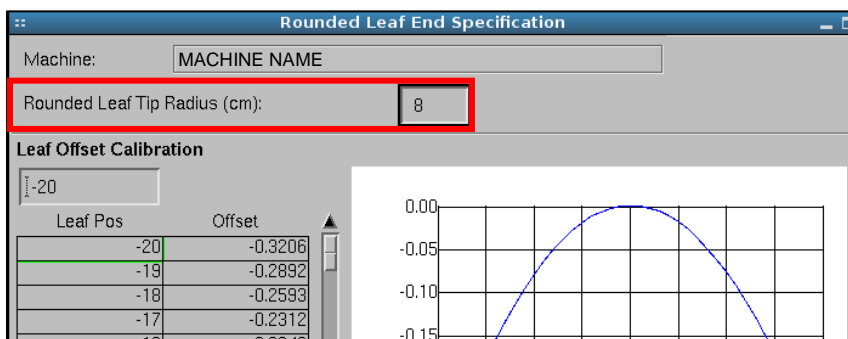


- Select “MLC...”. This will lead you to the MLC Editor window.
- Record if MLC model uses rounded leaves.

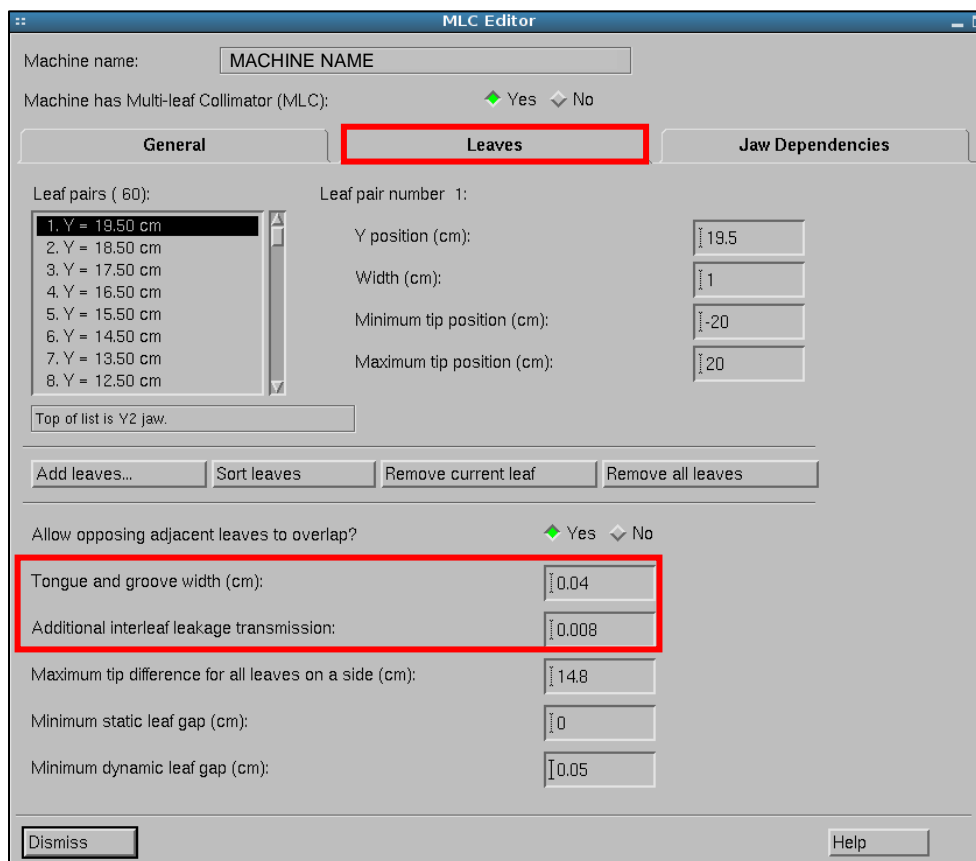


- If applicable, in the General tab of the MLC Editor window, select “Rounded leaf end specs...”. This will open the “Rounded Leaf End Specification” window.

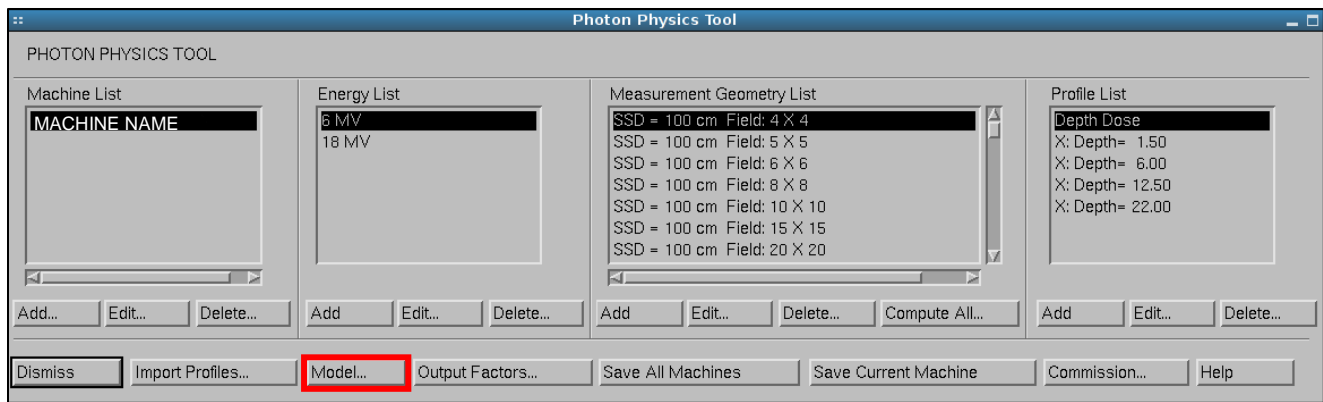
- Record the value for “**Rounded Leaf Tip Radius (cm)**” at the top of the window. Close the window.



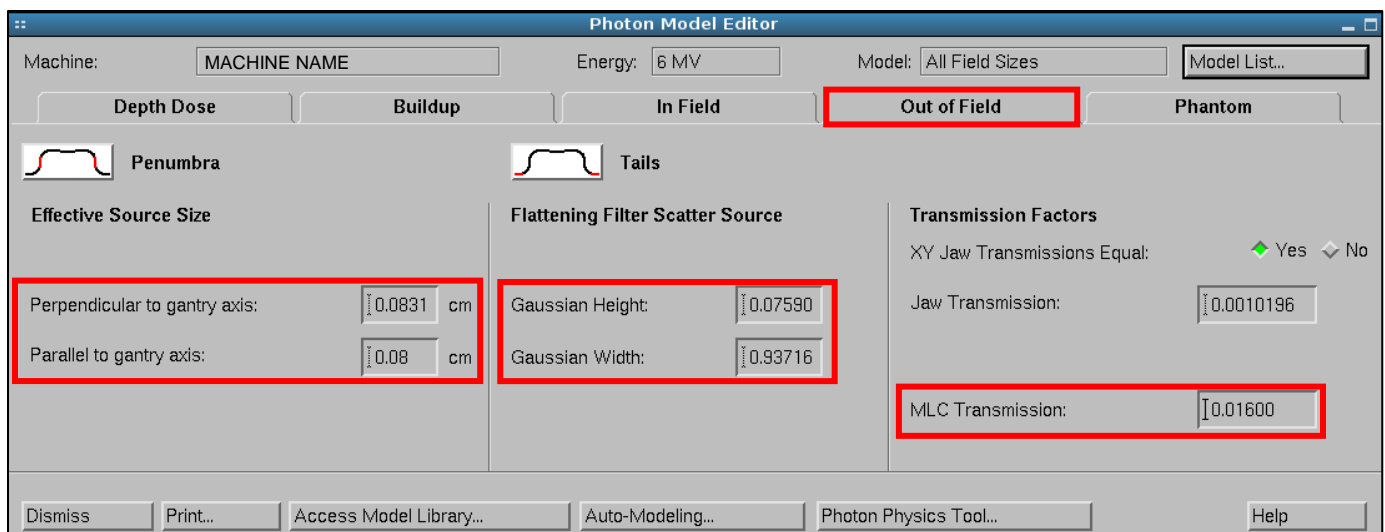
- In the MLC Editor window, select the “Leaves” tab.
  - Record values for **Tongue and groove width** and **Additional interleaf leakage transmission** values. Close the MLC Editor window.



- Back in the “Photon Physics Tool” window, select “Model...”
  - If the “Model...” option is unavailable, ensure that measurements have been read in for the machine model (you should see populated values in the “Measurement Geometry List” and “Profile List”). You may need to select “Read Measured Data.”



- In the Photon Model Editor window, select the “Out of Field” tab.
  - Record values for **Effective source size** (X and Y), **Flattening Filter Gaussian Height and Width**, and **MLC Transmission**.



## How to Access Requested TPS Parameters: RayStation Users

*Note: These instructions were developed using RayStation 5.*

- Open RayPhysics and select the “Beam Commissioning” tab at the top of the window.
- On the left hand side of the machine tree view, select the commissioned machine and the photon energy that you use for IMRT.
  - At the top of the parameter workspace on the right, select the tab for “Fluence”.
  - Record values for **Primary source X and Y widths, MLC transmission, MLC Tongue and groove, Leaf tip width, and MLC position offset, gain, and curvature.\***
    - Record either the x- or y-defined MLC calibrations, depending on how you MLC is configured in the TPS.

RayPhysics 5

Beam Commissioning

Varian 2110, 6 MV

Commissioned machines

Varian 2110 [15 Jun 2016, 11:53:17 (hr:min:sec)]

Photon energies

6.0 MV

Electron energies

6.0 MeV

9.0 MeV

12.0 MeV

16.0 MeV

6x6

10x10

15x15

20x20

25x25

Enter edit mode

Import curves...

Import cone dose curves...

Remove curves...

Export curves...

Energy Spectrum

Off Axis

Fluence

Output Factor Corrections

Wedge Factor Corrections

Sources

Source	Eff. dist. to source [cm]	X width [cm]	Y width [cm]	Weight
Primary	-	0.086	0.085	-
Flattening filter	10.00	2.528	-	0.07618
Electrons	-	50.000	-	0.02899

Weight of flattening filter electron source

0.100

☐ Flattening filter free

Collimator position

Collimator	Eff. dist. to source [cm]	Transmission
Y-jaws	35.70	-
X-jaws	44.50	0.01000
MLC	53.80	0.01600

Collimator calibration

Collimator	Offset [cm]	Gain	Curvature [1/cm]
Y-jaws	0.000	0.0000	0.00000
X-jaws	0.000	0.0000	0.00000
MLC x-position	0.000	0.0000	0.00080
MLC y-position	-	0.0000	-

Additional MLC parameters

Tongue and groove [cm]

0.050

Leaf tip width [cm]

0.300

Curves

Modulation

Open

Modulation

None

Type

Depth

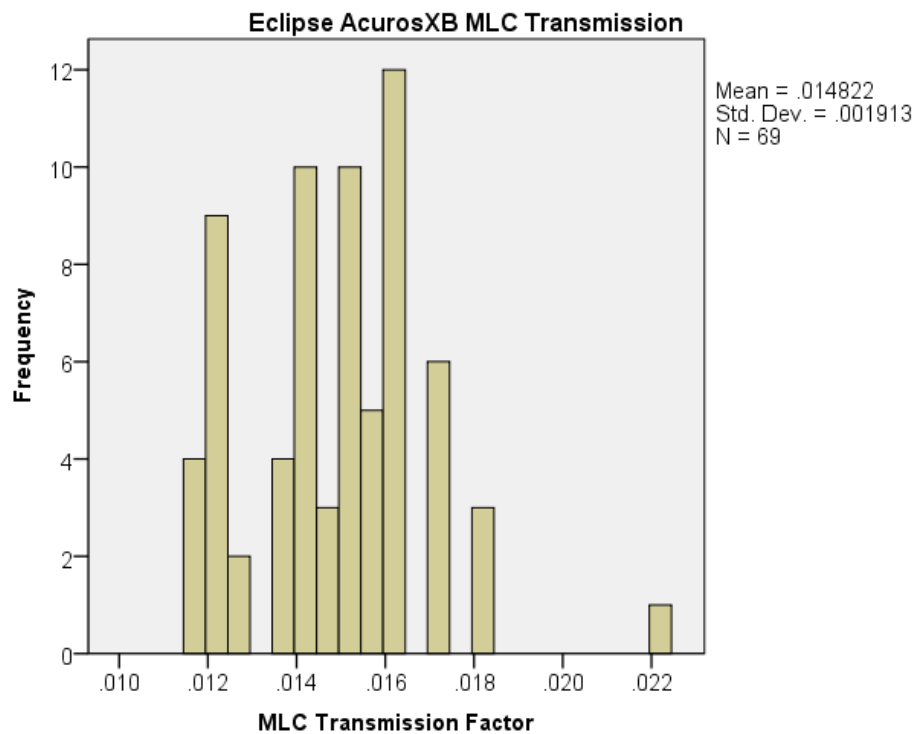
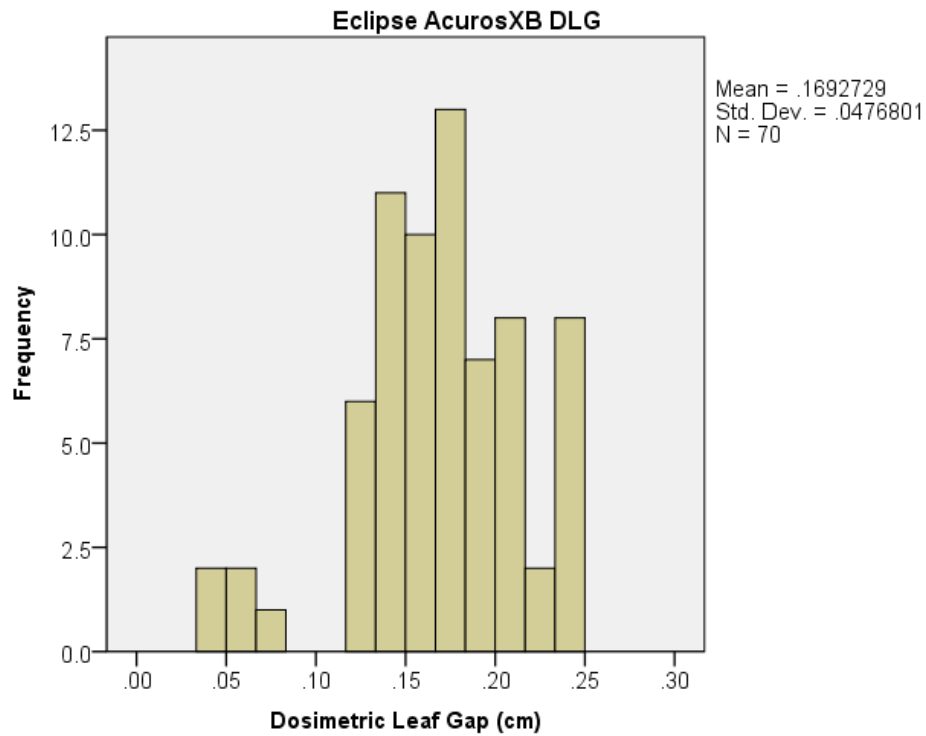
Field size [cm]

All

Depth [cm]

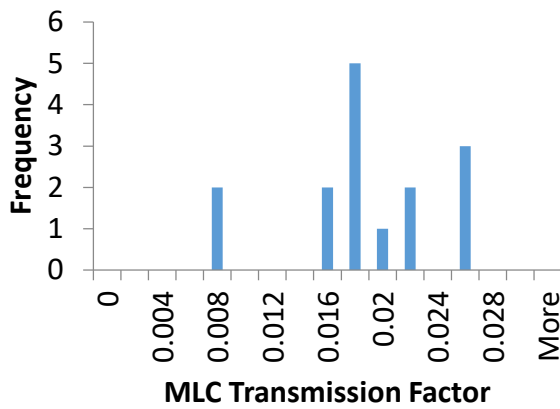
All

Compute selected

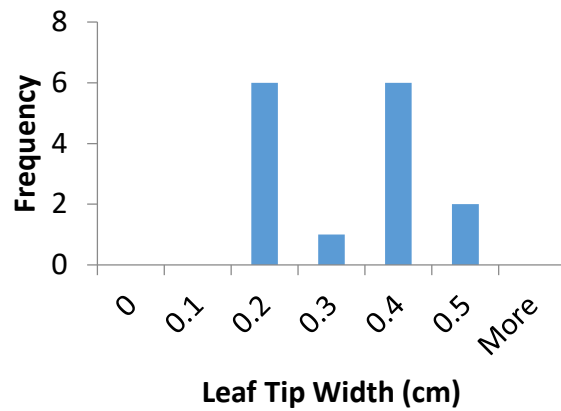


**Figure C-1.** Survey responses for the standard 6 MV Varian Base class for Eclipse AcurosXB: dosimetric leaf gap (top), and MLC transmission factor (bottom).

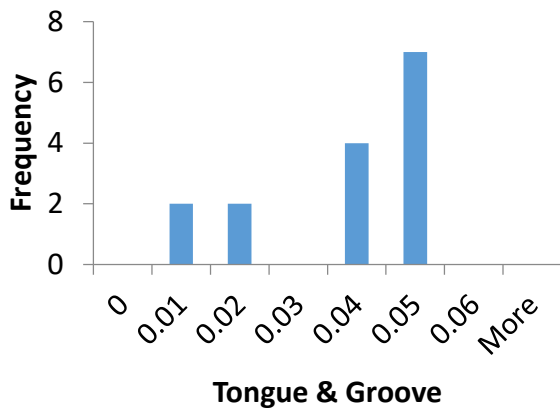
**a.** Varian Base (6X) - 120 MLC  
RayStation: MLC Transmission



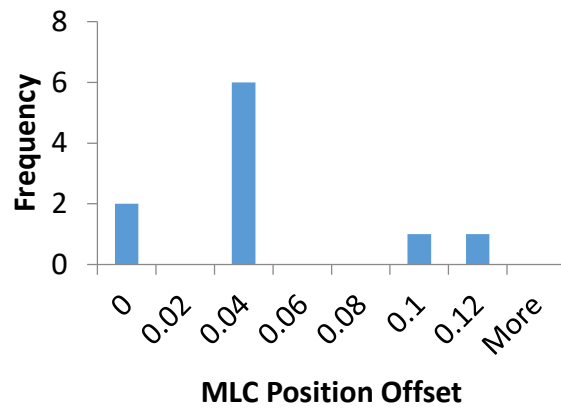
**b.** Varian Base (6X) - 120 MLC  
RayStation: Leaf Tip Width



**c.** Varian Base (6X) - 120 MLC  
RayStation: Tongue & Groove



**d.** Varian Base (6X) - 120 MLC  
RayStation: MLC Offset



**Figure C-2.** Survey responses for the 6 MV Varian Base class for RayStation: a) MLC transmission factor, b) leaf tip width, c) tongue and groove, d) MLC position offset.

**Table C-1.** Beam modeling summary statistics for 6MV beams modeled in Eclipse AAA.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Varian 2100	Varian 2300	Varian 6EX	Varian 600	Trilogy SRS	Elekta Agility	Elekta MLCi
Effective Target Spot Size X [mm]	N =	312	19	226	61	8	4	2	4	22	4	5
	2.5th	0.000	0.000	0.000	0.000	-	-	-	-	-	-	-
	25th	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000	0.000	0.000	0.500
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>1.000</b>
	75th	0.000	0.000	0.000	0.000	0.000	0.375	-	0.000	0.000	1.500	2.500
	97.5th	1.250	0.000	1.750	1.225	-	-	-	-	-	-	-
	St. Dev.	2.352	0.000	2.765	0.302	0.000	0.250	0.000	0.000	0.266	1.000	1.140
Effective Target Spot Size Y [mm]	N =	312	19	226	61	8	4	2	4	22	4	5
	2.5th	0.000	0.000	0.000	0.000	-	-	-	-	-	-	-
	25th	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000	0.000	0.000	0.000
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
	75th	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000	0.000	0.000	1.750
	97.5th	1.000	0.000	1.900	1.088	-	-	-	-	-	-	-
	St. Dev.	1.807	0.000	2.760	0.225	0.000	0.000	0.000	0.000	0.112	0.000	1.095
Dosimetric Leaf Gap [cm]	N =	327	23	233	61	8	2	2	4	21	4	4
	2.5th	0.1000	-	0.0868	0.0046	-	-	-	-	-	-	-
	25th	0.1550	0.0636	0.1140	0.0399	0.1800	-	-	0.1595	0.1500	0.0023	0.0475
	50th	<b>0.1700</b>	<b>0.0800</b>	<b>0.1330</b>	<b>0.0590</b>	<b>0.1940</b>	<b>0.1575</b>	<b>0.1800</b>	<b>0.2090</b>	<b>0.1700</b>	<b>0.0105</b>	<b>0.1000</b>
	75th	0.1900	0.1400	0.1600	0.0910	0.2165	-	-	0.2300	0.1800	0.0270	0.4000
	97.5th	0.2300	-	0.2211	0.1945	-	-	-	-	-	-	-
	St. Dev.	0.1993	0.0481	0.0894	0.0448	0.0279	0.0035	0.0566	0.0385	0.3184	0.0135	0.2142
MLC Transmission Factor	N =	326	23	233	61	8	2	2	4	21	4	5
	2.5th	0.0118	-	0.0114	0.0039	-	-	-	-	-	-	-
	25th	0.0145	0.0114	0.0141	0.0117	0.0129	-	-	0.0130	0.0140	0.0037	0.0069
	50th	<b>0.0158</b>	<b>0.0125</b>	<b>0.0150</b>	<b>0.0120</b>	<b>0.0159</b>	<b>0.0171</b>	<b>0.0147</b>	<b>0.0138</b>	<b>0.0150</b>	<b>0.0062</b>	<b>0.0150</b>
	75th	0.0165	0.0146	0.0153	0.0128	0.0168	-	-	0.0155	0.0160	0.0063	0.0173
	97.5th	0.0200	-	0.0189	0.0374	-	-	-	-	-	-	-
	St. Dev.	0.0020	0.0026	0.0023	0.0057	0.0026	0.0030	0.0016	0.0014	0.0024	0.0017	0.0074

**Table C-2.** Beam modeling summary statistics for 6MV beams modeled in Eclipse AcurosXB.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Varian 2100	Trilogy SRS	Elekta Agility	Elekta MLCi
Effective Target Spot Size X [mm]	N =	69	5	68	31	2	5	1	1
	2.5th	0.000	-	0.000	-	-	-	-	-
	25th	1.000	1.000	1.000	1.000	-	1.000	-	-
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	<b>0.750</b>	<b>1.000</b>	<b>1.000</b>	<b>3.000</b>
	75th	1.125	1.250	1.000	1.000	-	1.375	-	-
	97.5th	1.800	-	1.500	-	-	-	-	-
	St. Dev.	0.397	0.224	0.346	0.269	1.061	0.224	-	-
Effective Target Spot Size Y [mm]	N =	69	5	68	31	2	5	1	1
	2.5th	0.000	-	0.000	-	-	-	-	-
	25th	0.500	0.500	1.000	1.000	-	0.750	-	-
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	<b>0.500</b>	<b>1.000</b>	<b>0.000</b>	<b>3.000</b>
	75th	1.000	1.000	1.000	1.000	-	1.000	-	-
	97.5th	1.800	-	1.500	-	-	-	-	-
	St. Dev.	0.458	0.447	0.412	0.341	0.707	0.224	-	-
Dosimetric Leaf Gap (DLG)	N =	70	5	69	29	2	5	1	1
	2.5th	0.0400	-	0.0968	-	-	-	-	-
	25th	0.1432	0.0818	0.1250	0.0470	-	0.0690	-	-
	<b>50th</b>	<b>0.1700</b>	<b>0.1000</b>	<b>0.1390</b>	<b>0.0650</b>	<b>0.2250</b>	<b>0.1600</b>	<b>0.0300</b>	<b>0.1240</b>
	75th	0.2000	0.1200	0.1650	0.1005	-	0.1871	-	-
	97.5th	0.2500	-	0.4420	-	-	-	-	-
	St. Dev.	0.0477	0.0270	0.1136	0.0477	0.0354	0.0632	-	-
MLC Transmission Factor	N =	69	5	69	29	2	5	1	1
	2.5th	0.0117	-	0.0119	-	-	-	-	-
	25th	0.0138	0.0110	0.0143	0.0117	-	0.0112	-	-
	<b>50th</b>	<b>0.0150</b>	<b>0.0120</b>	<b>0.0150</b>	<b>0.0120</b>	<b>0.0142</b>	<b>0.0117</b>	<b>0.0042</b>	<b>0.0100</b>
	75th	0.0160	0.0121	0.0159	0.0125	-	0.0163	-	-
	97.5th	0.0190	-	0.0200	-	-	-	-	-
	St. Dev.	0.0019	0.0009	0.0021	0.0029	0.0025	0.0028	-	-



**Table C-3.** Beam modeling summary statistics for 6MV beams modeled in RayStation.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Trilogy SRS	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Primary Source X Width [cm]	N =	15	2	12	1	41	2	5
	2.5th	-	-	-	-	0.090	-	-
	25th	0.050	-	0.060	-	0.118	-	0.132
	50th	<b>0.057</b>	<b>0.089</b>	<b>0.083</b>	<b>0.040</b>	<b>0.130</b>	<b>0.120</b>	<b>0.160</b>
	75th	0.106	-	0.140	-	0.143	-	0.170
	97.5th	-	-	-	-	0.206	-	-
	St. Dev.	0.031	0.055	0.037	-	0.024	0.000	0.024
Primary Source Y Width [cm]	N =	15	2	12	1	41	2	5
	2.5th	-	-	-	-	0.003	-	-
	25th	0.052	-	0.055	-	0.070	-	0.089
	50th	<b>0.070</b>	<b>0.083</b>	<b>0.083</b>	<b>0.040</b>	<b>0.080</b>	<b>0.060</b>	<b>0.115</b>
	75th	0.075	-	0.110	-	0.123	-	0.150
	97.5th	-	-	-	-	0.259	-	-
	St. Dev.	0.017	0.032	0.024	-	0.050	0.000	0.032
MLC Transmission	N =	15	2	12	1	41	2	5
	2.5th	-	-	-	-	0.0000	-	-
	25th	0.0150	-	0.0147	-	0.0015	-	0.0075
	50th	<b>0.0180</b>	<b>0.0163</b>	<b>0.0150</b>	<b>0.0200</b>	<b>0.0030</b>	<b>0.0095</b>	<b>0.0100</b>
	75th	0.0220	-	0.0197	-	0.0070	-	0.0110
	97.5th	-	-	-	-	0.0160	-	-
	St. Dev.	0.0056	0.0053	0.0044	-	0.0041	0.0000	0.0026
Tongue & Groove [cm]	N =	15	2	12	1	41	2	5
	2.5th	-	-	-	-	0.000	-	-
	25th	0.020	-	0.040	-	0.050	-	0.025
	50th	<b>0.040</b>	<b>0.050</b>	<b>0.050</b>	<b>0.050</b>	<b>0.050</b>	<b>0.0550</b>	<b>0.060</b>
	75th	0.050	-	0.050	-	0.075	-	0.105
	97.5th	-	-	-	-	0.110	-	-
	St. Dev.	0.015	0.000	0.013	-	0.027	0.0000	0.054
Leaf Tip Width [cm]	N =	15	2	12	1	41	2	5
	2.5th	-	-	-	-	0.100	-	-
	25th	0.200	-	0.200	-	0.150	-	0.200
	50th	<b>0.320</b>	<b>0.300</b>	<b>0.200</b>	<b>0.200</b>	<b>0.200</b>	<b>0.550</b>	<b>0.200</b>
	75th	0.400	-	0.275	-	0.450	-	0.250
	97.5th	-	-	-	-	0.985	-	-
	St. Dev.	0.110	0.141	0.092	-	0.214	0.000	0.045
MLC Position Offset [cm]	N =	10	0	9	1	32	2	3
	2.5th	-	-	-	-	-	-	-
	25th	0.017	-	0.005	-	0.005	-	-
	50th	<b>0.040</b>	-	<b>0.022</b>	<b>0.000</b>	<b>0.010</b>	<b>-0.020</b>	<b>0.020</b>
	75th	0.055	-	0.048	-	0.024	-	-
	97.5th	-	-	-	-	-	-	-
	St. Dev.	0.039	-	0.022	-	0.049	0.000	0.025

Table C-3, continued.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Trilogy SRS	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
	<b>N =</b>	10	0	9	1	32	2	3
<b>MLC Position Gain</b>	2.5th	-	-	-	-	-	-	-
	25th	0.0000	-	-0.0004	-	0.0010	-	-
	<b>50th</b>	<b>0.0015</b>	-	<b>0.0000</b>	<b>0.0000</b>	<b>0.0038</b>	<b>0.0040</b>	<b>0.0020</b>
	75th	0.0049	-	0.0020	-	0.0079	-	-
	97.5th	-	-	-	-	-	-	-
	St. Dev.	0.0060	-	0.0050	-	0.0304	0.0000	0.0042
	<b>N =</b>	10	0	9	1	32	2	3
<b>MLC Position Curvature [1/cm]</b>	2.5th	-	-	-	-	-	-	-
	25th	0.0000	-	0.0000	-	0.0000	-	-
	<b>50th</b>	<b>0.0000</b>	-	<b>0.0008</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0005</b>
	75th	0.0009	-	0.0008	-	0.0002	-	-
	97.5th	-	-	-	-	-	-	-
	St. Dev.	0.0005	-	0.0004	-	0.0336	0.0000	0.0003

**Table C-4.** Beam modeling summary statistics for 6MV beams modeled in Pinnacle.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Varian 2100	Varian 2300	Varian 6EX	Trilogy SRS	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Effective Source Size X [cm]	N =	55	1	25	7	3	1	1	2	69	5	27
	2.5th	0.0100	-	-	-	-	-	-	-	0.0100	-	-
	25th	0.0471	-	0.0344	0.0250	-	-	-	-	0.0500	0.0100	0.0400
	50th	<b>0.0800</b>	<b>0.0100</b>	<b>0.0800</b>	<b>0.0400</b>	<b>0.0200</b>	<b>0.0600</b>	<b>0.1393</b>	<b>0.0202</b>	<b>0.0900</b>	<b>0.0200</b>	<b>0.1019</b>
	75th	0.1106	-	0.1387	0.0815	-	-	-	-	0.1350	0.0548	0.1500
	97.5th	0.2200	-	-	-	-	-	-	-	0.3125	-	-
	St. Dev.	0.0498	-	0.0610	0.0261	0.0346	-	-	0.0112	0.0645	0.0233	0.0588
Effective Source Size Y [cm]	N =	55	1	25	7	3	1	1	2	69	5	27
	2.5th	0.0100	-	-	-	-	-	-	-	0.0100	-	-
	25th	0.0387	-	0.0472	0.0250	-	-	-	-	0.0400	0.0217	0.0150
	50th	<b>0.0750</b>	<b>0.0100</b>	<b>0.0831</b>	<b>0.0387</b>	<b>0.0200</b>	<b>0.0782</b>	<b>0.1000</b>	<b>0.0157</b>	<b>0.0690</b>	<b>0.0450</b>	<b>0.0800</b>
	75th	0.1000	-	0.1088	0.0608	-	-	-	-	0.1180	0.0625	0.1431
	97.5th	0.1716	-	-	-	-	-	-	-	0.2441	-	-
	St. Dev.	0.0427	-	0.0467	0.0267	0.0361	-	-	0.0062	0.0583	0.0239	0.0588
MLC Transmission	N =	46	1	23	6	3	1	1	1	65	5	28
	2.5th	0.0092	-	-	-	-	-	-	-	0.0010	-	-
	25th	0.0139	-	0.0110	0.0123	-	-	-	-	0.0013	0.0056	0.0061
	50th	<b>0.0160</b>	<b>0.0140</b>	<b>0.0150</b>	<b>0.0137</b>	<b>0.0100</b>	<b>0.0230</b>	<b>0.0160</b>	<b>0.0122</b>	<b>0.0030</b>	<b>0.0060</b>	<b>0.0098</b>
	75th	0.0180	-	0.0201	0.0148	-	-	-	-	0.0046	0.0070	0.0148
	97.5th	0.0465	-	-	-	-	-	-	-	0.0121	-	-
	St. Dev.	0.0065	-	0.0053	0.0019	0.0048	-	-	-	0.0025	0.0008	0.0067
Rounded Leaf Tip Radius [cm]	N =	54	1	25	7	4	1	1	2	70	4	28
	2.5th	6.0	-	-	-	-	-	-	-	15.9	-	-
	25th	8.0	-	8.0	16.0	8.0	-	-	-	17.0	12.2	15.0
	50th	<b>8.0</b>	<b>16.0</b>	<b>8.0</b>	<b>16.0</b>	<b>8.0</b>	<b>8.0</b>	<b>8.0</b>	<b>8.0</b>	<b>17.0</b>	<b>15.1</b>	<b>15.0</b>
	75th	9.0	-	9.0	16.0	50.0	-	-	-	17.0	18.0	15.0
	97.5th	14.5	-	-	-	-	-	-	-	25.0	-	-
	St. Dev.	1.9	-	2.3	4.3	28.0	-	-	0.0	1.8	3.3	1.0

Table C-4, continued.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Varian 2100	Varian 2300	Varian 6EX	Trilogy SRS	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Tongue and Groove Width [cm]	N =	55	1	25	7	4	1	1	2	71	5	28
	2.5th	0.010	-	-	-	-	-	-	-	0.004	-	-
	25th	0.100	-	0.040	0.005	0.070	-	-	-	0.100	0.040	0.050
	50th	<b>0.100</b>	<b>16.000</b>	<b>0.040</b>	<b>0.040</b>	<b>0.085</b>	<b>0.070</b>	<b>0.040</b>	<b>0.100</b>	<b>0.100</b>	<b>0.050</b>	<b>0.100</b>
	75th	0.100	-	0.100	0.040	0.100	-	-	-	0.200	0.120	0.100
	97.5th	0.110	-	-	-	-	-	-	-	0.200	-	-
	St. Dev.	0.028	-	0.039	0.020	0.017	-	-	0.000	0.063	0.043	0.046
Additional Tongue and Groove Transmission	N =	55	1	25	7	3	1	1	2	71	5	28
	2.5th	0.000	-	-	-	-	-	-	-	0.000	-	-
	25th	0.008	-	0.005	0.003	-	-	-	-	0.000	0.006	0.007
	50th	<b>0.010</b>	<b>0.005</b>	<b>0.010</b>	<b>0.005</b>	<b>0.030</b>	<b>0.030</b>	<b>0.008</b>	<b>0.010</b>	<b>0.001</b>	<b>0.007</b>	<b>0.010</b>
	75th	0.010	-	0.010	0.005	-	-	-	-	0.002	0.007	0.010
	97.5th	0.038	-	-	-	-	-	-	-	0.002	-	-
	St. Dev.	0.007	-	0.005	0.001	0.012	-	-	0.000	0.001	0.001	0.015
Flattening Filter Gaussian Height	N =	55	1	25	7	3	1	1	2	69	5	27
	2.5th	0.0403	-	-	-	-	-	-	-	0.0428	-	-
	25th	0.0710	-	0.0665	0.0664	-	-	-	-	0.0612	0.0660	0.0700
	50th	<b>0.0785</b>	<b>0.0729</b>	<b>0.0739</b>	<b>0.0800</b>	<b>0.0800</b>	<b>0.0900</b>	<b>0.0650</b>	<b>0.0692</b>	<b>0.0750</b>	<b>0.0750</b>	<b>0.0800</b>
	75th	0.0907	-	0.0857	0.0910	-	-	-	-	0.0846	0.0775	0.0859
	97.5th	0.4549	-	-	-	-	-	-	-	0.1120	-	-
	St. Dev.	0.0827	-	0.0129	0.0141	0.0169	-	-	0.0131	0.0175	0.0075	0.0750
Flattening Filter Gaussian Width	N =	55	1	25	7	3	1	1	2	69	5	27
	2.5th	0.4522	-	-	-	-	-	-	-	1.2766	-	-
	25th	1.1161	-	1.2576	1.2028	-	-	-	-	1.4500	1.4607	1.4272
	50th	<b>1.3000</b>	<b>1.3408</b>	<b>1.5890</b>	<b>1.3000</b>	<b>1.5000</b>	<b>1.0000</b>	<b>0.9545</b>	<b>2.1757</b>	<b>1.6656</b>	<b>1.8000</b>	<b>1.8000</b>
	75th	1.5000	-	1.8440	1.7987	-	-	-	-	2.1414	2.0000	2.0972
	97.5th	5.0000	-	-	-	-	-	-	-	3.3763	-	-
	St. Dev.	0.7997	-	0.7099	0.6827	0.1443	-	-	0.8142	0.5190	0.2729	0.6655

**Table C-5.** Beam modeling summary statistics for 6 FFF beams modeled in Eclipse AAA.

Parameter	Percentile	Varian TrueBeam	Varian TrueBeam (HD MLC)	Varian Halcyon	Elekta Agility
Effective Target Spot Size X [mm]	N =	129	46	5	2
	2.5th	0.000	0.000	-	-
	25th	0.000	0.000	0.000	-
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
	75th	0.375	0.000	0.000	-
	97.5th	1.600	1.000	-	-
	St. Dev.	3.532	0.309	0.000	0.000
Effective Target Spot Size Y [mm]	N =	129	46	5	2
	2.5th	0.000	0.000	-	-
	25th	0.000	0.000	0.000	-
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
	75th	0.000	0.000	0.000	-
	97.5th	1.900	0.930	-	-
	St. Dev.	3.536	0.170	0.000	0.000
Dosimetric Leaf Gap [cm]	N =	133	48	6	2
	2.5th	0.0214	0.0084	-	-
	25th	0.0960	0.0310	0.0100	-
	50th	<b>0.1200</b>	<b>0.0473</b>	<b>0.0100</b>	<b>0.0365</b>
	75th	0.1545	0.0900	0.0100	-
	97.5th	0.2263	0.1678	-	-
	St. Dev.	0.0687	0.0393	0.0000	0.0007
MLC Transmission Factor	N =	133	48	6	2
	2.5th	0.0080	0.0017	-	-
	25th	0.0120	0.0096	0.0035	-
	50th	<b>0.0127</b>	<b>0.0100</b>	<b>0.0047</b>	<b>0.0038</b>
	75th	0.0136	0.0109	0.0047	-
	97.5th	0.0200	0.0130	.	-
	St. Dev.	0.0074	0.0021	0.0019	0.0003

**Table C-6.** Beam modeling summary statistics for 6 FFF beams modeled in Eclipse AcurosXB.

Parameter	Percentile	Varian TrueBeam	Varian TrueBeam (HD MLC)
	<b>N =</b>	<b>33</b>	<b>21</b>
<b>Effective Target Spot Size X [mm]</b>	2.5th	-	-
	25th	1.000	1.000
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000
	97.5th	-	-
	St. Dev.	0.339	0.158
	<b>N =</b>	<b>33</b>	<b>21</b>
<b>Effective Target Spot Size Y [mm]</b>	2.5th	-	-
	25th	1.000	0.875
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000
	97.5th	-	-
	St. Dev.	0.448	0.321
	<b>N =</b>	<b>34</b>	<b>20</b>
<b>Dosimetric Leaf Gap (DLG)</b>	2.5th	-	-
	25th	0.1030	0.0390
	<b>50th</b>	<b>0.1357</b>	<b>0.0560</b>
	75th	0.1696	0.0900
	97.5th	-	-
	St. Dev.	0.1305	0.0442
	<b>N =</b>	<b>34</b>	<b>20</b>
<b>MLC Transmission Factor</b>	2.5th	-	-
	25th	0.0121	0.0090
	<b>50th</b>	<b>0.0130</b>	<b>0.0100</b>
	75th	0.0138	0.0104
	97.5th	-	-
	St. Dev.	0.0022	0.0016

**Table C-7.** Beam modeling summary statistics for 6 FFF beams modeled in RayStation.

Parameter	Percentile	Varian TrueBeam	Elekta Agility
Primary Source X Width [cm]	<b>N =</b>	8	10
	2.5th	-	-
	25th	0.056	0.130
	<b>50th</b>	<b>0.068</b>	<b>0.140</b>
	75th	0.077	0.144
	97.5th	-	-
	St. Dev.	0.031	0.023
Primary Source Y Width [cm]	<b>N =</b>	8	10
	2.5th	-	-
	25th	0.065	0.059
	<b>50th</b>	<b>0.108</b>	<b>0.090</b>
	75th	0.108	0.150
	97.5th	-	-
	St. Dev.	0.023	0.072
MLC Transmission	<b>N =</b>	8	10
	2.5th	-	-
	25th	0.0005	0.0010
	<b>50th</b>	<b>0.0038</b>	<b>0.0027</b>
	75th	0.0177	0.0043
	97.5th	-	-
	St. Dev.	0.0086	0.0016
Tongue & Groove [cm]	<b>N =</b>	8	10
	2.5th	-	-
	25th	0.040	0.050
	<b>50th</b>	<b>0.040</b>	<b>0.050</b>
	75th	0.040	0.050
	97.5th	-	-
	St. Dev.	0.012	0.009
Leaf Tip Width [cm]	<b>N =</b>	8	10
	2.5th	-	-
	25th	0.300	0.100
	<b>50th</b>	<b>0.750</b>	<b>0.100</b>
	75th	1.000	0.200
	97.5th	-	-
	St. Dev.	0.370	0.106
MLC Position Offset [cm]	<b>N =</b>	7	5
	2.5th	-	-
	25th	0.002	0.010
	<b>50th</b>	<b>0.002</b>	<b>0.020</b>
	75th	0.042	0.023
	97.5th	-	-
	St. Dev.	0.020	0.007

Parameter	Percentile	Varian TrueBeam	Elekta Agility
MLC Position Gain	<b>N =</b>	7	5
	2.5th	-	-
	25th	-0.0015	-0.0046
	<b>50th</b>	<b>0.0000</b>	<b>0.0000</b>
	75th	0.0000	0.0005
	97.5th	-	-
	St. Dev.	0.0007	0.0027
MLC Position Curvature [1/cm]	<b>N =</b>	7	5
	2.5th	-	-
	25th	0.0000	0.0000
	<b>50th</b>	<b>0.0017</b>	<b>0.0000</b>
	75th	0.0017	0.0000
	97.5th	-	-
	St. Dev.	0.0009	0.0000

**Table C-8.** Beam modeling summary statistics for 6 FFF beams modeled in Pinnacle. For brevity, 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles are not reported due to limited survey samples.

Parameter	Percentile	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility
Effective Source Size X [cm]	<b>N =</b>	12	4	13
	25th	0.0334	0.0379	0.0500
	<b>50th</b>	<b>0.0559</b>	<b>0.0562</b>	<b>0.0538</b>
	75th	0.0963	0.0964	0.0847
	St. Dev.	0.0400	0.0328	0.0364
Effective Source Size Y [cm]	<b>N =</b>	12	4	13
	25th	0.0530	0.0286	0.0600
	<b>50th</b>	<b>0.0978</b>	<b>0.0500</b>	<b>0.0834</b>
	75th	0.1143	0.1023	0.1513
	St. Dev.	0.0370	0.0401	0.0440
MLC Transmission	<b>N =</b>	11	4	13
	25th	0.0110	0.0106	0.0015
	<b>50th</b>	<b>0.0138</b>	<b>0.0113</b>	<b>0.0020</b>
	75th	0.0200	0.0147	0.0030
	St. Dev.	0.0075	0.0023	0.0014
Rounded Leaf Tip Radius [cm]	<b>N =</b>	12	4	13
	25th	8.0	16.0	17.0
	<b>50th</b>	<b>8.0</b>	<b>16.0</b>	<b>17.0</b>
	75th	10.0	16.0	17.0
	St. Dev.	2.4	0.0	2.2
Tongue and Groove Width [cm]	<b>N =</b>	12	4	13
	25th	0.018	0.005	0.053
	<b>50th</b>	<b>0.070</b>	<b>0.023</b>	<b>0.100</b>
	75th	0.100	0.040	0.200
	St. Dev.	0.042	0.020	0.075
Additional Tongue and Groove Transmission	<b>N =</b>	12	4	13
	25th	0.005	0.004	0.000
	<b>50th</b>	<b>0.010</b>	<b>0.005</b>	<b>0.002</b>
	75th	0.010	0.005	0.002
	St. Dev.	0.003	0.001	0.001
Flattening Filter Gaussian Height	<b>N =</b>	12	4	13
	25th	0.0285	0.0170	0.0300
	<b>50th</b>	<b>0.0408</b>	<b>0.0359</b>	<b>0.0350</b>
	75th	0.0534	0.0403	0.0415
	St. Dev.	2.5271	0.0133	0.0111
Flattening Filter Gaussian Width	<b>N =</b>	12	4	13
	25th	1.2719	0.7035	1.0000
	<b>50th</b>	<b>2.1634</b>	<b>1.8754</b>	<b>1.3146</b>
	75th	2.5835	3.1050	1.5300
	St. Dev.	1.0868	1.2957	0.3522



**Table C-9.** Beam modeling summary statistics for 10 MV beams modeled in Eclipse AAA.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility	Elekta MLCi
Effective Target Spot Size X [mm]	N =	61	6	156	46	1	1
	2.5th	0.000	0.000	0.000	0.000	-	-
	25th	0.000	0.000	0.000	0.000	-	-
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>3.000</b>
	75th	0.000	0.000	0.000	0.000	-	-
	97.5th	1.000	0.000	1.611	1.000	-	-
	St. Dev.	0.250	0.000	3.309	0.258	-	-
Effective Target Spot Size Y [mm]	N =	61	6	156	46	1	1
	2.5th	0.000	0.000	0.000	0.000	-	-
	25th	0.000	0.000	0.000	0.000	-	-
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>2.500</b>
	75th	0.000	0.000	0.000	0.000	-	-
	97.5th	1.000	0.000	1.900	0.165	-	-
	St. Dev.	0.180	0.000	3.311	0.029	-	-
Dosimetric Leaf Gap [cm]	N =	63	8	159	47	1	1
	2.5th	0.1480	-	0.0700	0.0168	-	-
	25th	0.1700	0.0325	0.1240	0.0473	-	-
	50th	<b>0.1900</b>	<b>0.0920</b>	<b>0.1490</b>	<b>0.0600</b>	<b>0.0290</b>	<b>0.5000</b>
	75th	0.2100	0.1775	0.1770	0.1090	-	-
	97.5th	0.2930	-	0.2375	0.2080	-	-
	St. Dev.	0.0320	0.0764	0.0951	0.0431	-	-
MLC Transmission Factor	N =	63	8	159	47	1	1
	2.5th	0.0131	-	0.0140	0.0028	-	-
	25th	0.0170	0.0139	0.0166	0.0130	-	-
	50th	<b>0.0180</b>	<b>0.0142</b>	<b>0.0170</b>	<b>0.0140</b>	<b>0.0066</b>	<b>0.0170</b>
	75th	0.0190	0.0165	0.0179	0.0145	-	-
	97.5th	0.0232	-	0.0244	0.0185	-	-
	St. Dev.	0.0021	0.0026	0.0019	0.0024	-	-

**Table C-10.** Beam modeling summary statistics for 10 MV beams modeled in Eclipse AcurosXB.

Parameter	Percentile	Varian Base	Varian TrueBeam	Varian TrueBeam (HD MLC)
Effective Target Spot Size X [mm]	<b>N =</b>	16	49	16
	2.5th	-	0.000	-
	25th	1.000	1.000	1.000
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000	1.000
	97.5th	-	1.500	-
	St. Dev.	0.437	0.346	0.316
Effective Target Spot Size Y [mm]	<b>N =</b>	16	49	16
	2.5th	-	0.000	-
	25th	0.250	1.000	1.000
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000	1.000
	97.5th	-	1.500	-
	St. Dev.	0.447	0.395	0.375
Dosimetric Leaf Gap (DLG)	<b>N =</b>	16	50	15
	2.5th	-	0.1137	-
	25th	0.1676	0.1570	0.0496
	<b>50th</b>	<b>0.1900</b>	<b>0.1700</b>	<b>0.0620</b>
	75th	0.2100	0.2025	0.1050
	97.5th	-	1.2339	-
	St. Dev.	0.0356	0.2120	0.0478
MLC Transmission Factor	<b>N =</b>	16	50	15
	2.5th	-	0.0147	-
	25th	0.0169	0.0166	0.0130
	<b>50th</b>	<b>0.0170</b>	<b>0.0170</b>	<b>0.0137</b>
	75th	0.0178	0.0176	0.0150
	97.5th	-	0.0255	-
	St. Dev.	0.0008	0.0020	0.0016

**Table C-11.** Beam modeling summary statistics for 10 MV beams modeled in RayStation.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Elekta Agility	Elekta MLCi
Primary Source X Width [cm]	N =	4	2	3	24	2
	25th	0.063	-	-	0.124	-
	50th	<b>0.076</b>	<b>0.090</b>	<b>0.050</b>	<b>0.138</b>	<b>0.125</b>
	75th	0.122	-	-	0.161	-
	St. Dev.	0.034	0.057	0.015	0.026	0.049
Primary Source Y Width [cm]	N =	4	2	3	24	2
	25th	0.055	-	-	0.064	-
	50th	<b>0.076</b>	<b>0.086</b>	<b>0.084</b>	<b>0.100</b>	<b>0.150</b>
	75th	0.104	-	-	0.150	-
	St. Dev.	0.026	0.036	0.017	0.057	0.099
MLC Transmission	N =	4	2	3	24	2
	25th	0.0123	-	-	0.0025	-
	50th	<b>0.0195</b>	<b>0.0218</b>	<b>0.0156</b>	<b>0.0044</b>	<b>0.0075</b>
	75th	0.0271	-	-	0.0050	-
	St. Dev.	0.0079	0.0108	0.0065	0.0022	0.0035
Tongue & Groove [cm]	N =	4	2	3	24	2
	25th	0.020	-	-	0.050	-
	50th	<b>0.050</b>	<b>0.050</b>	<b>0.050</b>	<b>0.050</b>	<b>0.025</b>
	75th	0.050	-	-	0.100	-
	St. Dev.	0.020	0.000	0.023	0.094	0.035
Leaf Tip Width [cm]	N =	4	2	3	24	2
	25th	0.058	-	-	0.125	-
	50th	<b>0.200</b>	<b>0.250</b>	<b>0.200</b>	<b>0.200</b>	<b>0.200</b>
	75th	0.425	-	-	0.500	-
	St. Dev.	0.203	0.071	0.173	0.210	0.000
MLC Position Offset [cm]	N =	3	0	2	17	1
	25th	-	-	-	0.000	-
	50th	<b>0.020</b>	-	<b>0.010</b>	<b>0.000</b>	<b>0.200</b>
	75th	-	-	-	0.005	-
	St. Dev.	0.053	-	0.014	0.012	-
MLC Position Gain	N =	3	0	2	17	1
	25th	-	-	-	0.0000	-
	50th	<b>0.0000</b>	-	<b>0.0100</b>	<b>0.0020</b>	<b>0.0200</b>
	75th	-	-	-	0.0025	-
	St. Dev.	0.0115	-	0.0141	0.0017	-
MLC Position Curvature [1/cm]	N =	3	0	2	17	1
	25th	-	-	-	0.0000	-
	50th	<b>0.0000</b>	-	<b>0.0000</b>	<b>0.0001</b>	<b>0.0035</b>
	75th	-	-	-	0.0002	-
	St. Dev.	0.0005	-	0.0000	0.0001	-

**Table C-12.** Beam modeling summary statistics for 10 MV beams modeled in Pinnacle.

Parameter	Percentile	Varian Base	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Effective Source Size X [cm]	N =	6	16	4	41	2	15
	2.5th	-	-	-	0.0100	-	-
	25th	0.0361	0.0819	0.0206	0.0349	-	0.0788
	50th	<b>0.0818</b>	<b>0.1093</b>	<b>0.0294</b>	<b>0.0900</b>	<b>0.0351</b>	<b>0.1090</b>
	75th	0.1134	0.1400	0.0391	0.1000	-	0.1285
	97.5th	-	-	-	0.2624	-	-
	St. Dev.	0.0459	0.0502	0.0099	0.0660	0.0355	0.0616
Effective Source Size Y [cm]	N =	6	16	4	41	2	15
	2.5th	-	-	-	0.0100	-	-
	25th	0.0308	0.0578	0.0209	0.0275	-	0.0460
	50th	<b>0.0600</b>	<b>0.0989</b>	<b>0.0319</b>	<b>0.0700</b>	<b>0.0328</b>	<b>0.0747</b>
	75th	0.1134	0.1377	0.0557	0.1000	-	0.1463
	97.5th	-	-	-	0.2051	-	-
	St. Dev.	0.0501	0.0452	0.0186	0.0525	0.0153	0.0725
MLC Transmission	N =	6	15	4	40	2	15
	2.5th	-	-	-	0.0010	-	-
	25th	0.0150	0.0125	0.0121	0.0016	-	0.0050
	50th	<b>0.0180</b>	<b>0.0197</b>	<b>0.0138</b>	<b>0.0036</b>	<b>0.0032</b>	<b>0.0080</b>
	75th	0.0202	0.0221	0.0163	0.0049	-	0.0126
	97.5th	-	-	-	0.0079	-	-
	St. Dev.	0.0031	0.0058	0.0022	0.0016	0.0003	0.0067
Rounded Leaf Tip Radius [cm]	N =	6	16	4	41	1	15
	2.5th	-	-	-	12.3	-	-
	25th	8.0	8.0	16.0	17.0	-	15.0
	50th	<b>9.0</b>	<b>8.0</b>	<b>16.0</b>	<b>17.0</b>	<b>18.0</b>	<b>15.0</b>
	75th	12.0	10.0	16.0	17.0	-	15.0
	97.5th	-	-	-	25.0	-	-
	St. Dev.	2.0	2.2	0.0	2.3	-	0.5
Tongue and Groove Width [cm]	N =	6	15	4	41	2	15
	2.5th	-	-	-	0.003	-	-
	25th	0.093	0.040	0.005	0.100	-	0.050
	50th	<b>0.100</b>	<b>0.040</b>	<b>0.023</b>	<b>0.100</b>	<b>0.075</b>	<b>0.100</b>
	75th	0.100	0.100	0.040	0.200	-	0.100
	97.5th	-	-	-	0.200	-	-
	St. Dev.	0.012	0.037	0.020	0.059	0.064	0.049
Additional Tongue and Groove Transmission	N =	6	15	4	41	2	15
	2.5th	-	-	-	0.000	-	-
	25th	0.009	0.005	0.004	0.000	-	0.010
	50th	<b>0.010</b>	<b>0.010</b>	<b>0.005</b>	<b>0.001</b>	<b>0.006</b>	<b>0.010</b>
	75th	0.015	0.010	0.005	0.002	-	0.030
	97.5th	-	-	-	0.002	-	-
	St. Dev.	0.009	0.007	0.001	0.001	0.001	0.016

Table C-12, continued.

Parameter	Percentile	Varian Base	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Flattening Filter Gaussian Height	N =	6	16	4	41	2	15
	2.5th	-	-	-	0.0600	-	-
	25th	0.0707	0.0782	0.0749	0.0650	-	0.0600
	50th	<b>0.0752</b>	<b>0.0837</b>	<b>0.0846</b>	<b>0.0733</b>	<b>0.0820</b>	<b>0.0715</b>
	75th	0.0855	0.0865	0.0969	0.0849	-	0.0890
	97.5th	-	-	-	0.1490	-	-
	St. Dev.	0.0083	0.0088	0.0114	0.0192	0.0028	0.0183
Flattening Filter Gaussian Width	N =	6	16	4	41	2	15
	2.5th	-	-	-	1.0941	-	-
	25th	1.2096	1.3590	1.2889	1.4300	-	1.4352
	50th	<b>1.3084</b>	<b>1.5517</b>	<b>1.7013</b>	<b>1.5874</b>	<b>1.5548</b>	<b>1.7569</b>
	75th	1.6905	1.7743	2.7480	1.8000	-	2.0000
	97.5th	-	-	-	2.4850	-	-
	St. Dev.	0.5454	0.4158	0.7920	0.3181	0.2053	0.5979

**Table C-13.** Beam modeling summary statistics for 10 FFF beams modeled in Eclipse AAA.

Parameter	Percentile	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility
Effective Target Spot Size X [mm]	<b>N =</b>	109	42	1
	2.5th	0.000	0.000	-
	25th	0.000	0.000	-
	<b>50th</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
	75th	0.000	0.000	-
	97.5th	1.500	0.970	-
	St. Dev.	3.838	0.193	-
Effective Target Spot Size Y [mm]	<b>N =</b>	109	42	1
	2.5th	0.000	0.000	-
	25th	0.000	0.000	-
	<b>50th</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
	75th	0.000	0.000	-
	97.5th	1.700	0.570	-
	St. Dev.	3.843	0.097	-
Dosimetric Leaf Gap [cm]	<b>N =</b>	109	43	1
	2.5th	0.0785	0.0037	-
	25th	0.1135	0.0410	-
	<b>50th</b>	<b>0.1370</b>	<b>0.0789</b>	<b>0.0470</b>
	75th	0.1750	0.1110	-
	97.5th	0.2345	0.4717	-
	St. Dev.	0.0981	0.0793	-
MLC Transmission Factor	<b>N =</b>	109	43	1
	2.5th	0.0118	0.0081	-
	25th	0.0145	0.0119	-
	<b>50th</b>	<b>0.0151</b>	<b>0.0124</b>	<b>0.0032</b>
	75th	0.0160	0.0130	-
	97.5th	0.0185	0.0258	-
	St. Dev.	0.0016	0.0026	-

**Table C-14.** Beam modeling summary statistics for 10 FFF beams modeled in Eclipse AcurosXB.

Parameter	Percentile	Varian TrueBeam	Varian TrueBeam (HD MLC)
Effective Target Spot Size X [mm]	<b>N =</b>	30	19
	2.5th	-	-
	25th	1.000	1.000
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000
	97.5th	-	-
	St. Dev.	0.274	0.289
Effective Target Spot Size Y [mm]	<b>N =</b>	30	19
	2.5th	-	-
	25th	1.000	1.000
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000
	97.5th	-	-
	St. Dev.	0.381	0.344
Dosimetric Leaf Gap (DLG)	<b>N =</b>	31	18
	2.5th	-	-
	25th	0.1230	0.0383
	<b>50th</b>	<b>0.1460</b>	<b>0.0550</b>
	75th	0.1800	0.0824
	97.5th	-	-
	St. Dev.	0.1706	0.0418
MLC Transmission Factor	<b>N =</b>	31	18
	2.5th	-	-
	25th	0.0149	0.0121
	<b>50th</b>	<b>0.0151</b>	<b>0.0125</b>
	75th	0.0160	0.0130
	97.5th	-	-
	St. Dev.	0.0017	0.0038

**Table C-15.** Beam modeling summary statistics for 10 FFF beams modeled in RayStation. For brevity, 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles are not reported due to limited survey samples.

Parameter	Percentile	Varian TrueBeam	Elekta Agility
<b>Primary Source X Width [cm]</b>	<b>N =</b>	1	8
	25th	-	0.103
	<b>50th</b>	<b>0.060</b>	<b>0.120</b>
	75th	-	0.140
	St. Dev.	-	0.026
<b>Primary Source Y Width [cm]</b>	<b>N =</b>	1	8
	25th	-	0.030
	<b>50th</b>	<b>0.070</b>	<b>0.100</b>
	75th	-	0.100
	St. Dev.	-	0.045
<b>MLC Transmission</b>	<b>N =</b>	1	8
	25th	-	0.0018
	<b>50th</b>	<b>0.0137</b>	<b>0.0050</b>
	75th	-	0.0084
	St. Dev.	-	0.0032
<b>Tongue &amp; Groove [cm]</b>	<b>N =</b>	1	8
	25th	-	0.013
	<b>50th</b>	<b>0.050</b>	<b>0.050</b>
	75th	-	0.050
	St. Dev.	-	0.023
<b>Leaf Tip Width [cm]</b>	<b>N =</b>	1	8
	25th	-	0.033
	<b>50th</b>	<b>0.200</b>	<b>0.150</b>
	75th	-	0.425
	St. Dev.	-	0.197
<b>MLC Position Offset [cm]</b>	<b>N =</b>	1	4
	25th	-	0.000
	<b>50th</b>	<b>0.030</b>	<b>0.000</b>
	75th	-	0.000
	St. Dev.	-	0.000
<b>MLC Position Gain</b>	<b>N =</b>	1	4
	25th	-	0.0000
	<b>50th</b>	<b>0.0000</b>	<b>0.0000</b>
	75th	-	0.0000
	St. Dev.	-	0.0000
<b>MLC Position Curvature [1/cm]</b>	<b>N =</b>	1	4
	25th	-	0.0000
	<b>50th</b>	<b>0.0000</b>	<b>0.0002</b>
	75th	-	0.0004
	St. Dev.	-	0.0002



**Table C-16.** Beam modeling summary statistics for 10 FFF beams modeled in Pinnacle. For brevity, 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles are not reported due to limited survey samples.

Parameter	Percentile	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility
Effective Source Size X [cm]	<b>N =</b>	6	5	6
	25th	0.0697	0.0875	0.0500
	<b>50th</b>	<b>0.1293</b>	<b>0.1206</b>	<b>0.0500</b>
	75th	0.1665	0.1933	0.0958
	St. Dev.	0.0692	0.0697	0.0605
Effective Source Size Y [cm]	<b>N =</b>	6	5	6
	25th	0.0464	0.0593	0.0500
	<b>50th</b>	<b>0.0974</b>	<b>0.0608</b>	<b>0.0500</b>
	75th	0.1450	0.1562	0.1153
	St. Dev.	0.0761	0.0848	0.0601
MLC Transmission	<b>N =</b>	6	5	5
	25th	0.0097	0.0121	0.0026
	<b>50th</b>	<b>0.0124</b>	<b>0.0139</b>	<b>0.0036</b>
	75th	0.0217	0.0148	0.0036
	St. Dev.	0.0079	0.0015	0.0009
Rounded Leaf Tip Radius [cm]	<b>N =</b>	6	5	6
	25th	8.0	16.0	17.0
	<b>50th</b>	<b>8.0</b>	<b>16.0</b>	<b>17.0</b>
	75th	11.5	21.0	17.0
	St. Dev.	3.2	4.5	0.0
Tongue and Groove Width [cm]	<b>N =</b>	6	5	6
	25th	0.031	0.005	0.075
	<b>50th</b>	<b>0.070</b>	<b>0.040</b>	<b>0.100</b>
	75th	0.100	0.040	0.125
	St. Dev.	0.041	0.019	0.063
Additional Tongue and Groove Transmission	<b>N =</b>	6	5	6
	25th	0.004	0.005	0.000
	<b>50th</b>	<b>0.010</b>	<b>0.005</b>	<b>0.002</b>
	75th	0.015	0.019	0.002
	St. Dev.	0.010	0.011	0.001
Flattening Filter Gaussian Height	<b>N =</b>	6	5	6
	25th	0.0118	0.0150	0.0200
	<b>50th</b>	<b>0.0250</b>	<b>0.0216</b>	<b>0.0200</b>
	75th	0.0269	0.1168	0.0265
	St. Dev.	0.0103	0.0799	0.0041
Flattening Filter Gaussian Width	<b>N =</b>	6	5	6
	25th	1.8719	1.1390	1.3850
	<b>50th</b>	<b>2.3631</b>	<b>3.0000</b>	<b>1.3850</b>
	75th	2.8500	4.5826	1.6743
	St. Dev.	0.6870	1.7261	0.4243

**Table C-17.** Beam modeling summary statistics for 15 MV beams modeled in Eclipse AAA.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility	Elekta MLCi
Effective Target Spot Size X [mm]	N =	69	3	16	23	1	1
	2.5th	0.000	-	0.000	-	-	-
	25th	0.000	-	0.000	0.000	-	-
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>1.000</b>
	75th	0.000	-	0.000	0.000	-	-
	97.5th	1.313	-	2.369	-	-	-
	St. Dev.	0.363	0.000	3.827	0.229	-	-
Effective Target Spot Size Y [mm]	N =	69	3	116	23	1	1
	2.5th	0.000	-	0.000	-	-	-
	25th	0.000	-	0.000	0.000	-	-
	50th	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
	75th	0.000	-	0.000	0.000	-	-
	97.5th	1.125	-	2.323	-	-	-
	St. Dev.	0.261	0.000	3.826	0.209	-	-
Dosimetric Leaf Gap [cm]	N =	73	4	117	24	1	0
	2.5th	0.0618	-	0.0960	-	-	-
	25th	0.1650	0.0538	0.1263	0.0463	-	-
	50th	<b>0.1800</b>	<b>0.0675</b>	<b>0.1530</b>	<b>0.0526</b>	<b>0.0900</b>	-
	75th	0.2000	0.0925	0.1800	0.0948	-	-
	97.5th	0.2650	-	0.3270	-	-	-
	St. Dev.	0.0403	0.0210	0.1431	0.0429	-	-
MLC Transmission Factor	N =	73	4	117	24	1	1
	2.5th	0.0130	-	0.0140	-	-	-
	25th	0.0156	0.0108	0.0165	0.0131	-	-
	50th	<b>0.0170</b>	<b>0.0130</b>	<b>0.0170</b>	<b>0.0138</b>	<b>0.0042</b>	<b>0.0000</b>
	75th	0.0180	0.0190	0.0174	0.0140	-	-
	97.5th	0.0250	-	0.0250	-	-	-
	St. Dev.	0.0024	0.0047	0.0020	0.0025	-	-

**Table C-18.** Beam modeling summary statistics for 15 MV beams modeled in Eclipse AcurosXB.

Parameter	Percentile	Varian Base	Varian TrueBeam	Varian TrueBeam (HD MLC)
Effective Target Spot Size X [mm]	<b>N =</b>	17	29	11
	2.5th	-	-	-
	25th	1.000	1.000	1.000
	<b>50th</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.500	1.000	1.000
	97.5th	-	-	-
	St. Dev.	0.385	0.041	0.350
Effective Target Spot Size Y [mm]	<b>N =</b>	17	29	11
	2.5th	-	-	-
	25th	0.000	1.000	1.000
	<b>50th</b>	<b>0.500</b>	<b>1.000</b>	<b>1.000</b>
	75th	1.000	1.000	1.000
	97.5th	-	-	-
	St. Dev.	0.483	0.518	0.405
Dosimetric Leaf Gap (DLG)	<b>N =</b>	17	30	11
	2.5th	-	-	-
	25th	0.1610	0.1549	0.0380
	<b>50th</b>	<b>0.1706</b>	<b>0.1900</b>	<b>0.0505</b>
	75th	0.2000	0.2050	0.1100
	97.5th	-	-	-
	St. Dev.	0.0371	0.1241	0.0436
MLC Transmission Factor	<b>N =</b>	17	30	11
	2.5th	-	-	-
	25th	0.0156	0.0160	0.0135
	<b>50th</b>	<b>0.0160</b>	<b>0.0170</b>	<b>0.0140</b>
	75th	0.0180	0.0181	0.0145
	97.5th	-	-	-
	St. Dev.	0.0035	0.0024	0.0014

**Table C-19.** Beam modeling summary statistics for 15 MV beams modeled in RayStation. For brevity, 2.5th and 97.5th percentiles are not reported due to limited survey samples.

Parameter	Percentile	Varian TrueBeam	Elekta Agility	Elekta MLCi
Primary Source X Width [cm]	N =	4	18	5
	25th	0.040	0.095	0.085
	50th	<b>0.070</b>	<b>0.100</b>	<b>0.120</b>
	75th	0.104	0.115	0.145
	St. Dev.	0.036	0.045	0.035
Primary Source Y Width [cm]	N =	4	18	5
	25th	0.040	0.073	0.057
	50th	<b>0.055</b>	<b>0.100</b>	<b>0.100</b>
	75th	0.092	0.120	0.150
	St. Dev.	0.029	0.040	0.061
MLC Transmission	N =	4	18	5
	25th	0.0160	0.0010	0.0075
	50th	<b>0.0190</b>	<b>0.0070</b>	<b>0.0100</b>
	75th	0.0238	0.0070	0.0125
	St. Dev.	0.0041	0.0030	0.0035
Tongue & Groove [cm]	N =	4	18	5
	25th	0.028	0.050	0.025
	50th	<b>0.050</b>	<b>0.050</b>	<b>0.050</b>
	75th	0.050	0.075	0.100
	St. Dev.	0.015	0.025	0.055
Leaf Tip Width [cm]	N =	4	18	5
	25th	0.174	0.200	0.010
	50th	<b>0.250</b>	<b>0.200</b>	<b>0.200</b>
	75th	0.300	0.200	0.250
	St. Dev.	0.069	0.157	0.129
MLC Position Offset [cm]	N =	2	13	3
	25th	-	0.000	-
	50th	<b>0.000</b>	<b>0.100</b>	<b>0.050</b>
	75th	-	0.100	-
	St. Dev.	0.000	0.050	0.050
MLC Position Gain	N =	2	13	3
	25th	-	0.0065	-
	50th	<b>0.0000</b>	<b>0.0750</b>	<b>0.0000</b>
	75th	-	0.0750	-
	St. Dev.	0.0000	0.0361	0.0026
MLC Position Curvature [1/cm]	N =	2	13	3
	25th	-	0.0002	-
	50th	<b>0.0004</b>	<b>0.1000</b>	<b>0.0002</b>
	75th	-	0.1000	-
	St. Dev.	0.0006	0.0518	0.0026

**Table C-20.** Beam modeling summary statistics for 15 MV beams modeled in Pinnacle.

Parameter	Percentile	Varian Base	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Effective Source Size X [cm]	N =	10	9	1	13	2	5
	25th	0.0275	0.0227	-	0.0309	-	0.0806
	50th	<b>0.0411</b>	<b>0.1370</b>	<b>0.0200</b>	<b>0.1425</b>	<b>0.0800</b>	<b>0.1106</b>
	75th	0.0786	0.1663	-	0.1723	-	0.1928
	St. Dev.	0.0261	0.0727	-	0.0851	0.0566	0.0584
Effective Source Size Y [cm]	N =	10	9	1	13	2	5
	25th	0.0242	0.0310	-	0.0450	-	0.0720
	50th	<b>0.0275</b>	<b>0.1014</b>	<b>0.0200</b>	<b>0.1000</b>	<b>0.0746</b>	<b>0.1450</b>
	75th	0.0857	0.1262	-	0.2000	-	0.1491
	St. Dev.	0.0447	0.0466	-	0.0855	0.0489	0.0584
MLC Transmission	N =	9	9	1	10	2	5
	25th	0.0126	0.0144	-	0.0010	-	0.0045
	50th	<b>0.0165</b>	<b>0.0184</b>	<b>0.0150</b>	<b>0.0038</b>	<b>0.0110</b>	<b>0.0081</b>
	75th	0.0165	0.0292	-	0.0062	-	0.0116
	St. Dev.	0.0025	0.0077	-	0.0027	0.0057	0.0037
Rounded Leaf Tip Radius [cm]	N =	10	8	1	13	1	5
	25th	8.0	8.0	-	17.0	-	15.0
	50th	<b>8.0</b>	<b>9.0</b>	<b>16.0</b>	<b>17.0</b>	<b>12.2</b>	<b>15.0</b>
	75th	8.0	10.0	-	17.0	-	16.0
	St. Dev.	0.0	2.7	-	1.4	-	0.9
Tongue and Groove Width [cm]	N =	10	9	1	13	2	5
	25th	0.093	0.040	-	0.080	-	0.020
	50th	<b>0.100</b>	<b>0.040</b>	<b>0.005</b>	<b>0.200</b>	<b>0.045</b>	<b>0.100</b>
	75th	0.100	0.100	-	0.200	-	0.100
	St. Dev.	0.020	0.029	-	0.073	0.021	0.046
Additional Tongue and Groove Transmission	N =	10	9	1	13	2	5
	25th	0.009	0.005	-	0.000	-	0.004
	50th	<b>0.015</b>	<b>0.008</b>	<b>0.003</b>	<b>0.000</b>	<b>0.033</b>	<b>0.010</b>
	75th	0.020	0.010	-	0.001	-	0.010
	St. Dev.	0.009	0.008	-	0.001	0.039	0.004
Flattening Filter Gaussian Height	N =	10	9	1	13	2	5
	25th	0.0548	0.0731	-	0.0789	-	0.0659
	50th	<b>0.0585</b>	<b>0.0828</b>	<b>0.0828</b>	<b>0.0844</b>	<b>0.0410</b>	<b>0.0800</b>
	75th	0.0728	0.0901	-	0.0923	-	0.2625
	St. Dev.	0.0168	0.0124	-	0.0173	0.0472	0.1582
Flattening Filter Gaussian Width	N =	10	9	1	13	2	5
	25th	1.0702	1.2103	-	1.2230	-	0.7991
	50th	<b>1.1405</b>	<b>1.4702</b>	<b>3.0000</b>	<b>1.4249</b>	<b>1.5197</b>	<b>1.1200</b>
	75th	1.3798	1.7133	-	2.2462	-	2.3873
	St. Dev.	0.2008	0.5678	-	0.5668	0.2549	0.8767

**Table C-21.**Beam modeling summary statistics for 18 MV beams modeled in Eclipse AAA.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta MLCi
Effective Target Spot Size X [mm]	<b>N =</b>	72	4	37	7	1
	2.5th	0.000	-	-	-	-
	25th	0.000	0.000	0.000	0.000	-
	<b>50th</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>3.000</b>
	75th	0.000	0.000	0.000	0.000	-
	97.5th	2.988	-	-	-	-
	St. Dev.	1.204	0.000	0.397	0.000	-
Effective Target Spot Size Y [mm]	<b>N =</b>	72	4	37	7	1
	2.5th	0.000	-	-	-	-
	25th	0.000	0.000	0.000	0.000	-
	<b>50th</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>2.500</b>
	75th	0.000	0.000	0.000	0.000	-
	97.5th	2.988	-	-	-	-
	St. Dev.	1.197	0.000	0.229	0.000	-
Dosimetric Leaf Gap [cm]	<b>N =</b>	73	4	37	6	1
	2.5th	-1.5449	-	-	-	-
	25th	0.1684	0.0498	0.1180	0.0433	-
	<b>50th</b>	<b>0.1830</b>	<b>0.0785</b>	<b>0.1600</b>	<b>0.0845</b>	<b>0.5000</b>
	75th	0.2000	0.1335	0.2000	0.1614	-
	97.5th	0.2400	-	-	-	-
	St. Dev.	0.3390	0.0455	0.0420	0.0591	-
MLC Transmission Factor	<b>N =</b>	73	4	37	6	1
	2.5th	0.0119	-	-	-	-
	25th	0.0150	0.0126	0.0150	0.0128	-
	<b>50th</b>	<b>0.0170</b>	<b>0.0127</b>	<b>0.0158</b>	<b>0.0132</b>	<b>0.0160</b>
	75th	0.0195	0.0167	0.0170	0.0158	-
	97.5th	0.0259	-	-	-	-
	St. Dev.	0.0032	0.0027	0.0037	0.0016	-

**Table C-22.** Beam modeling summary statistics for 18 MV beams modeled in Eclipse AcurosXB.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)
Effective Target Spot Size X [mm]	N =	10	0	8	1
	2.5th	-	-	-	-
	25th	1.000	-	1.000	-
	50th	<b>1.000</b>	-	<b>1.000</b>	<b>1.000</b>
	75th	1.000	-	1.000	-
	97.5th	-	-	-	-
	St. Dev.	0.158	-	0.200	-
Effective Target Spot Size Y [mm]	N =	10	0	8	1
	2.5th	-	-	-	-
	25th	1.000	-	0.850	-
	50th	<b>1.000</b>	-	<b>1.000</b>	<b>1.000</b>
	75th	1.000	-	1.000	-
	97.5th	-	-	-	-
	St. Dev.	0.316	-	0.351	-
Dosimetric Leaf Gap (DLG)	N =	11	0	8	1
	2.5th	-	-	-	-
	25th	0.1480	-	0.1242	-
	50th	<b>0.1530</b>	-	<b>0.1700</b>	<b>0.0640</b>
	75th	0.1800	-	0.2075	-
	97.5th	-	-	-	-
	St. Dev.	0.0514	-	0.0426	-
MLC Transmission Factor	N =	11	0	8	4
	2.5th	-	-	-	-
	25th	0.0131	-	0.0146	-
	50th	<b>0.0167</b>	-	<b>0.0158</b>	<b>0.0130</b>
	75th	0.0170	-	0.0177	-
	97.5th	-	-	-	-
	St. Dev.	0.0052	-	0.0015	-

**Table C-23.** Beam modeling summary statistics for 18 MV beams modeled in RayStation. For brevity, 2.5th and 97.5th percentiles are not reported due to limited survey samples.

Parameter	Percentile	Varian Base	Varian Base (HD MLC)	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility
Primary Source X Width [cm]	N =	3	0	3	0	5
	25th	-	-	-	-	0.100
	50th	<b>0.040</b>	-	<b>0.090</b>	-	<b>0.100</b>
	75th	-	-	-	-	0.145
	St. Dev.	0.052	-	0.012	-	0.027
Primary Source Y Width [cm]	N =	3	0	3	0	5
	25th	-	-	-	-	0.100
	50th	<b>0.060</b>	-	<b>0.100</b>	-	<b>0.100</b>
	75th	-	-	-	-	0.130
	St. Dev.	0.033	-	0.016	-	0.022
MLC Transmission	N =	3	0	3	0	5
	25th	-	-	-	-	0.0055
	50th	<b>0.0180</b>	-	<b>0.0150</b>	-	<b>0.0100</b>
	75th	-	-	-	-	0.0258
	St. Dev.	0.0044	-	0.0058	-	0.0133
Tongue & Groove [cm]	N =	3	0	3	0	5
	25th	-	-	-	-	0.050
	50th	<b>0.050</b>	-	<b>0.050</b>	-	<b>0.050</b>
	75th	-	-	-	-	0.075
	St. Dev.	0.017	-	0.000	-	0.022
Leaf Tip Width [cm]	N =	3	0	3	0	5
	25th	-	-	-	-	0.010
	50th	<b>0.200</b>	-	<b>0.200</b>	-	<b>0.010</b>
	75th	-	-	-	-	0.200
	St. Dev.	0.070	-	0.058	-	0.104
MLC Position Offset [cm]	N =	2	0	3	0	5
	25th	-	-	-	-	0.009
	50th	<b>0.060</b>	-	<b>0.054</b>	-	<b>0.029</b>
	75th	-	-	-	-	0.030
	St. Dev.	0.085	-	0.031	-	0.012
MLC Position Gain	N =	2	0	3	0	5
	25th	-	-	-	-	0.0012
	50th	<b>0.0000</b>	-	<b>0.0031</b>	-	<b>0.0012</b>
	75th	-	-	-	-	0.0063
	St. Dev.	0.0000	-	0.0018	-	0.0032
MLC Position Curvature [1/cm]	N =	2	0	3	0	5
	25th	-	-	-	-	-0.0003
	50th	<b>0.0004</b>	-	<b>0.0007</b>	-	<b>0.0001</b>
	75th	-	-	-	-	0.0001
	St. Dev.	0.0006	-	0.0000	-	0.0003



**Table C-24.**Beam modeling summary statistics for 18 MV beams modeled in Pinnacle. For brevity, 2.5th and 97.5th percentiles are not reported due to limited survey samples.

Parameter	Percentile	Varian Base	Varian TrueBeam	Varian TrueBeam (HD MLC)	Elekta Agility	Elekta Beam Modulator	Elekta MLCi
Effective Source Size X [cm]	<b>N =</b>	11	2	1	16	2	11
	25th	0.0100	-	-	0.0118	-	0.0100
	<b>50th</b>	<b>0.1094</b>	<b>0.1196</b>	<b>0.0400</b>	<b>0.0600</b>	<b>0.0225</b>	<b>0.0531</b>
	75th	0.1368	-	-	0.0858	-	0.0856
	St. Dev.	0.0688	0.0276	-	0.0491	0.0000	0.0422
Effective Source Size Y [cm]	<b>N =</b>	11	2	1	16	2	11
	25th	0.0100	-	-	0.0113	-	0.0100
	<b>50th</b>	<b>0.0583</b>	<b>0.1180</b>	<b>0.0400</b>	<b>0.0388</b>	<b>0.0231</b>	<b>0.0200</b>
	75th	0.0975	-	-	0.0800	-	0.1456
	St. Dev.	0.0456	0.0255	-	0.0337	0.0000	0.0663
MLC Transmission	<b>N =</b>	6	2	1	15	2	11
	25th	0.0150	-	-	0.0020	-	0.0040
	<b>50th</b>	<b>0.0184</b>	<b>0.0278</b>	<b>0.0170</b>	<b>0.0035</b>	<b>0.0040</b>	<b>0.0100</b>
	75th	0.0241	-	-	0.0043	-	0.0139
	St. Dev.	0.0057	0.0031	-	0.0014	0.0000	0.0099
Rounded Leaf Tip Radius [cm]	<b>N =</b>	10	2	1	16	2	11
	25th	7.5	-	-	17.0	-	15.0
	<b>50th</b>	<b>8.0</b>	<b>8.0</b>	<b>26.0</b>	<b>17.0</b>	<b>18.0</b>	<b>15.0</b>
	75th	8.0	-	-	17.0	-	15.0
	St. Dev.	1.6	0.0	-	2.0	0.0	0.0
Tongue and Groove Width [cm]	<b>N =</b>	11	2	1	16	2	11
	25th	0.040	-	-	0.019	-	0.063
	<b>50th</b>	<b>0.100</b>	<b>0.023</b>	<b>0.040</b>	<b>0.080</b>	<b>0.120</b>	<b>0.100</b>
	75th	0.120	-	-	0.200	-	0.150
	St. Dev.	0.040	0.025	-	0.083	0.000	0.059
Additional Tongue and Groove Transmission	<b>N =</b>	11	2	1	16	2	11
	25th	0.005	-	-	0.000	-	0.010
	<b>50th</b>	<b>0.008</b>	<b>0.008</b>	<b>0.007</b>	<b>0.002</b>	<b>0.007</b>	<b>0.010</b>
	75th	0.010	-	-	0.046	-	0.030
	St. Dev.	0.003	0.004	-	0.026	0.000	0.016
Flattening Filter Gaussian Height	<b>N =</b>	11	2	1	16	2	11
	25th	0.0700	-	-	0.0750	-	0.0703
	<b>50th</b>	<b>0.0712</b>	<b>0.0816</b>	<b>0.0600</b>	<b>0.0800</b>	<b>0.0750</b>	<b>0.0802</b>
	75th	0.0867	-	-	0.0844	-	0.0872
	St. Dev.	0.0091	0.0022	-	0.0093	0.0000	0.0099
Flattening Filter Gaussian Width	<b>N =</b>	11	2	1	16	2	11
	25th	0.7986	-	-	1.3590	-	1.0800
	<b>50th</b>	<b>1.0000</b>	<b>1.2308</b>	<b>0.7000</b>	<b>1.7130</b>	<b>1.5000</b>	<b>1.3271</b>
	75th	1.0746	-	-	2.2000	-	1.5476
	St. Dev.	0.2401	0.3806	-	0.3809	0.0000	0.3279

## **Appendix D: Supplement to Chapter 5**

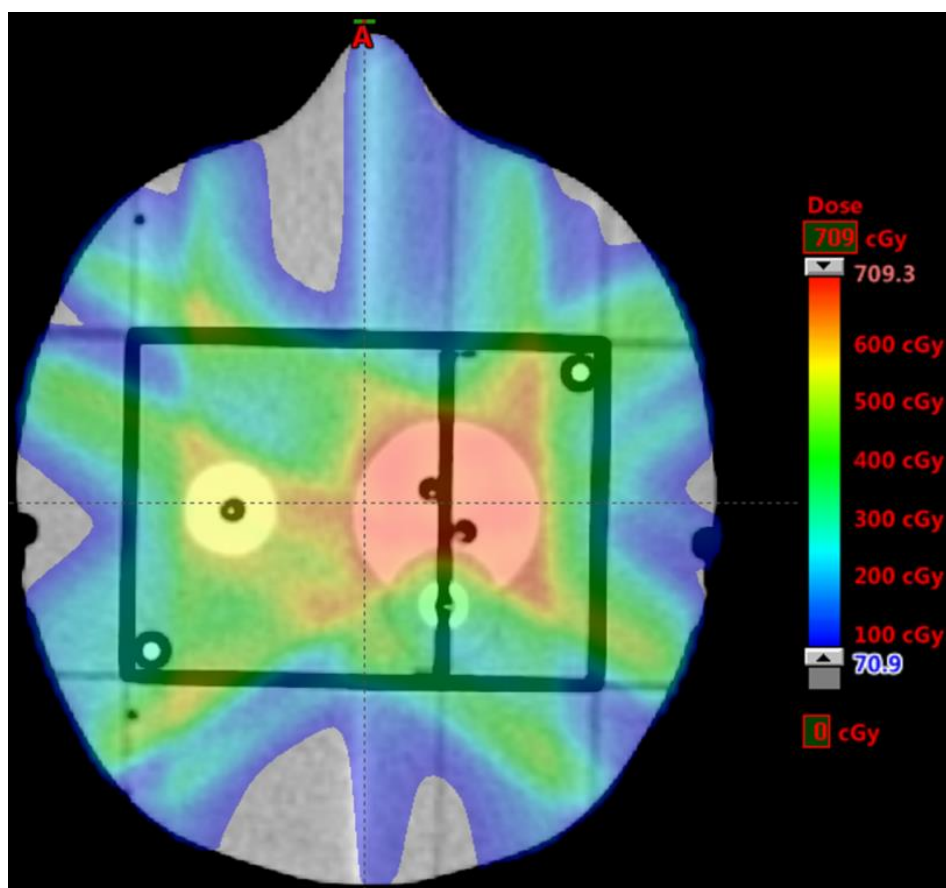
This appendix serves as the supplement to  
Chapter 5: Sensitivity Analyses of Common Beam Modeling Parameters

The following are supplementary information that further inform the work and conclusions of Chapter 5. Here exist data related to the plans generated for use in the sensitivity analyses, the values used for beam modeling manipulations, and additional figures describing the behavior of dose changes in each of the test plans. Tables D-1 through D-5 and Figures D-1 through D-10 describe the general plan information as it relates to IROC dose constraints. Tables D-6 and D-7 demonstrate the agreement of the baseline plan models with phantom performance. Tables D-8 through D-11 show the dosimetric agreement with reference data for the original, un-tuned beam models and final models in Eclipse and RayStation. Tables D-12 through D-14 show relevant information for how the PDD, off-axis factor, and jaw-defined small field output factor error statistics were developed. Finally, Figures D-11 through D-19 show how individual parameters influence dose in Eclipse and RayStation, while Figure D-20 shows supplemental interplay relationships not explicitly demonstrated in Chapter 5.

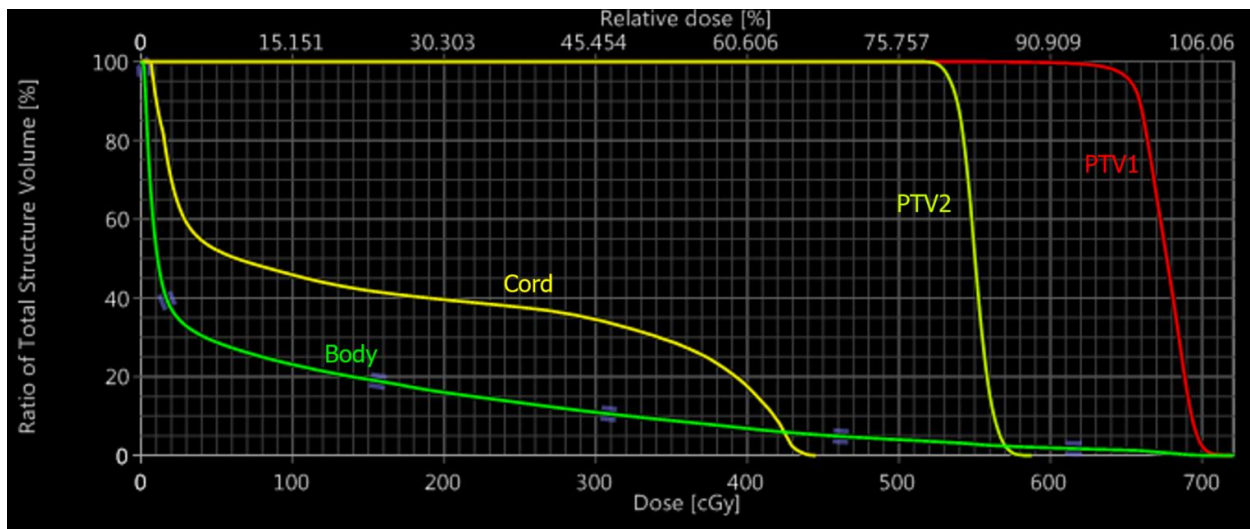
DICOM files for the phantom plans, as well as the water phantom for beam modeling validation, are archived at: J:\Everyone\Mallory\Dissertation\Appendix D.

**Table D-1.** Plan information for the IMRT 5-field plan used in determining the relative sensitivity of beam modeling parameters.

	MU	3741
	Max Dose	721 cGy
<b>PTV_66</b>	Mean Dose	676 cGy
	D99	630 cGy
	D95	653 cGy
<b>PTV_54</b>	Mean Dose	554 cGy
	D99	526 cGy
	D95	534 cGy
<b>Cord</b>	D0.2cc	440 cGy



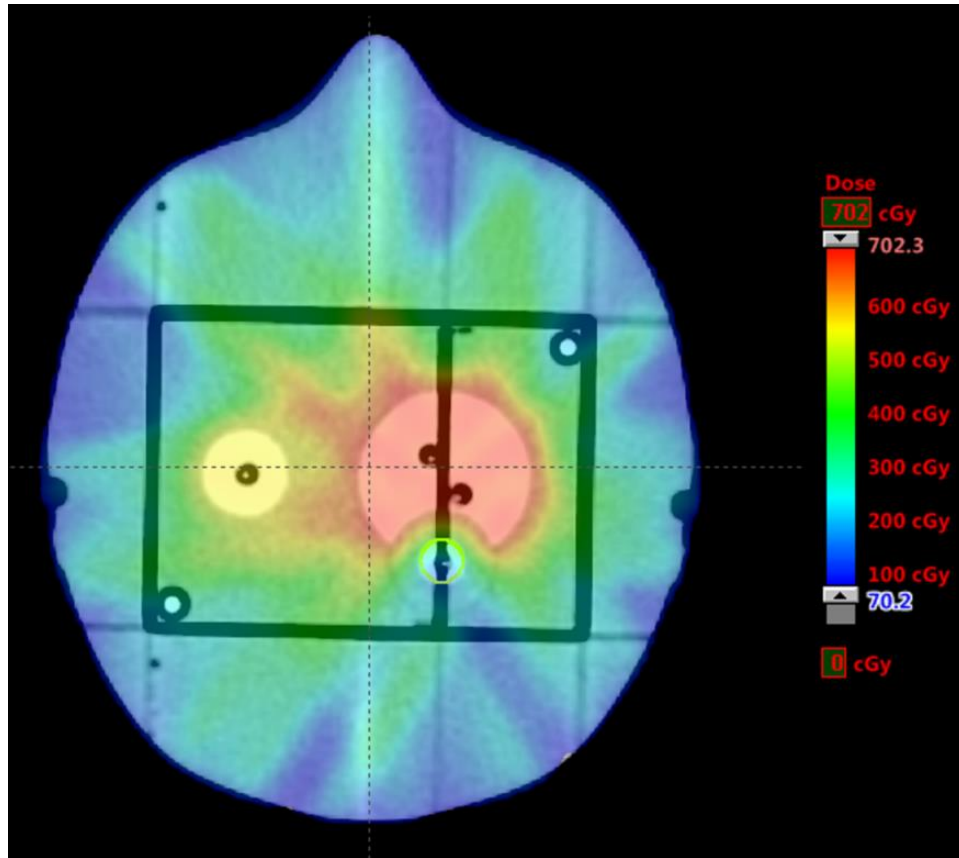
**Figure D-1.** Axial slice of the IROC H&N phantom with the IMRT 5-field plan.



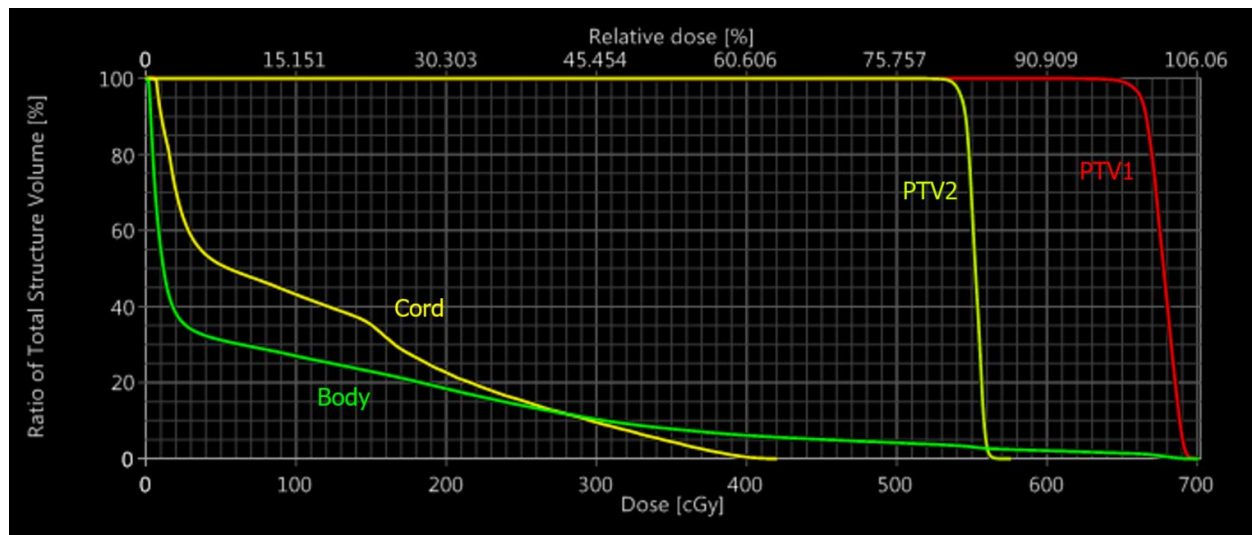
**Figure D-2.** DVH information for the IMRT 5-field plan

**Table D-2.** Plan information for the IMRT 7-field plan used in determining the relative sensitivity of beam modeling parameters.

	MU	2470
	Max Dose	702 cGy
<b>PTV_66</b>	Mean Dose	677 cGy
	D99	651 cGy
	D95	662 cGy
<b>PTV_54</b>	Mean Dose	552 cGy
	D99	537 cGy
	D95	542 cGy
<b>Cord</b>	D0.2cc	405 cGy



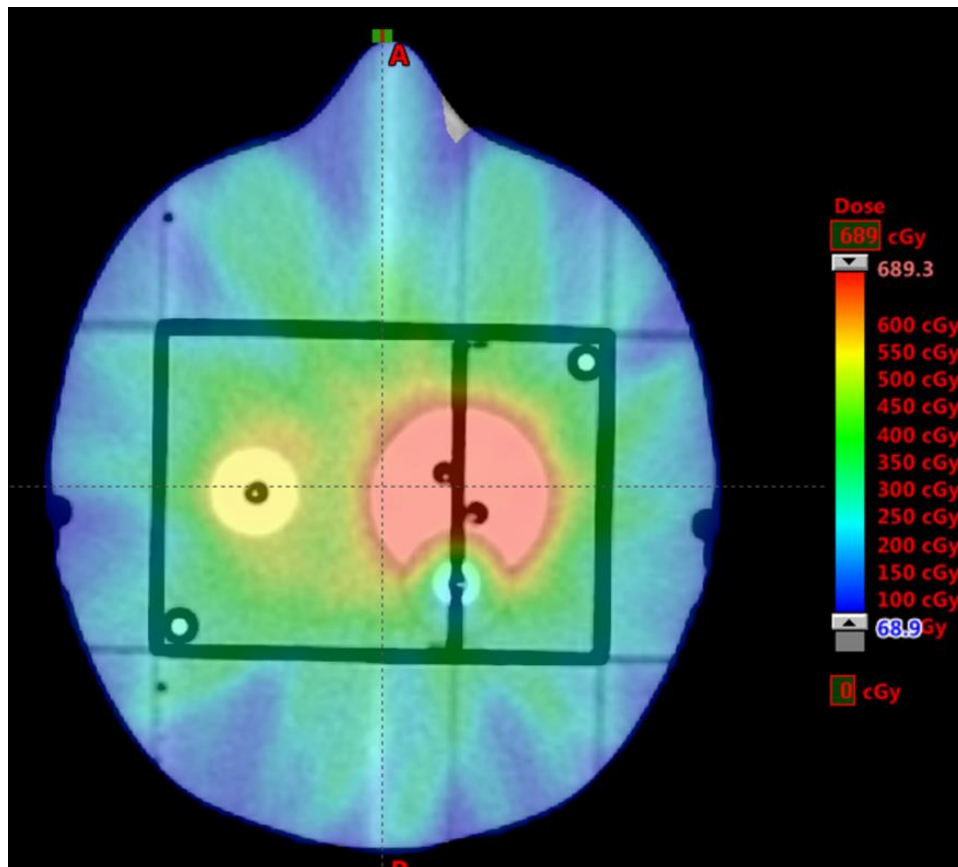
**Figure D-3.** Axial slice of the IROC H&N phantom with the IMRT 7-field plan.



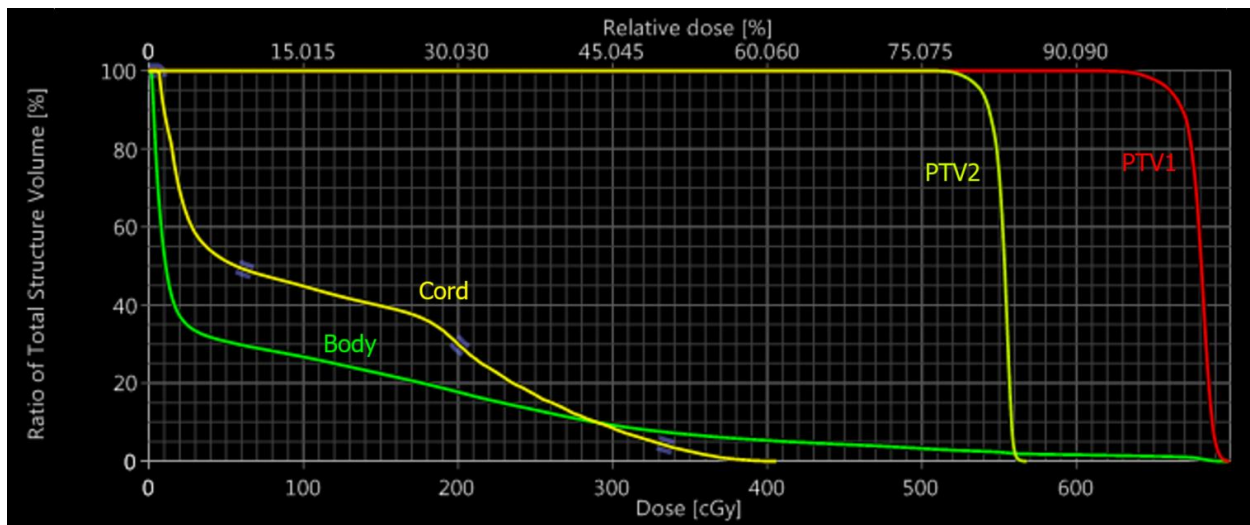
**Figure D-4.** DVH information for the IMRT 7-field plan.

**Table D-3.** Plan information for the IMRT 9-field plan used in determining the relative sensitivity of beam modeling parameters.

MU		2729
Max Dose		699 cGy
<b>PTV_66</b>	Mean Dose	679 cGy
	D99	639 cGy
	D95	660 cGy
<b>PTV_54</b>	Mean Dose	552 cGy
	D99	519 cGy
	D95	538 cGy
<b>Cord</b>	D0.2cc	403 cGy



**Figure D-5.** Axial slice of the IROC H&N phantom with the IMRT 9-field plan.

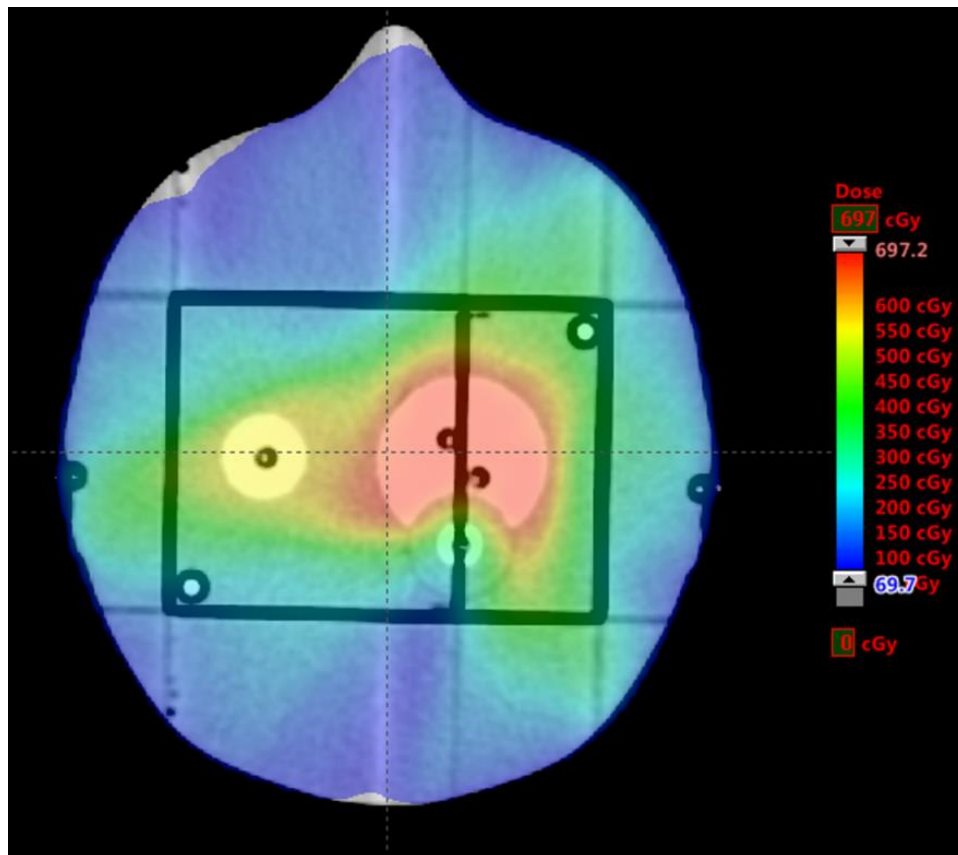


**Figure D-6.** DVH information for the IMRT 9-field plan.

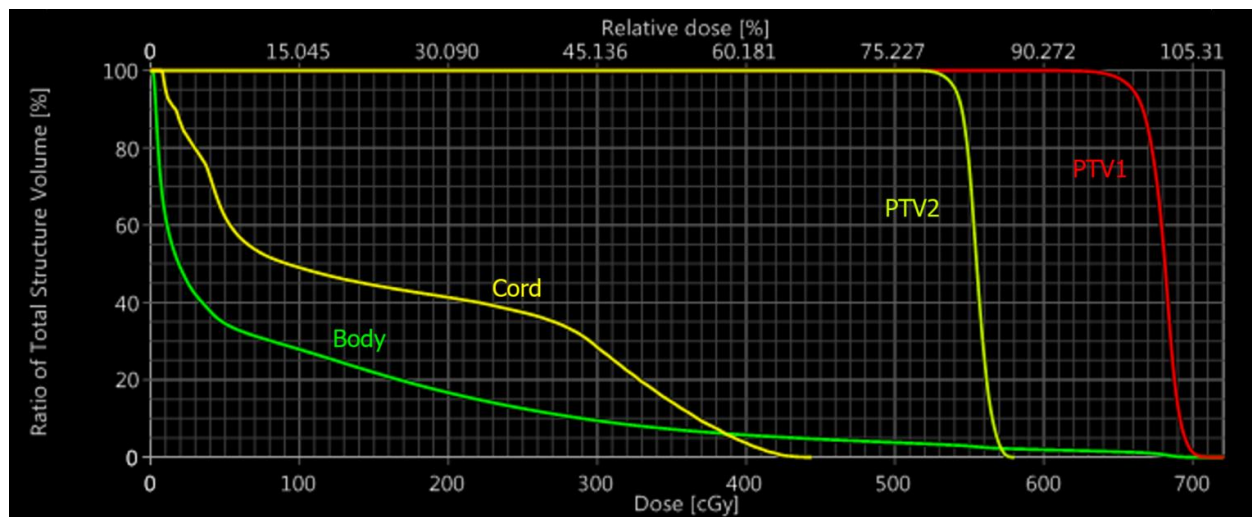
**Table D-4.** Plan information for the VMAT 1-arc plan used in determining the relative sensitivity of beam modeling parameters.

MU		1990
	Max Dose	721 cGy
<b>PTV_66</b>	Mean Dose	680 cGy
	D99	643 cGy
	D95	660 cGy
<b>PTV_54</b>	Mean Dose	555 cGy
	D99	530 cGy
	D95	540 cGy
<b>Cord</b>	D0.2cc	431 cGy





**Figure D-7.** Axial slice of the IROC H&N phantom with the VMAT 1-arc plan.

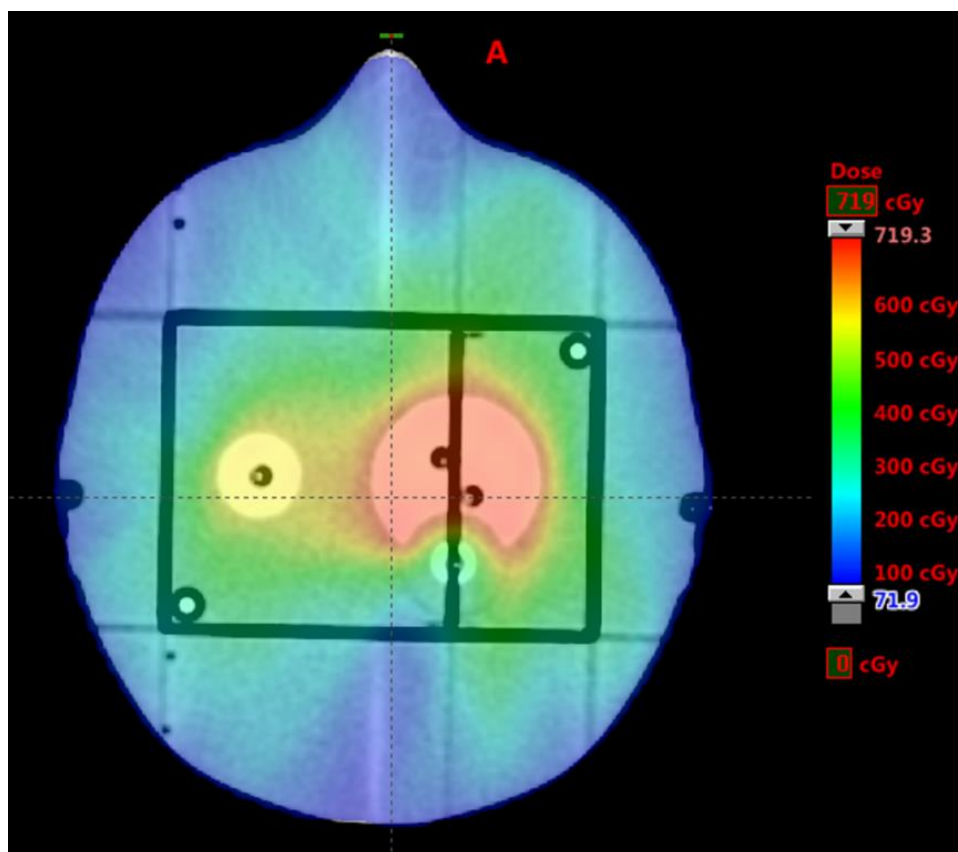


**Figure D-8.** DVH information for the VMAT 1-arc plan.

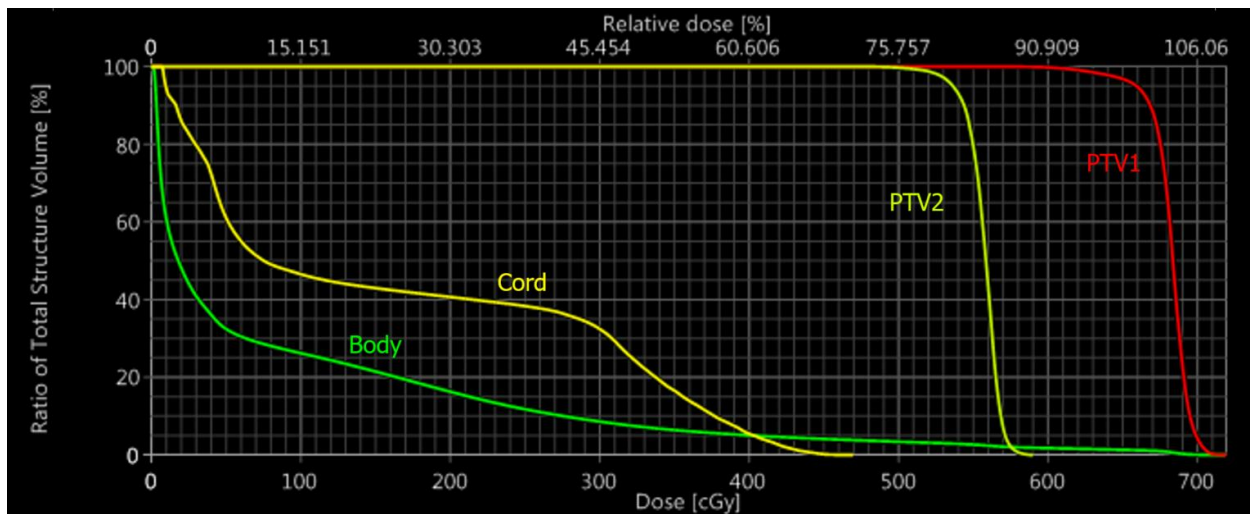


**Table D-5.** Plan information for the VMAT 2-arc plan used in determining the relative sensitivity of beam modeling parameters.

	MU	2130
	Max Dose	719 cGy
<b>PTV_66</b>	Mean Dose	682 cGy
	D99	622 cGy
	D95	659 cGy
<b>PTV_54</b>	Mean Dose	557 cGy
	D99	516 cGy
	D95	536 cGy
<b>Cord</b>	D0.2cc	449 cGy



**Figure D-9.** Axial slice of the IROC H&N phantom with the VMAT 2-arc plan.



**Figure D-10.** DVH information for the VMAT 2-arc plan.

**Table D-6.** Verification of the baseline beam model in Eclipse through agreement between TPS-calculated dose and measured dose for the VMAT 2-arc and IMRT 9-field plans. Note that all but one point express agreement within 2%.

TLD	VMAT 2-Arc Plan			IMRT 9-Field Plan		
	TPS (cGy)	TLD (cGy)	Ratio	TPS (cGy)	TLD (cGy)	Ratio
PTV1 sup. ant.	676.5	683	1.01	671.4	670	1.00
PTV1 inf. ant.	682.7	675	0.99	673.3	672	1.00
PTV1 sup. post.	687.8	699	1.02	680.8	672	0.99
PTV1 inf. post.	685.3	680	0.99	682.7	683	1.00
PTV2 sup.	559.1	550	0.98	550.7	554	1.01
PTV2 inf.	552.0	529	0.96	552.2	556	1.01

**Table D-7.** Verification of the baseline beam model in RayStation through agreement between TPS-calculated dose and measured dose for the VMAT 2-arc and IMRT 9-field plans. Note that all but one point express agreement within 2%.

TLD	VMAT 2-Arc Plan			IMRT 9-Field Plan		
	TPS (cGy)	TLD (cGy)	Ratio	TPS (cGy)	TLD (cGy)	Ratio
PTV1 sup. ant.	681	683	1.00	675	670	0.99
PTV1 inf. ant.	683	675	0.99	675	672	1.00
PTV1 sup. post.	692	699	1.01	681	672	0.99
PTV1 inf. post.	694	680	0.98	684	683	1.00
PTV2 sup.	560	550	0.98	556	554	1.00
PTV2 inf.	556	529	0.95	556	556	1.00

**Table D-8.** Comparison of dosimetric characteristics for the un-tuned Eclipse Varian Base 6X model.

Parameter	Reference Value (Median)	Model Value	Difference
<b>Non-Dosimetric Parameters</b>			
SS X	0	0	0.00%
SS Y	0	0	0.00%
MLC Transmission	0.016	0.015	-6.25%
DLG	0.17	0.2	17.65%
<b>PDD Factors</b>			
PDD 10x10 (norm dmax) - 5 cm	0.861	0.861	0.00%
PDD 10x10 (norm dmax) - 10 cm	0.664	0.663	-0.15%
PDD 10x10 (norm dmax) - 15 cm	0.504	0.506	0.40%
PDD 10x10 (norm dmax) - 20 cm	0.381	0.384	0.79%
PDD 10x10 (norm 10) - 1.5 cm	1.507	1.509	0.13%
PDD 10x10 (norm 10) - 5 cm	1.297	1.3	0.23%
PDD 10x10 (norm 10) - 15 cm	0.759	0.763	0.53%
PDD 10x10 (norm 10) - 20 cm	0.574	0.579	0.87%
<b>Output Factors (norm 10x10@dmax)</b>			
6x6	0.957	0.96	0.31%
15x15	1.032	1.03	-0.19%
20x20	1.054	1.05	-0.38%
30x30	1.08	1.077	-0.28%
<b>Off Axis Factors</b>			
5 cm	1.029	1.024	-0.49%
10 cm	1.042	1.05	0.77%
15 cm	1.055	1.059	0.38%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>			
2x2	0.807	0.822	1.86%
3x3	0.852	0.867	1.76%
4x4	0.888	0.894	0.68%
6x6	0.938	0.94	0.21%
<b>SBRT Output Factors</b>			
2x2	0.778	0.788	1.29%
3x3	0.82	0.831	1.34%
4x4	0.855	0.865	1.17%
6x6	0.915	0.917	0.22%
Avg Err Total Deviation			0.50%
			14.417%

**Table D-9.** Dosimetric characteristics for the final tuned Eclipse Varian Base 6X model.

Parameter	Reference Value (Median)	Model Value	Difference
<b>Non-Dosimetric Parameters</b>			
SS X	0	0	0.00%
SS Y	0	0	0.00%
MLC Transmission	0.0158	0.0158	0.00%
DLG	0.17	0.17	0.00%
<b>PDD Factors</b>			
PDD 10x10 (norm dmax) - 5 cm	0.861	0.8618	0.09%
PDD 10x10 (norm dmax) - 10 cm	0.664	0.663339	-0.10%
PDD 10x10 (norm dmax) - 15 cm	0.504	0.50609	0.41%
PDD 10x10 (norm dmax) - 20 cm	0.381	0.383986	0.78%
PDD 10x10 (norm 10) - 1.5 cm	1.507	1.5075	0.03%
PDD 10x10 (norm 10) - 5 cm	1.297	1.299	0.15%
PDD 10x10 (norm 10) - 15 cm	0.759	0.763	0.53%
PDD 10x10 (norm 10) - 20 cm	0.574	0.579	0.87%
<b>Output Factors (norm 10x10@dmax)</b>			
6x6	0.957	0.955	-0.21%
15x15	1.032	1.032	0.00%
20x20	1.054	1.053	-0.09%
30x30	1.08	1.078	-0.19%
<b>Off Axis Factors</b>			
5 cm	1.029	1.025	-0.39%
10 cm	1.042	1.052	0.96%
15 cm	1.055	1.062	0.66%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>			
2x2	0.807	0.827	2.48%
3x3	0.852	0.868	1.88%
4x4	0.888	0.897	1.01%
6x6	0.938	0.941	0.32%
<b>SBRT Output Factors</b>			
2x2	0.778	0.778	0.00%
3x3	0.82	0.814	-0.73%
4x4	0.855	0.851	-0.47%
6x6	0.915	0.914	-0.11%
<b>AVG ERROR</b>			<b>0.344%</b>
<b>Total</b>			
<b>Deviation</b>			<b>12.476%</b>

**Table D-10.** Dosimetric characteristics for the un-tuned RayStation Varian Base 6X model.

Parameter	Reference Value (Median)	Model Value	Difference
<b>Non-Dosimetric Parameters</b>			
SS X	0.057	0.086	50.88%
SS Y	0.07	0.087	24.29%
MLC Transmission	0.018	0.0146	-18.89%
Tongue and Groove	0.04	0.02	-50.00%
Leaf Tip Width	0.32	0.165	-48.44%
MLC Offset	0.04	0.075	87.50%
MLC Gain	0.0015	0.003	100.00%
MLC Curvature	0	0	0.00%
<b>PDD Factors</b>			
PDD 10x10 (norm dmax) - 5 cm	0.861	0.868	0.81%
PDD 10x10 (norm dmax) - 10 cm	0.664	0.673	1.36%
PDD 10x10 (norm dmax) - 15 cm	0.504	0.511	1.39%
PDD 10x10 (norm dmax) - 20 cm	0.381	0.385	1.05%
PDD 10x10 (norm 10) - 1.5 cm	1.507	1.487	-1.33%
PDD 10x10 (norm 10) - 5 cm	1.297	1.29	-0.54%
PDD 10x10 (norm 10) - 15 cm	0.759	0.759	0.00%
PDD 10x10 (norm 10) - 20 cm	0.574	0.572	-0.35%
<b>Output Factors (norm 10x10@dmax)</b>			
6x6	0.957	0.961	0.42%
15x15	1.032	1.033	0.10%
20x20	1.054	1.055	0.09%
30x30	1.08	1.081	0.09%
<b>Off Axis Factors</b>			
5 cm	1.029	1.034	0.49%
10 cm	1.042	1.048	0.58%
15 cm	1.055	1.059	0.38%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>			
2x2	0.807	0.828	2.60%
3x3	0.852	0.865	1.53%
4x4	0.888	0.896	0.90%
6x6	0.938	0.945	0.75%
<b>SBRT Output Factors</b>			
2x2	0.778	0.793	1.93%
3x3	0.82	0.826	0.73%
4x4	0.855	0.863	0.94%
6x6	0.915	0.92	0.55%
		Avg Err	0.63%
		Total Deviation	18.883%

**Table D-11.** Dosimetric characteristics for the final tuned RayStation Varian Base 6X model.

Parameter	Reference Value (Median)	Model Value	Difference
<b>Non-Dosimetric Parameters</b>			
SS X	0.057	0.057	0.00%
SS Y	0.07	0.070	0.00%
MLC Transmission	0.018	0.018	0.00%
Tongue and Groove	0.04	0.040	0.00%
Leaf Tip Width	0.32	0.320	0.00%
MLC Offset	0.04	0.040	0.00%
MLC Gain	0.0015	0.002	0.00%
MLC Curvature	0	0.000	0.00%
<b>PDD Factors</b>			
PDD 10x10 (norm dmax) - 5 cm	0.861	0.868	0.81%
PDD 10x10 (norm dmax) - 10 cm	0.664	0.670	0.90%
PDD 10x10 (norm dmax) - 15 cm	0.504	0.508	0.79%
PDD 10x10 (norm dmax) - 20 cm	0.381	0.381	0.00%
PDD 10x10 (norm 10) - 1.5 cm	1.507	1.493	-0.93%
PDD 10x10 (norm 10) - 5 cm	1.297	1.296	-0.08%
PDD 10x10 (norm 10) - 15 cm	0.759	0.758	-0.13%
PDD 10x10 (norm 10) - 20 cm	0.574	0.569	-0.87%
<b>Output Factors (norm 10x10@dmax)</b>			
6x6	0.957	0.957	0.00%
15x15	1.032	1.032	0.00%
20x20	1.054	1.054	0.00%
30x30	1.08	1.080	0.00%
<b>Off Axis Factors</b>			
5 cm	1.029	1.029	0.00%
10 cm	1.042	1.047	0.48%
15 cm	1.055	1.058	0.28%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>			
2x2	0.807	0.821	1.73%
3x3	0.852	0.861	1.06%
4x4	0.888	0.893	0.56%
6x6	0.938	0.942	0.43%
<b>SBRT Output Factors</b>			
2x2	0.778	0.778	0.00%
3x3	0.82	0.821	0.12%
4x4	0.855	0.855	0.00%
6x6	0.915	0.915	0.00%
		Avg Err	0.22%
		Total	
		Deviation	9.186%

**Table D-12.** Percentile information for PDD curve measurement errors as reported from IROC Houston site visit data (normalized to  $d_{\max}$ ). All errors introduced to the curves were such that the error at 20 cm depth was approximately as reported below.

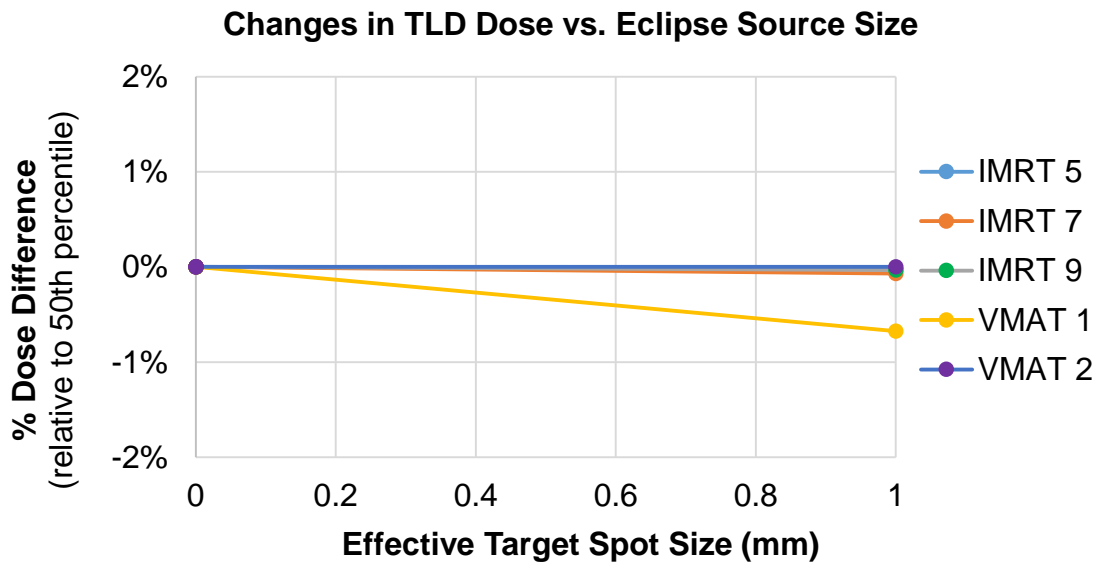
<b>N</b>	Valid	363	
	Missing	0	
<b>Mean</b>		1.00740	
<b>Median</b>		1.01000	
<b>Std. Deviation</b>		.049210	
<b>Range</b>		.945	<b>Change introduced</b>
<b>Percentiles</b>	2.5	.98700	<b>(-2.3%)</b>
	25	1.00300	<b>(-0.7%)</b>
	50	1.01000	<b>(+0%)</b>
	75	1.01600	<b>(+0.6%)</b>
	97.5	1.03490	<b>(+2.5%)</b>

**Table D-13.** Percentile information for jaw-defined small field output factor measurement errors as reported from IROC Houston site visit data (normalized to  $d_{10}$  for a 10x10 cm field). All errors introduced to the curves were such that the error was approximately as reported below.

SBRT Factors		2x2 cm	3x3 cm	4x4 cm	6x6 cm
<b>N</b>	Valid	47	57	57	58
	Missing	0	0	0	0
<b>Mean</b>		1.00117	1.00807	1.00684	1.00666
<b>Median</b>		1.00500	1.00900	1.00800	1.00500
<b>Std. Deviation</b>		.024203	.013578	.011870	.013324
<b>Range</b>		.138	.090	.079	.107
<b>Percentiles</b>	2.5	.92380	.96495	.96665	.98648
	25	.99500	1.00350	1.00300	1.00275
	50	1.00500	1.00900	1.00800	1.00500
	75	1.01400	1.01200	1.01250	1.01000
	97.5	1.05480	1.04415	1.03485	1.05690

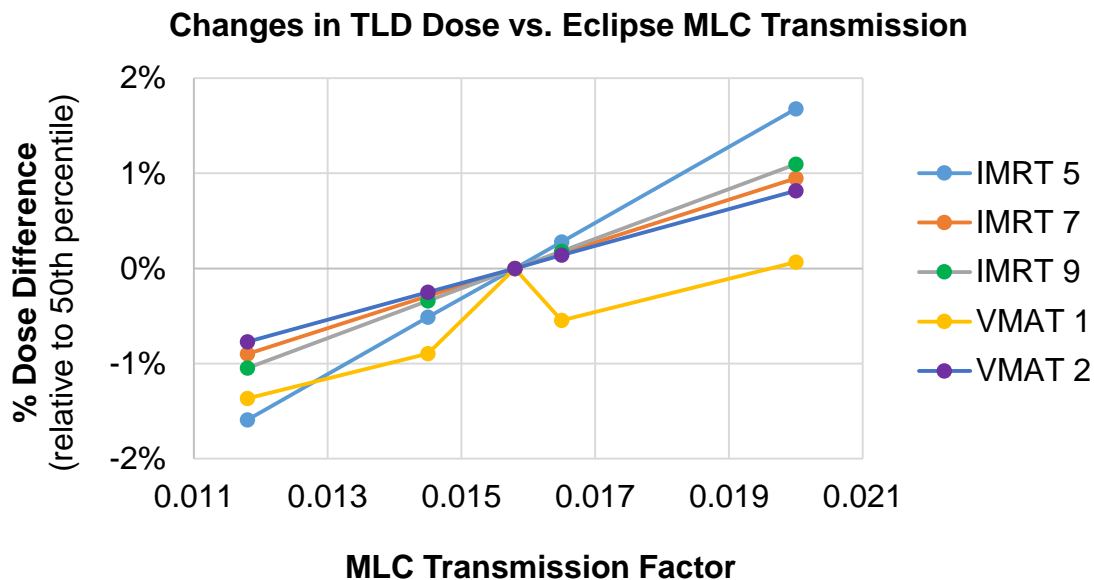
**Table D-14.** Percentile information for off-axis measurement errors as reported from IROC Houston site visit data (normalized to  $d_{\max}$  for a 40x40 cm field). All errors introduced to the curves were such that the error was approximately as reported below.

Off-Axis Factors		5 cm	10 cm	15 cm
<b>N</b>	Valid	341	347	330
	Missing	0	0	0
<b>Mean</b>		.99792	3.88663	.99737
<b>Median</b>		.99806	.99906	.99715
<b>Std. Deviation</b>		.006204	53.784619	.010901
<b>Percentiles</b>	2.5	.98653	.97954	.97399
	25	.99420	.99522	.99213
	50	.99806	.99906	.99715
	75	1.00195	1.00382	1.00317
	97.5	1.01177	1.01555	1.01734

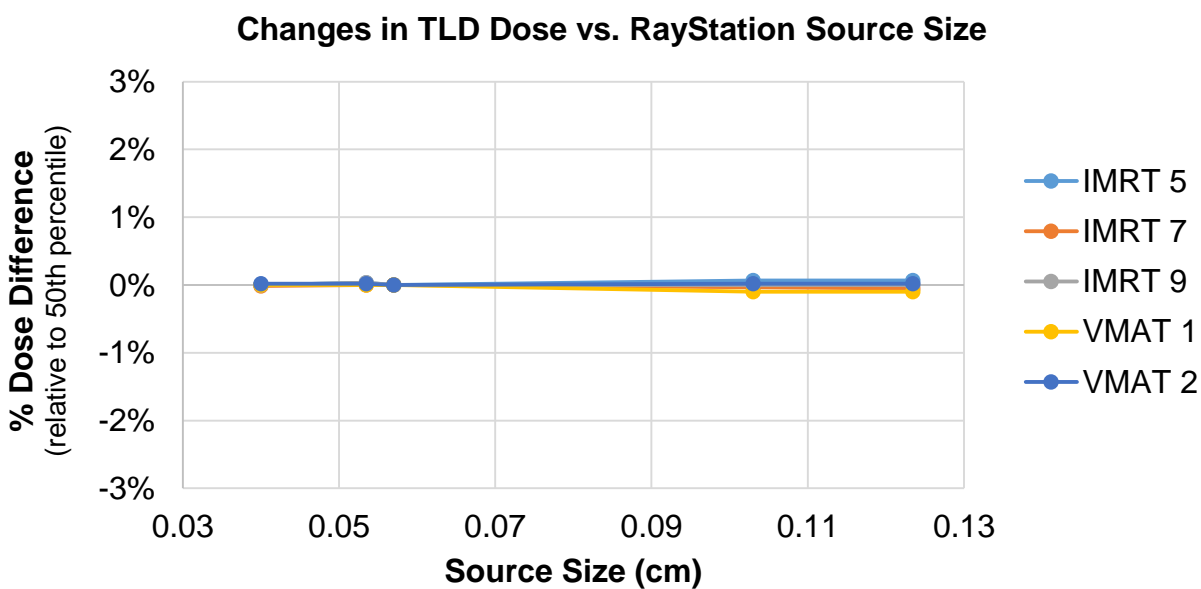


**Figure D-11.** Changes in average TLD dose calculated following manipulation of the effective target spot size in Eclipse for each of the five IROC-H head and neck phantom plans.

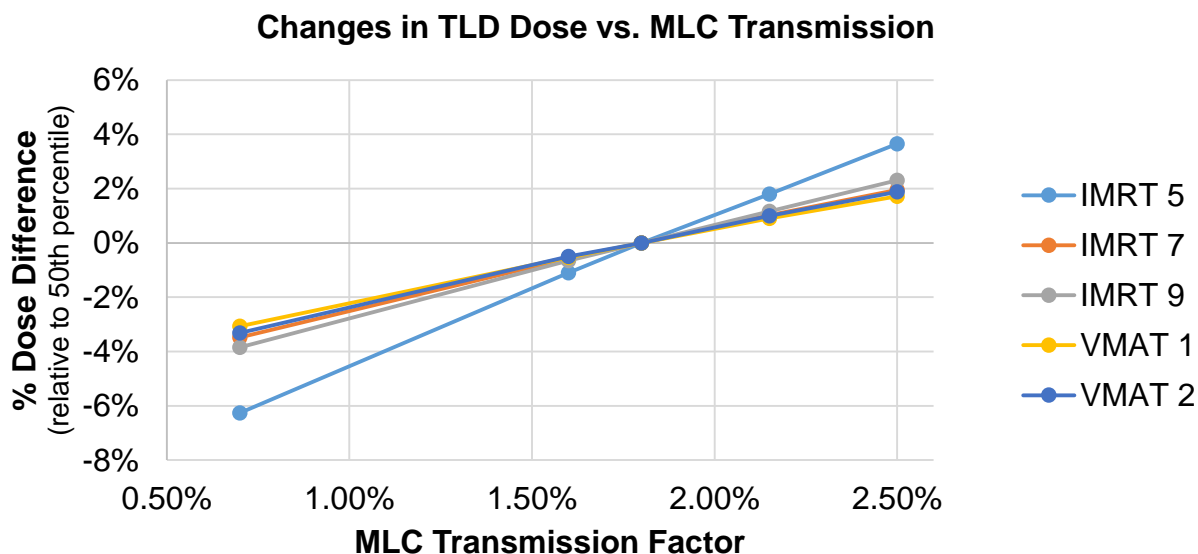




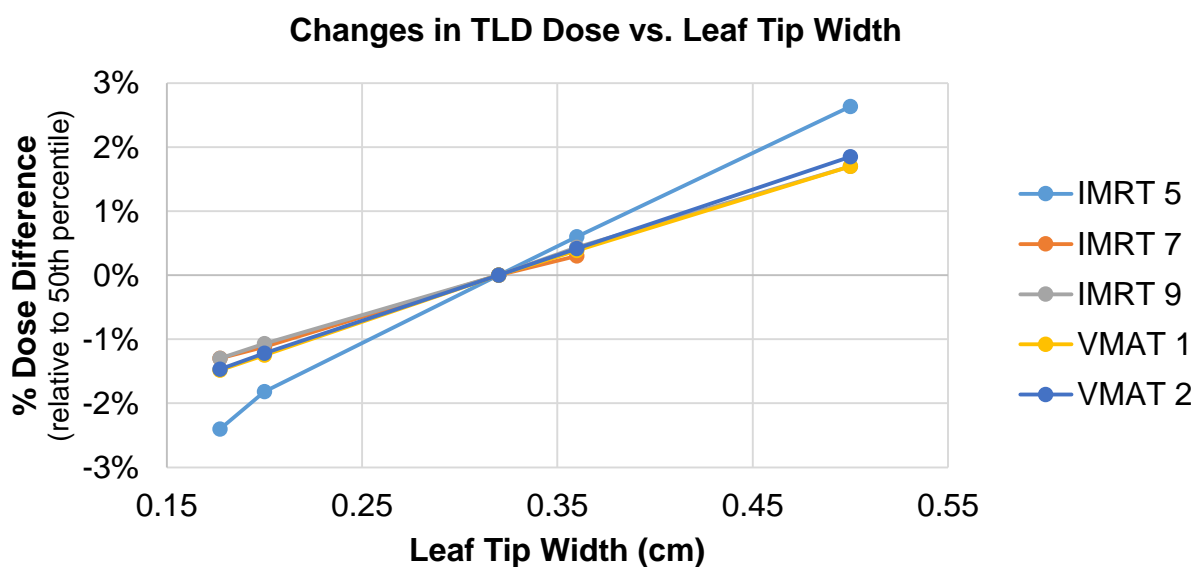
**Figure D-12.** Changes in average TLD dose calculated following manipulation of the MLC transmission factor in Eclipse for each of the five IROC-H head and neck phantom plans.



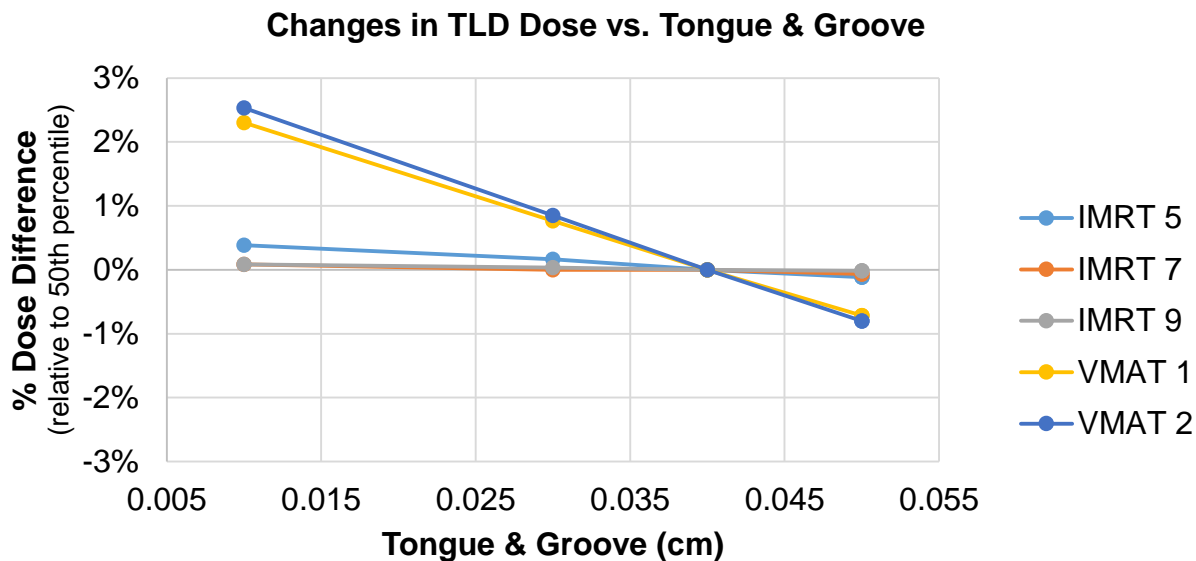
**Figure D-13.** Changes in average TLD dose calculated following manipulation of the source size in RayStation for each of the five IROC-H head and neck phantom plans.



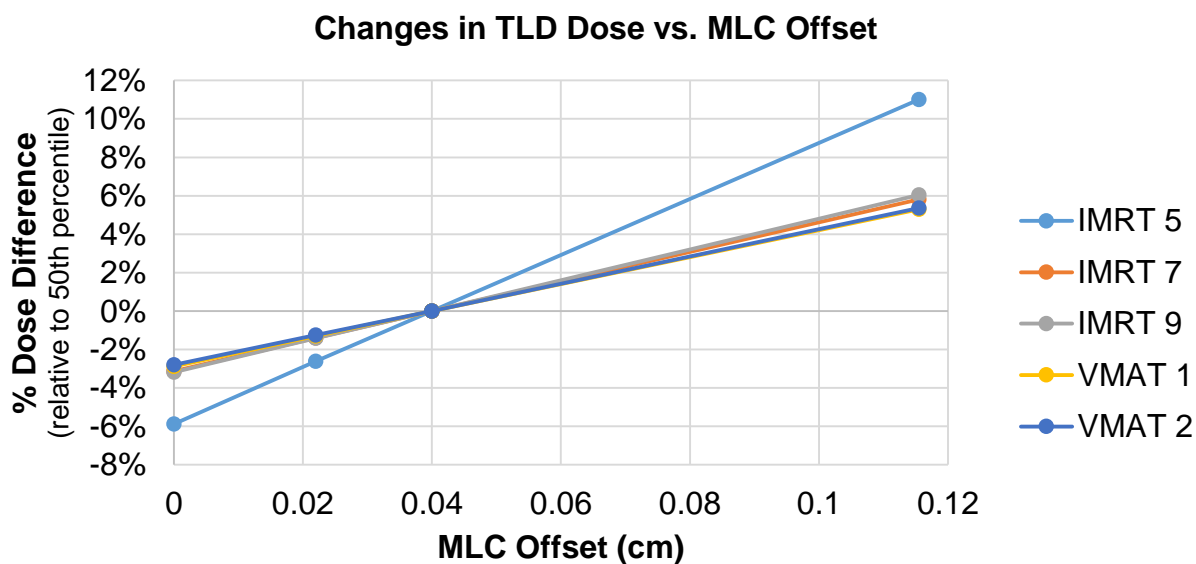
**Figure D-14.** Changes in average TLD dose calculated following manipulation of the MLC transmission factor in RayStation for each of the five IROC-H head and neck phantom plans.



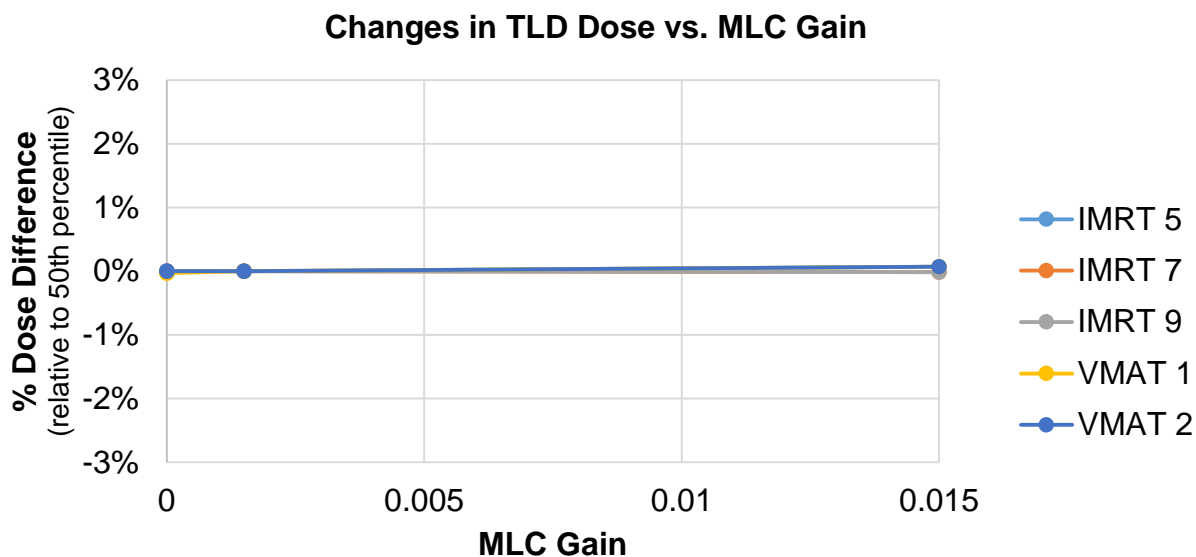
**Figure D-15.** Changes in average TLD dose calculated following manipulation of the leaf tip width in RayStation for each of the five IROC-H head and neck phantom plans.



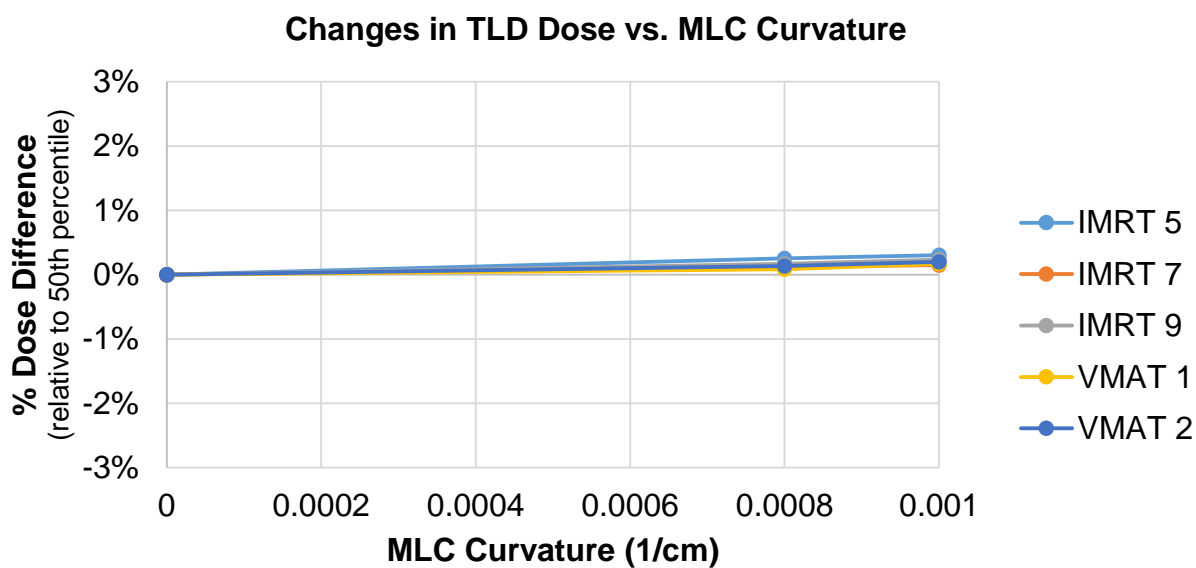
**Figure D-16.** Changes in average TLD dose calculated following manipulation of the tongue and groove in RayStation for each of the five IROC-H head and neck phantom plans.



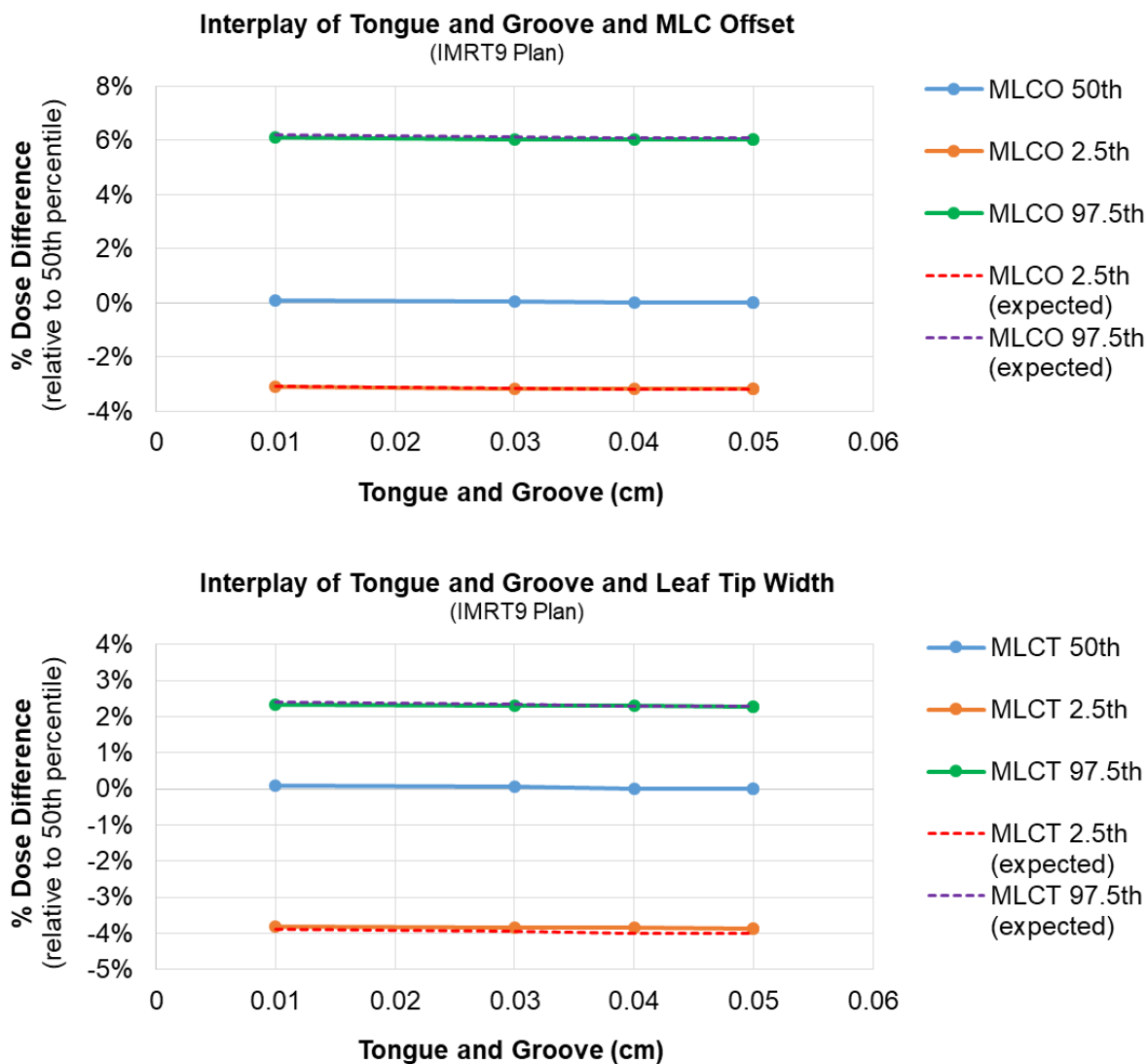
**Figure D-17.** Changes in average TLD dose calculated following manipulation of the MLC position offset in RayStation for each of the five IROC-H head and neck phantom plans.



**Figure D-18.** Changes in average TLD dose calculated following manipulation of the MLC position gain in RayStation for each of the five IROC-H head and neck phantom plans.



**Figure D-19.** Changes in average TLD dose calculated following manipulation of the MLC position curvature in RayStation for each of the five IROC-H head and neck phantom plans.



**Figure D-20.** Interplay between tongue and groove with MLC offset (top) and leaf tip width (bottom) in RayStation using the IMRT9 plan.

## **Appendix E: Supplement to Chapter 6**

This appendix serves as the supplement to  
Chapter 6: Relationships between Beam Modeling and Phantom Performance

This appendix contains supplementary information related to the analyses of Chapter 6. Here exist more comprehensive data referenced in the chapter text. First, Tables E-1 and E-2 outline the counts and calculated p-values of all pass/fail and poor/well-performed irradiation relationships with individual reported parameters. Next, Table E-3 describes the counting of poor-performing and failing phantom cases with atypical parameters among each TPS. Figures E-1 and E-2 depict two other cases of interest not discussed in the manuscript text. Film analysis for the identified institutions, along with a complete record of the institutions reviewed for beam modeling/phantom status correlations can be found at: J:\Everyone\Mallory\Dissertation\Appendix E. Finally Figures E-3 through E-5 contain additional recalculated IMRT H&N plan dose profiles that demonstrate systematic changes, but were not included in the manuscript text.

**Table E-1.** Fisher's exact test for typical/atypical parameters values versus poor- or well-performed irradiation status.

Parameter	Typical/ Good	Extrema/ Good	Typical/ Poor	Extrema/ Poor	Fisher's Exact p-value
<b>ECLIPSE AAA</b>					
Source Size X	167	24	30	3	0.7744
Source Size Y	185	6	32	1	1
DLG	150	41	21	13	<b>0.04845</b>
MLC Transmission	154	36	24	10	0.1714
<b>ECLIPSE ACUROSXB</b>					
Source Size X	30	3	2	2	0.08002
Source Size Y	31	2	2	2	<b>0.04998</b>
DLG	24	8	1	3	0.07563
MLC Transmission	21	11	4	0	0.2904
<b>PINNACLE</b>					
Source Size X	33	2	5	0	1
Source Size Y	33	2	5	0	1
Leaf Tip Radius	30	3	4	1	0.4456
Tongue & Groove	30	5	4	1	1
Add'l T&G	31	3	5	0	1
MLC Transmission	28	6	3	2	0.2677
FF Gaussian Height	30	5	4	1	1
FF Gaussian Width	25	10	4	1	1
<b>RAYSTATION</b>					
Source Size X	16	3	4	8	<b>0.007</b>
Source Size Y	14	5	8	4	0.7039
MLC Transmission	17	1	11	1	1
Leaf Tip Width	17	2	11	1	1
Tongue & Groove	18	1	10	2	0.5435
MLC Offset	4	3	1	2	1
MLC Gain	3	4	3	0	0.2
MLC Curvature	6	1	3	0	1

**Table E-2.** Fisher's exact test for typical/atypical parameters values versus passing or failing phantom status.

Parameter	Typical/ Pass	Extrema/ Pass	Typical/ Fail	Extrema/ Fail	Fisher's Exact p-value
<b>ECLIPSE AAA</b>					
Source Size X	178	25	19	2	1
Source Size Y	196	7	21	0	1
DLG	160	44	11	10	<b>0.01372</b>
MLC Transmission	162	41	16	5	0.7764
<b>ECLIPSE ACUROSXB</b>					
Source Size X	32	5	0	0	N/A
Source Size Y	33	4	0	0	N/A
DLG	25	11	0	0	N/A
MLC Transmission	28	8	0	0	N/A
<b>PINNACLE</b>					
Source Size X	35	2	3	0	1
Source Size Y	35	2	3	0	1
Leaf Tip Radius	32	3	2	1	0.2907
Tongue & Groove	31	6	3	0	1
Add'l T&G	33	3	3	0	1
MLC Transmission	30	6	1	2	0.1011
FF Gaussian Height	3	33	1	2	0.05364
FF Gaussian Width	9	28	1	2	0.1781
<b>RAYSTATION</b>					
Source Size X	19	7	1	4	<b>0.04156</b>
Source Size Y	19	7	3	2	0.6125
MLC Transmission	24	1	4	1	0.3103
Leaf Tip Width	24	2	4	1	0.4216
Tongue & Groove	24	2	4	1	0.4216
MLC Offset	5	4	0	1	1
MLC Gain	5	4	1	0	1
MLC Curvature	8	1	1	0	1



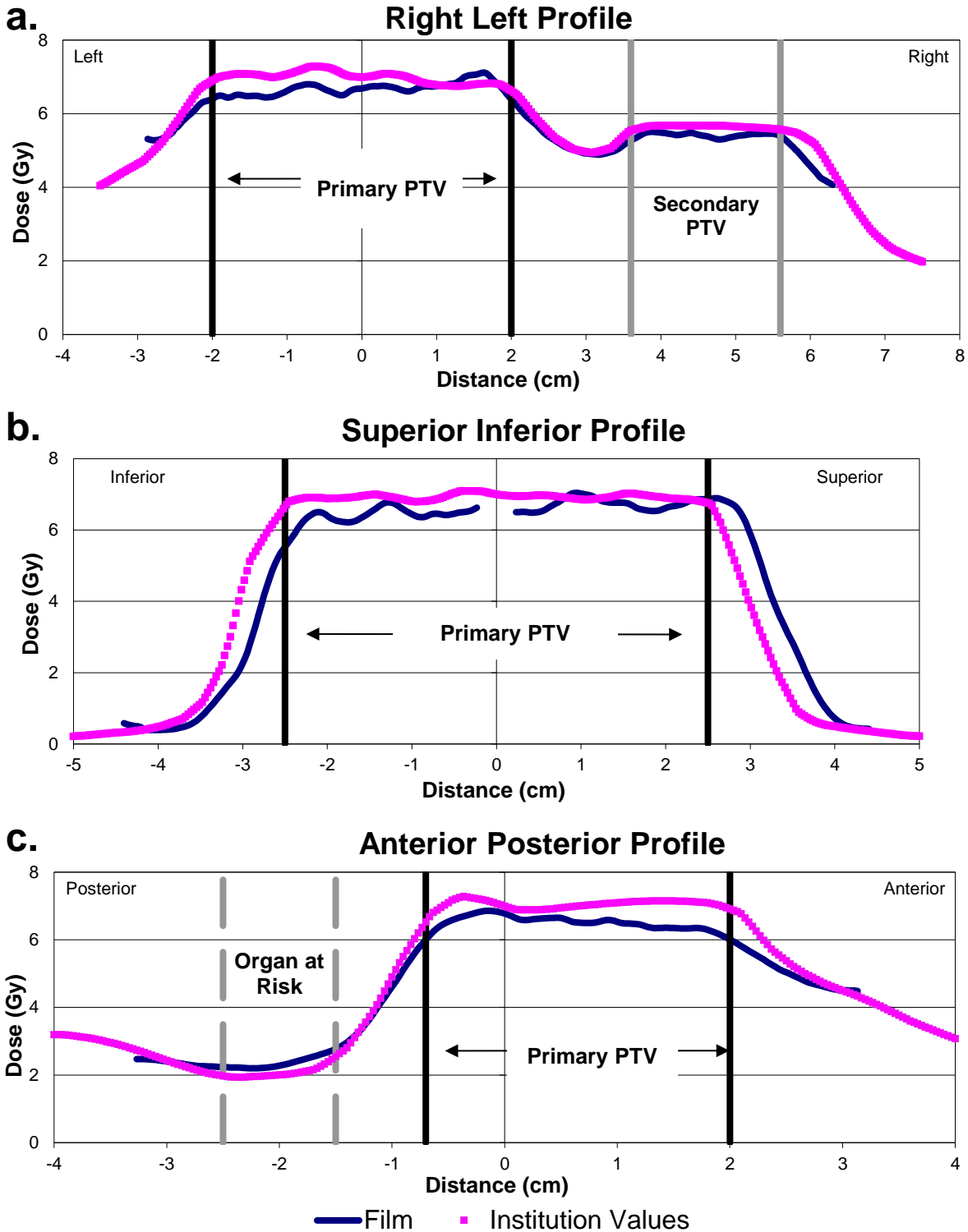
**Table E-3.** Counting of poor-performing and failing phantom cases with reported beam modeling parameters found important in dose calculation accuracy.

<b>Poorly Performed Irradiations</b>	<b>Without Atypical Parameters</b>	<b>Atypical Parameters</b>	<b>Important<sup>†</sup> Atypical Parameters</b>
Eclipse AAA	16	18	16
Eclipse AcurosXB	1	3	3
Pinnacle*	2	2	2
RayStation	5	7	3
Totals	24	30	24
<b>Failed Irradiations</b>			
Eclipse AAA	9	12	10
Eclipse AcurosXB	0	0	0
Pinnacle <sup>‡</sup>	0	3	3
RayStation	2*	4	3
Totals	11	19	16

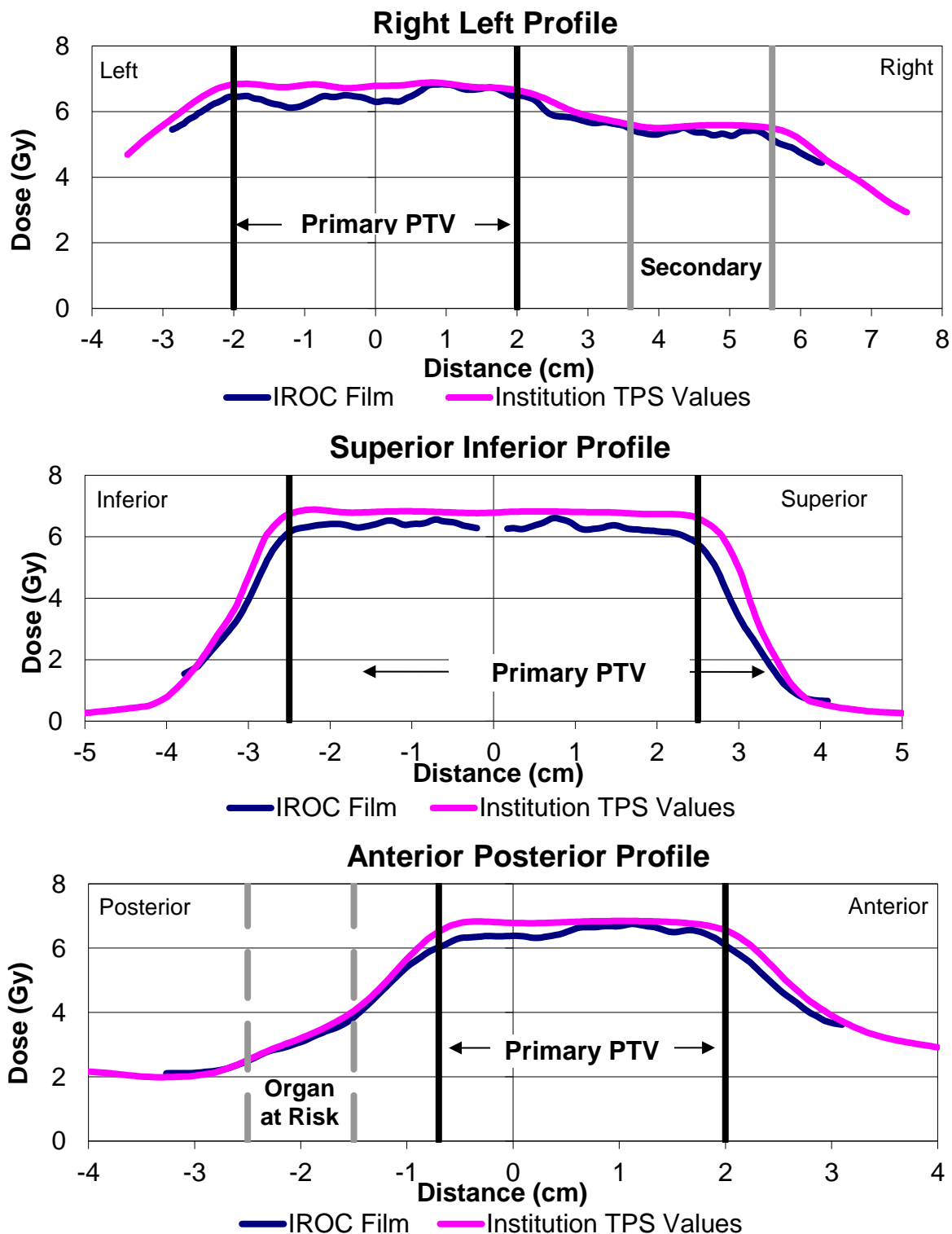
<sup>†</sup> Important parameters defined as those having a substantial effect on the dose calculation, on average up to 2% or greater for the most extreme values tested (see Chapter 5).

<sup>‡</sup> Based on literature, Gaussian factors and MLC transmission were deemed important for Pinnacle.<sup>14</sup>

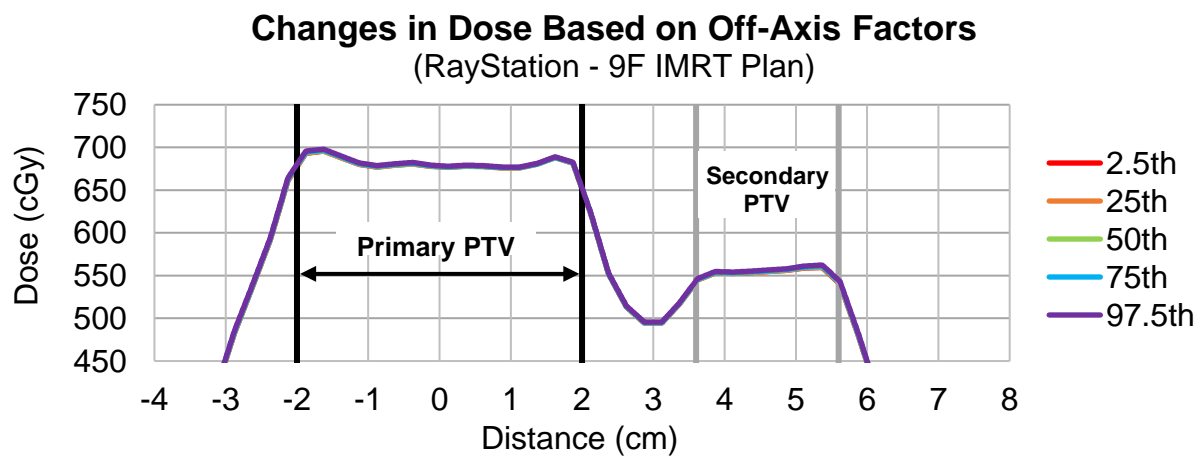
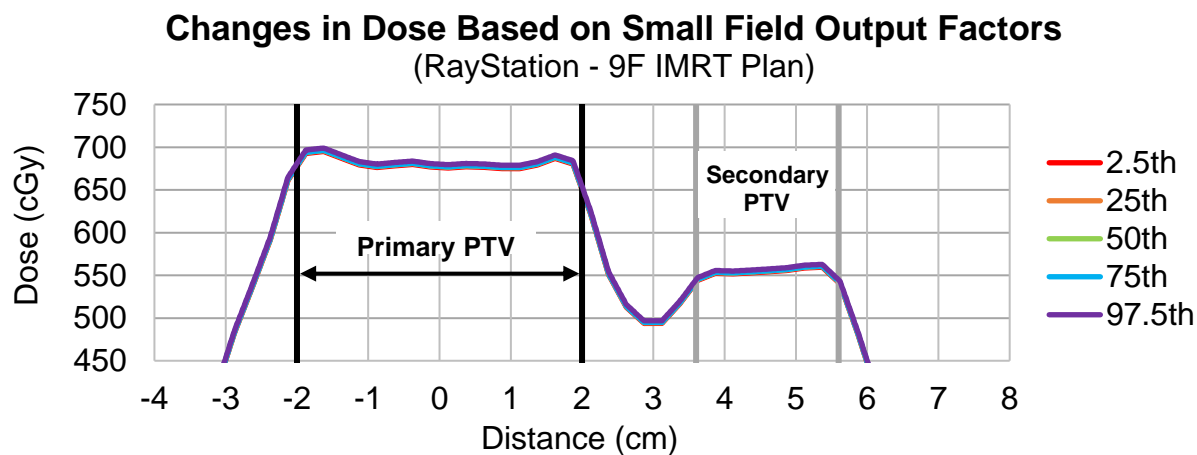
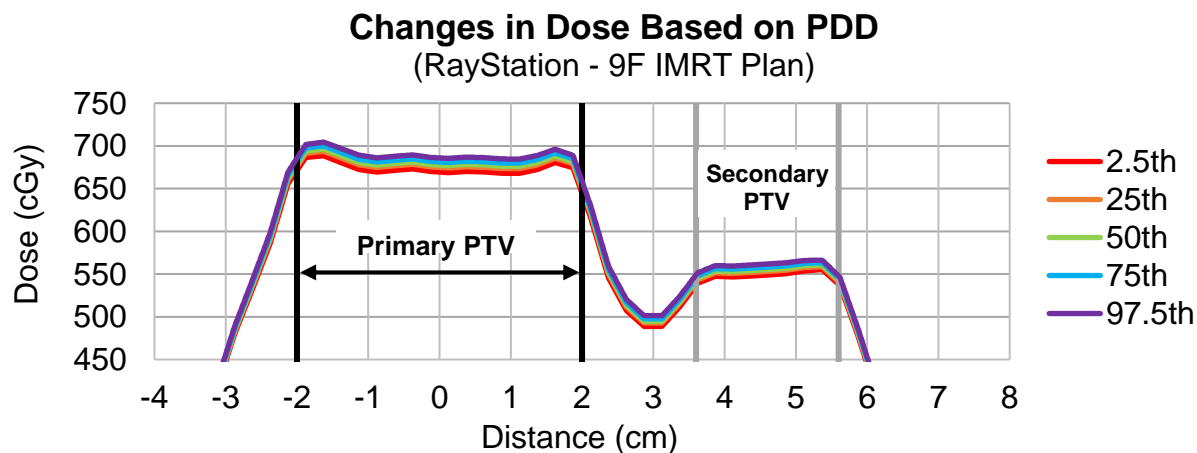
\* For one case, statistics could not be assessed due to limited samples from the original survey.



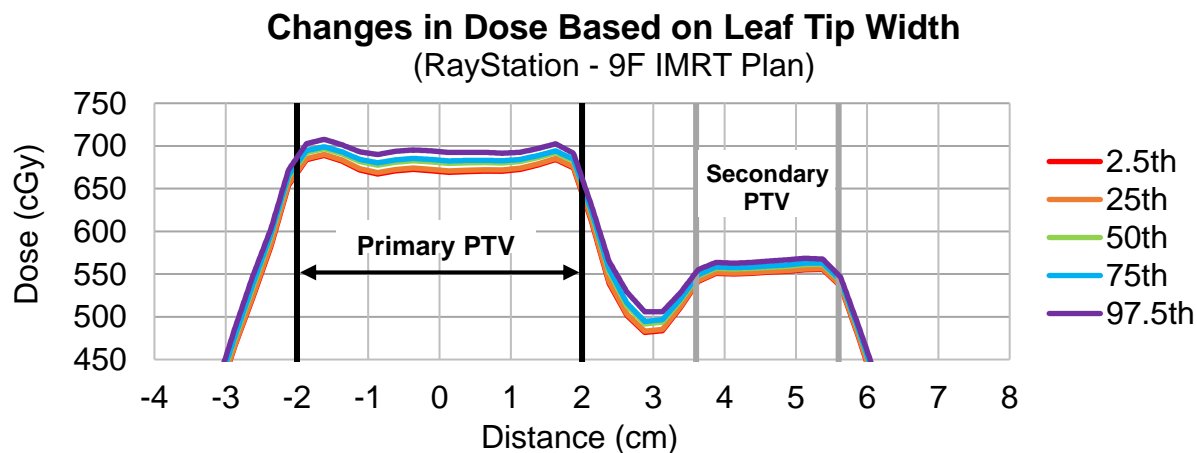
**Figure E-1.** Case of interest #1: Irradiation using RayStation with high MLC Offset (92<sup>nd</sup> percentile), leading them to greatly overestimate the dose delivered. Profiles shown are a) right-left, b) superior-inferior, and c) anterior-posterior. This institution reported 100% pixels passing for their IMRT QA. The institution performed poorly on a second attempt despite improving localization, presumably because the model was not adjusted.



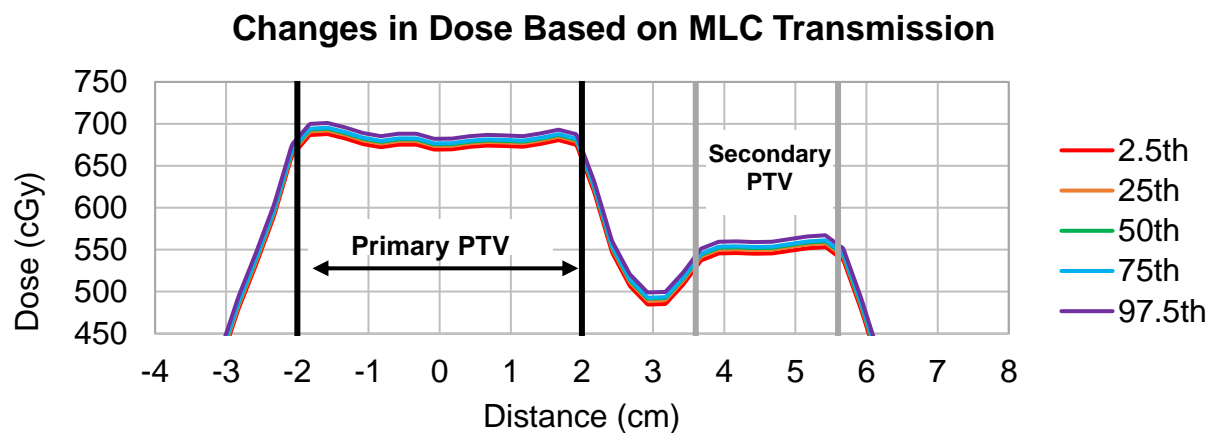
**Figure E-2.** Case of interest #2: Irradiation using RayStation with high MLC offset (100th percentile), that underestimated the dose delivered (purple) compared to film measurement (green). Dose profiles are a) left-right, b) superior-inferior, and c) anterior-posterior. This institution reported at least 99.2% pixels passing for their IMRT QA. A second attempt was made after adjusting the MLC offset, which improved the accuracy substantially.



**Figure E-3.** Recalculated IMRT H&N phantom profiles for dosimetric modeling parameters. From top to bottom: PDD, jaw-defined small field output factors, and off-axis factors.



**Figure E-4.** Recalculated IMRT H&N phantom profiles for RayStation’s leaf tip width.



**Figure E-5.** Recalculated IMRT H&N phantom profiles for Eclipse’s MLC transmission.

## **Appendix F: Supporting Data for Mobius3D Beam Models**

While not explicitly a part of the dissertation work, significant contributions were made to assist IROC Houston in its goal of automating phantom plan checks and identifying potential cases for which the TPS dose calculation is the most suspect cause for poor or failing phantom performance. As part of the research work accomplished, along with the assistance of both former and current lab members, multiple beam models were commissioned in the Mobius3D platform (Varian Medical Systems, Palo Alto, CA). The following pages contain data much like that of Tables D-8 through D-11, demonstrating the degree to which models were defined and tuned to best approximate IROC Houston reference data. These models were tuned to represent linear accelerators of “average” dosimetric performance. All models are based upon representative Linac class information previously published by James Kerns.<sup>55,56</sup>

It should be noted that several models were better suited for use with Mobius3D than others, given the ability to “tune” the model to the average performance. In general, lower energy beam models (e.g. 6 MV) were able to be better approximated more than those of higher energies. Fortunately, the vast majority of IROC phantom irradiations are performed with 6 MV beams, meaning that this work can characterize most cases. Additionally, MLC output factors were universally difficult to manage and tune (in most cases they were static).

The results presented in this appendix are for the most up-to-date Mobius3D server at IROC Houston (current IP Address: XX.XXX.XX.226) using Mobius3D version 2.1.2. Should additional upgrades be made to the system, the works included herein will need to be updated and validated.

## VARIAN BASE - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.861	0.861	0.00%	0.860	-0.12%
PDD 10x10 (norm dmax) - 10 cm	0.664	0.663	-0.15%	0.663	-0.15%
PDD 10x10 (norm dmax) - 15 cm	0.504	0.506	0.40%	0.505	0.20%
PDD 10x10 (norm dmax) - 20 cm	0.381	0.382	0.26%	0.379	-0.52%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.957	0.965	0.84%	0.959	0.21%
15x15	1.032	1.029	-0.29%	1.032	0.00%
20x20	1.054	1.052	-0.19%	1.054	0.00%
30x30	1.08	1.083	0.28%	1.079	-0.09%
<b>Off Axis Factors</b>					
5 cm	1.029	1.022	-0.68%	1.029	0.00%
10 cm	1.042	1.039	-0.29%	1.042	0.00%
15 cm	1.055	1.050	-0.47%	1.054	-0.09%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.807	0.788	-2.35%	0.786	-2.60%
3x3	0.852	0.845	-0.82%	0.839	-1.53%
4x4	0.888	0.882	-0.68%	0.877	-1.24%
6x6	0.938	0.940	0.21%	0.937	-0.11%
<b>SBRT Output Factors</b>					
2x2	0.778	0.783	0.64%	0.779	0.13%
3x3	0.82	0.831	1.34%	0.821	0.12%
4x4	0.855	0.866	1.29%	0.857	0.23%
6x6	0.915	0.923	0.87%	0.914	-0.11%
		<b>Average Error</b>	<b>0.63%</b>	<b>Average Error</b>	<b>0.39%</b>
		<b>Total Deviation</b>	<b>12.056%</b>	<b>Total Deviation</b>	<b>7.454%</b>

## TRUEBEAM - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.864	0.861	-0.35%	0.860	-0.46%
PDD 10x10 (norm dmax) - 10 cm	0.663	0.663	0.00%	0.663	0.00%
PDD 10x10 (norm dmax) - 15 cm	0.504	0.506	0.40%	0.504	0.00%
PDD 10x10 (norm dmax) - 20 cm	0.38	0.382	0.53%	0.378	-0.53%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.962	0.965	0.31%	0.960	-0.21%
15x15	1.03	1.029	-0.10%	1.030	0.00%
20x20	1.052	1.052	0.00%	1.052	0.00%
30x30	1.074	1.083	0.84%	1.074	0.00%
<b>Off Axis Factors</b>					
5 cm	1.024	1.022	-0.20%	1.024	0.00%
10 cm	1.042	1.039	-0.29%	1.042	0.00%
15 cm	1.054	1.050	-0.38%	1.053	-0.09%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.804	0.788	-1.99%	0.786	-2.24%
3x3	0.849	0.845	-0.47%	0.839	-1.18%
4x4	0.885	0.882	-0.34%	0.877	-0.90%
6x6	0.937	0.940	0.32%	0.937	0.00%
<b>SBRT Output Factors</b>					
2x2	0.781	0.783	0.26%	0.777	-0.51%
3x3	0.825	0.831	0.73%	0.824	-0.12%
4x4	0.856	0.866	1.17%	0.856	0.00%
6x6	0.914	0.923	0.98%	0.915	0.11%
		<b>Average Error</b>	<b>0.51%</b>	<b>Average Error</b>	<b>0.33%</b>
		<b>Total Deviation</b>	<b>9.637%</b>	<b>Total Deviation</b>	<b>6.355%</b>



# TRUEBEAM FFF - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.848	0.842	-0.71%	0.842	-0.71%
PDD 10x10 (norm dmax) - 10 cm	0.636	0.630	-0.94%	0.630	-0.94%
PDD 10x10 (norm dmax) - 15 cm	0.469	0.467	-0.43%	0.467	-0.43%
PDD 10x10 (norm dmax) - 20 cm	0.345	0.342	-0.87%	0.342	-0.87%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.976	0.973	-0.31%	0.965	-1.13%
15x15	1.017	1.018	0.10%	1.017	0.00%
20x20	1.029	1.035	0.58%	1.029	0.00%
30x30	1.041	1.053	1.15%	1.040	-0.10%
<b>Off Axis Factors</b>					
5 cm	0.909	0.913	0.44%	0.913	0.44%
10 cm	0.766	0.771	0.65%	0.771	0.65%
15 cm	0.642	0.649	1.09%	0.649	1.09%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.802	0.798	-0.50%	0.798	-0.50%
3x3	0.842	0.848	0.71%	0.848	0.71%
4x4	0.875	0.880	0.57%	0.880	0.57%
6x6	0.928	0.936	0.86%	0.936	0.86%
<b>SBRT Output Factors</b>					
2x2	0.786	0.792	0.76%	0.785	-0.13%
3x3	0.827	0.841	1.69%	0.828	0.12%
4x4	0.861	0.876	1.74%	0.863	0.23%
6x6	0.92	0.930	1.09%	0.922	0.22%
		Average Error	0.80%	Average Error	0.51%
		Total Deviation	15.202%	Total Deviation	9.696%

## VARIAN 2100 - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.864	0.860	-0.46%	0.862	-0.23%
PDD 10x10 (norm dmax) - 10 cm	0.667	0.663	-0.60%	0.665	-0.30%
PDD 10x10 (norm dmax) - 15 cm	0.508	0.506	-0.39%	0.507	-0.20%
PDD 10x10 (norm dmax) - 20 cm	0.385	0.382	-0.78%	0.381	-1.04%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.957	0.965	0.84%	0.958	0.10%
15x15	1.033	1.029	-0.39%	1.033	0.00%
20x20	1.055	1.052	-0.28%	1.055	0.00%
30x30	1.082	1.083	0.09%	1.082	0.00%
<b>Off Axis Factors</b>					
5 cm	1.029	1.022	-0.68%	1.029	0.00%
10 cm	1.041	1.039	-0.19%	1.040	-0.10%
15 cm	1.053	1.050	-0.28%	1.052	-0.09%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.808	0.788	-2.48%	0.787	-2.60%
3x3	0.853	0.845	-0.94%	0.841	-1.41%
4x4	0.889	0.882	-0.79%	0.879	-1.12%
6x6	0.94	0.940	0.00%	0.938	-0.21%
<b>SBRT Output Factors</b>					
2x2	0.776	0.783	0.90%	0.775	-0.13%
3x3	0.823	0.831	0.97%	0.824	0.12%
4x4	0.86	0.866	0.70%	0.861	0.12%
6x6	0.916	0.923	0.76%	0.914	-0.22%
		Average Error	0.66%	Average Error	0.42%
		Total Deviation	12.529%	Total Deviation	7.991%

## VARIAN 2300 - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.866	0.860	-0.69%	0.862	-0.46%
PDD 10x10 (norm dmax) - 10 cm	0.669	0.663	-0.90%	0.665	-0.60%
PDD 10x10 (norm dmax) - 15 cm	0.509	0.506	-0.59%	0.507	-0.39%
PDD 10x10 (norm dmax) - 20 cm	0.385	0.382	-0.78%	0.380	-1.30%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.957	0.965	0.84%	0.956	-0.10%
15x15	1.032	1.029	-0.29%	1.035	0.29%
20x20	1.054	1.052	-0.19%	1.054	0.00%
30x30	1.08	1.083	0.28%	1.081	0.09%
<b>Off Axis Factors</b>					
5 cm	1.03	1.022	-0.78%	1.027	-0.29%
10 cm	1.04	1.039	-0.10%	1.039	-0.10%
15 cm	1.05	1.050	0.00%	1.048	-0.19%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	N/A				
3x3	N/A				
4x4	N/A				
6x6	N/A				
<b>SBRT Output Factors</b>					
2x2	N/A				
3x3	N/A				
4x4	N/A				
6x6	N/A				
		Average Error	0.49%	Average Error	0.35%
		Total Deviation	5.425%	Total Deviation	3.817%

## TRILOGY SRS - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.862	0.860	-0.23%	0.860	-0.23%
PDD 10x10 (norm dmax) - 10 cm	0.662	0.663	0.15%	0.663	0.15%
PDD 10x10 (norm dmax) - 15 cm	0.5	0.506	1.20%	0.506	1.20%
PDD 10x10 (norm dmax) - 20 cm	0.375	0.382	1.87%	0.382	1.87%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.969	0.970	0.10%	0.967	-0.21%
15x15	1.024	1.024	0.00%	1.023	-0.10%
20x20	N/A				
30x30	N/A				
<b>Off Axis Factors</b>					
5 cm	N/A				
10 cm	N/A				
15 cm	N/A				
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.816	0.788	-3.43%	0.788	-3.43%
3x3	0.857	0.845	-1.40%	0.845	-1.40%
4x4	0.888	0.882	-0.68%	0.882	-0.68%
6x6	0.937	0.940	0.32%	0.940	0.32%
<b>SBRT Output Factors</b>					
2x2	0.792	0.792	0.00%	0.792	0.00%
3x3	0.833	0.846	1.56%	0.833	0.00%
4x4	0.867	0.875	0.92%	0.867	0.00%
6x6	0.924	0.928	0.43%	0.925	0.11%
		<b>Average Error</b>	<b>0.88%</b>	<b>Average Error</b>	<b>0.69%</b>
		<b>Total Deviation</b>	<b>12.297%</b>	<b>Total Deviation</b>	<b>9.689%</b>

## VARIAN 600 - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.857	0.860	0.35%	0.857	0.00%
PDD 10x10 (norm dmax) - 10 cm	0.662	0.663	0.15%	0.660	-0.30%
PDD 10x10 (norm dmax) - 15 cm	0.501	0.506	1.00%	0.502	0.20%
PDD 10x10 (norm dmax) - 20 cm	0.378	0.382	1.06%	0.377	-0.26%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.966	0.965	-0.10%	0.968	0.21%
15x15	1.028	1.029	0.10%	1.029	0.10%
20x20	1.046	1.052	0.57%	1.047	0.10%
30x30	1.065	1.083	1.69%	1.067	0.19%
<b>Off Axis Factors</b>					
5 cm	1.029	1.022	-0.68%	1.028	-0.10%
10 cm	1.043	1.039	-0.38%	1.043	0.00%
15 cm	1.053	1.050	-0.28%	1.053	0.00%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	N/A				
3x3	N/A				
4x4	N/A				
6x6	N/A				
<b>SBRT Output Factors</b>					
2x2	N/A				
3x3	N/A				
4x4	N/A				
6x6	N/A				
		<b>Average Error</b>	<b>0.58%</b>	<b>Average Error</b>	<b>0.13%</b>
		<b>Total Deviation</b>	<b>6.371%</b>	<b>Total Deviation</b>	<b>1.451%</b>

## VARIAN 6EX - 6X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.857	0.86	0.35%	0.859	0.23%
PDD 10x10 (norm dmax) - 10 cm	0.662	0.663	0.15%	0.662	0.00%
PDD 10x10 (norm dmax) - 15 cm	0.502	0.506	0.80%	0.504	0.40%
PDD 10x10 (norm dmax) - 20 cm	0.377	0.382	1.33%	0.377	0.00%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.964	0.965	0.10%	0.969	0.52%
15x15	1.023	1.029	0.59%	1.023	0.00%
20x20	1.041	1.052	1.06%	1.041	0.00%
30x30	1.059	1.083	2.27%	1.058	-0.09%
<b>Off Axis Factors</b>					
5 cm	1.032	1.022	-0.97%	1.029	-0.29%
10 cm	1.046	1.039	-0.67%	1.044	-0.19%
15 cm	1.062	1.05	-1.13%	1.056	-0.56%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.815	0.788	-3.31%	0.786	-3.56%
3x3	0.86	0.845	-1.74%	0.839	-2.44%
4x4	0.894	0.882	-1.34%	0.876	-2.01%
6x6	0.942	0.94	-0.21%	0.937	-0.53%
<b>SBRT Output Factors</b>					
2x2	N/A				
3x3	N/A				
4x4	N/A				
6x6	N/A				
		<b>Average Error</b>	<b>1.07%</b>	<b>Average Error</b>	<b>0.72%</b>
		<b>Total Deviation</b>	<b>16.017%</b>	<b>Total Deviation</b>	<b>10.836%</b>

## VARIAN BASE - 10X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.913	0.940	2.96%	0.918	0.55%
PDD 10x10 (norm dmax) - 10 cm	0.73	0.742	1.64%	0.728	-0.27%
PDD 10x10 (norm dmax) - 15 cm	0.582	0.587	0.86%	0.578	-0.69%
PDD 10x10 (norm dmax) - 20 cm	0.46	0.464	0.87%	0.458	-0.43%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.953	0.959	0.63%	0.948	-0.52%
15x15	1.033	1.029	-0.39%	1.034	0.10%
20x20	1.054	1.050	-0.38%	1.056	0.19%
30x30	1.083	1.080	-0.28%	1.086	0.28%
<b>Off Axis Factors</b>					
5 cm	1.029	1.022	-0.68%	1.032	0.29%
10 cm	1.044	1.034	-0.96%	1.047	0.29%
15 cm	1.053	1.038	-1.42%	1.059	0.57%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.825	0.786	-4.73%	0.795	-3.64%
3x3	0.881	0.882	0.11%	0.887	0.68%
4x4	0.918	0.926	0.87%	0.927	0.98%
6x6	0.959	0.966	0.73%	0.966	0.73%
<b>SBRT Output Factors</b>					
2x2	0.794	0.780	-1.76%	0.813	2.39%
3x3	0.846	0.845	-0.12%	0.847	0.12%
4x4	0.877	0.886	1.03%	0.886	1.03%
6x6	0.929	0.936	0.75%	0.924	-0.54%
		<b>Average Error</b>	<b>1.11%</b>	<b>Average Error</b>	<b>0.75%</b>
		<b>Total Deviation</b>	<b>21.169%</b>	<b>Total Deviation</b>	<b>14.284%</b>

## TRUEBEAM - 10X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.918	0.940	2.40%	0.931	1.42%
PDD 10x10 (norm dmax) - 10 cm	0.737	0.742	0.68%	0.737	0.00%
PDD 10x10 (norm dmax) - 15 cm	0.586	0.587	0.17%	0.586	0.00%
PDD 10x10 (norm dmax) - 20 cm	0.463	0.464	0.22%	0.465	0.43%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.956	0.959	0.31%	0.957	0.10%
15x15	1.032	1.029	-0.29%	1.032	0.00%
20x20	1.053	1.050	-0.28%	1.053	0.00%
30x30	1.077	1.080	0.28%	1.078	0.09%
<b>Off Axis Factors</b>					
5 cm	1.029	1.022	-0.68%	1.029	0.00%
10 cm	1.047	1.034	-1.24%	1.047	0.00%
15 cm	1.056	1.038	-1.70%	1.055	-0.09%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.823	0.786	-4.50%	0.783	-4.86%
3x3	0.88	0.882	0.23%	0.885	0.57%
4x4	0.916	0.926	1.09%	0.930	1.53%
6x6	0.958	0.966	0.84%	0.968	1.04%
<b>SBRT Output Factors</b>					
2x2	0.79	0.780	-1.27%	0.791	0.13%
3x3	0.849	0.845	-0.47%	0.844	-0.59%
4x4	0.884	0.886	0.23%	0.886	0.23%
6x6	0.933	0.936	0.32%	0.936	0.32%
		<b>Average Error</b>	<b>0.90%</b>	<b>Average Error</b>	<b>0.60%</b>
		<b>Total Deviation</b>	<b>17.191%</b>	<b>Total Deviation</b>	<b>11.404%</b>



## TRUEBEAM FFF - 10X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.908	0.911	0.33%	0.911	0.33%
PDD 10x10 (norm dmax) - 10 cm	0.712	0.709	-0.42%	0.709	-0.42%
PDD 10x10 (norm dmax) - 15 cm	0.554	0.551	-0.54%	0.551	-0.54%
PDD 10x10 (norm dmax) - 20 cm	0.43	0.427	-0.70%	0.427	-0.70%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.98	0.980	0.00%	0.975	-0.51%
15x15	1.015	1.012	-0.30%	1.014	-0.10%
20x20	1.026	1.024	-0.19%	1.026	0.00%
30x30	1.034	1.037	0.29%	1.035	0.10%
<b>Off Axis Factors</b>					
5 cm	0.823	0.824	0.12%	0.824	0.12%
10 cm	0.632	0.629	-0.47%	0.629	-0.47%
15 cm	0.497	0.496	-0.20%	0.496	-0.20%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.842	0.820	-2.61%	0.820	-2.61%
3x3	0.892	0.886	-0.67%	0.886	-0.67%
4x4	0.918	0.918	0.00%	0.918	0.00%
6x6	0.956	0.959	0.31%	0.959	0.31%
<b>SBRT Output Factors</b>					
2x2	0.825	0.820	-0.61%	0.824	-0.12%
3x3	0.881	0.881	0.00%	0.881	0.00%
4x4	0.907	0.916	0.99%	0.909	0.22%
6x6	0.948	0.954	0.63%	0.949	0.11%
		<b>Average Error</b>	<b>0.49%</b>	<b>Average Error</b>	<b>0.40%</b>
		<b>Total Deviation</b>	<b>9.400%</b>	<b>Total Deviation</b>	<b>7.540%</b>

## VARIAN BASE - 15X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.943	0.948	0.53%	0.950	0.74%
PDD 10x10 (norm dmax) - 10 cm	0.767	0.771	0.52%	0.773	0.78%
PDD 10x10 (norm dmax) - 15 cm	0.617	0.621	0.65%	0.624	1.13%
PDD 10x10 (norm dmax) - 20 cm	0.496	0.497	0.20%	0.500	0.81%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.959	0.957	-0.21%	0.965	0.63%
15x15	1.03	1.032	0.19%	1.028	-0.19%
20x20	1.05	1.059	0.86%	1.054	0.38%
30x30	1.075	1.088	1.21%	1.075	0.00%
<b>Off Axis Factors</b>					
5 cm	1.038	1.037	-0.10%	1.038	0.00%
10 cm	1.049	1.041	-0.76%	1.047	-0.19%
15 cm	1.061	1.049	-1.13%	1.055	-0.57%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.831	0.793	-4.57%	0.795	-4.33%
3x3	0.893	0.891	-0.22%	0.891	-0.22%
4x4	0.929	0.934	0.54%	0.934	0.54%
6x6	0.965	0.970	0.52%	0.970	0.52%
<b>SBRT Output Factors</b>					
2x2	0.799	0.771	-3.50%	0.787	-1.50%
3x3	0.865	0.847	-2.08%	0.865	0.00%
4x4	0.902	0.893	-1.00%	0.902	0.00%
6x6	0.949	0.942	-0.74%	0.949	0.00%
		<b>Average Error</b>	<b>1.03%</b>	<b>Average Error</b>	<b>0.66%</b>
		<b>Total Deviation</b>	<b>19.535%</b>	<b>Total Deviation</b>	<b>12.537%</b>

## TRUEBEAM - 15X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.946	0.948	0.21%	0.950	0.42%
PDD 10x10 (norm dmax) - 10 cm	0.769	0.771	0.26%	0.774	0.65%
PDD 10x10 (norm dmax) - 15 cm	0.62	0.621	0.16%	0.624	0.65%
PDD 10x10 (norm dmax) - 20 cm	0.499	0.497	-0.40%	0.500	0.20%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.953	0.957	0.42%	0.955	0.21%
15x15	1.035	1.032	-0.29%	1.035	0.00%
20x20	1.06	1.059	-0.09%	1.061	0.09%
30x30	1.085	1.088	0.28%	1.088	0.28%
<b>Off Axis Factors</b>					
5 cm	1.036	1.037	0.10%	1.036	0.00%
10 cm	1.045	1.041	-0.38%	1.045	0.00%
15 cm	1.056	1.049	-0.66%	1.054	-0.19%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.815	0.793	-2.70%	0.795	-2.45%
3x3	0.885	0.891	0.68%	0.891	0.68%
4x4	0.924	0.934	1.08%	0.934	1.08%
6x6	0.964	0.970	0.62%	0.970	0.62%
<b>SBRT Output Factors</b>					
2x2	0.784	0.771	-1.66%	0.784	0.00%
3x3	0.855	0.847	-0.94%	0.854	-0.12%
4x4	0.892	0.893	0.11%	0.892	0.00%
6x6	0.938	0.942	0.43%	0.940	0.21%
		<b>Average Error</b>	<b>0.60%</b>	<b>Average Error</b>	<b>0.41%</b>
		<b>Total Deviation</b>	<b>11.471%</b>	<b>Total Deviation</b>	<b>7.855%</b>

## VARIAN BASE - 18X

Parameter	Reference Values (Median)	Default Model	Default Error	Tuned Model	Tuned Error
<b>PDD Factors</b>					
PDD 10x10 (norm dmax) - 5 cm	0.963	0.968	0.52%	0.965	0.21%
PDD 10x10 (norm dmax) - 10 cm	0.793	0.795	0.25%	0.797	0.50%
PDD 10x10 (norm dmax) - 15 cm	0.647	0.648	0.15%	0.650	0.46%
PDD 10x10 (norm dmax) - 20 cm	0.527	0.526	-0.19%	0.527	0.00%
<b>Output Factors (norm 10x10@dmax)</b>					
6x6	0.943	0.950	0.74%	0.945	0.21%
15x15	1.041	1.053	1.15%	1.056	1.44%
20x20	1.066	1.073	0.66%	1.071	0.47%
30x30	1.094	1.104	0.91%	1.101	0.64%
<b>Off Axis Factors</b>					
5 cm	1.029	1.034	0.49%	1.030	0.10%
10 cm	1.044	1.046	0.19%	1.045	0.10%
15 cm	1.054	1.054	0.00%	1.054	0.00%
<b>MLC Output Factors (vs. 10x10 at 10 cm)</b>					
2x2	0.806	0.778	-3.47%	0.779	-3.35%
3x3	0.884	0.884	0.00%	0.885	0.11%
4x4	0.929	0.933	0.43%	0.933	0.43%
6x6	0.97	0.973	0.31%	0.973	0.31%
<b>SBRT Output Factors</b>					
2x2	0.767	0.739	-3.65%	0.766	-0.13%
3x3	0.847	0.853	0.71%	0.872	2.95%
4x4	0.891	0.900	1.01%	0.907	1.80%
6x6	0.942	0.944	0.21%	0.940	-0.21%
		<b>Average Error</b>	<b>0.79%</b>	<b>Average Error</b>	<b>0.71%</b>
		<b>Total Deviation</b>	<b>15.054%</b>	<b>Total Deviation</b>	<b>13.424%</b>

## References

1. American Society for Radiation Oncology (ASTRO). Accessed April 11, 2017. <https://www.astro.org/>
2. Kutcher GJ, Coia L, Gillin M, Hanson WF, Leibel S, Morton RJ, Palta JR, Purdy JA, Reinstein LE, Svensson GK, Weller M, Wingfield L. Comprehensive Qa for Radiation Oncology: Report of Aapm Radiation Therapy Committee Task Group 40. *Med Phys*. 1994;21(4):581-618. doi:10.1118/1.597316
3. International Commission on Radiation Units and Measurement. *Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures*; 1976.
4. Carson ME, Molineu A, Taylor PA, Followill DS, Stingo FC, Kry SF. Examining credentialing criteria and poor performance indicators for IROC Houston's anthropomorphic head and neck phantom. *Med Phys*. 2016;43(12):6491-6496. doi:10.1118/1.4967344
5. Ibbott GS, Molineu A, Followill DS. Independent evaluations of IMRT through the use of an anthropomorphic phantom. *Technol Cancer Res Treat*. 2006;5(5):481-487.
6. Molineu A, Followill DS, Balter PA, Hanson WF, Gillin MT, Huq MS, Eisbruch A, Ibbott GS. Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the Radiation Therapy Oncology Group. *Int J Radiat Oncol Biol Phys*. 2005;63(2):577-583. doi:10.1016/j.ijrobp.2005.05.021
7. Followill DS, Evans DR, Cherry C, Molineu A, Fisher G, Hanson WF, Ibbott GS. Design, development, and implementation of the radiological physics center's pelvis and thorax anthropomorphic quality assurance phantoms. *Med Phys*. 2007;34(6):2070-2076. doi:10.1118/1.2737158
8. Molineu A, Hernandez N, Nguyen T, Ibbott G, Followill D. Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom. *Med Phys*. 2013;40(2):022101. doi:10.1118/1.4773309
9. McVicker D, Yin F-F, Adamson JD. On the sensitivity of TG-119 and IROC credentialing to TPS commissioning errors. *J Appl Clin Med Phys*. 2016;17(1).
10. Kry SF, Molineu A, Kerns JR, Faught AM, Huang JY, Pulliam KB, Tonigan J, Alvarez P, Stingo F, Followill DS. Institutional patient-specific IMRT QA does not predict unacceptable plan delivery. *Int J Radiat Oncol Biol Phys*. 2014;90(5). doi:10.1016/j.ijrobp.2014.08.334
11. Nelms BE, Zhen H, Tomé WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys*. 2011;38(2):1037-1044. doi:10.1118/1.3544657

12. Nelms BE, Chan MF, Jarry G, Lemire M, Lowden J, Hampton C, Feygelman V. Evaluating IMRT and VMAT dose accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels. *Med Phys*. 2013;40(11):111722. doi:10.1118/1.4826166
13. Steers JM, Fraass BA. IMRT QA: Selecting gamma criteria based on error detection sensitivity. *Med Phys*. 2016;43(4):1982. doi:10.1118/1.4943953
14. Starkschall G, Steadham RE, Popple RA, Ahmad S, Rosen II. Beam-commissioning methodology for a three-dimensional convolution/superposition photon dose algorithm. *J Appl Clin Med Phys*. 2000;1(1):8-27. doi:10.1120/JACMP.V1I1.2651
15. Fogliata A, Nicolini G, Clivio A, Vanetti E, Cozzi L. Accuracy of Acuros XB and AAA dose calculation for small fields with reference to RapidArc® stereotactic treatments. *Med Phys*. 2011;38(11):6228-6237. doi:10.1118/1.3654739
16. Kerns JR, Stingo F, Followill DS, Howell RM, Melancon A, Kry SF. Treatment Planning System Calculation Errors Are Present in Most Imaging and Radiation Oncology Core-Houston Phantom Failures. *Int J Radiat Oncol Biol Phys*. 2017;98(5):1197-1203. doi:10.1016/j.ijrobp.2017.03.049
17. Ibbott GS, Haworth A, Followill DS. Quality assurance for clinical trials. *Front Oncol*. 2013;3:311. doi:10.3389/fonc.2013.00311
18. Ibbott GS, Followill DS, Molineu HA, Lowenstein JR, Alvarez PE, Roll JE. Challenges in Credentialing Institutions and Participants in Advanced Technology Multi-institutional Clinical Trials. *Int J Radiat Oncol Biol Phys*. 2008;71(1 SUPPL.):71-75. doi:10.1016/j.ijrobp.2007.08.083
19. Kirby TH, Hanson WF, Gastorf RJ, Chu CH, Shalek RJ. Mailable TLD system for photon and electron therapy beams. *Int J Radiat Oncol*. 1986;12(2):261-265. doi:10.1016/0360-3016(86)90107-0
20. Niroomand-Rad A, Blackwell CR, Coursey BM, Gall KP, Galvin JM, McLaughlin WL, Meigooni AS, Nath R, Rodgers JE, Soares CG. Radiochromic film dosimetry: recommendations of AAPM Radiation Therapy Committee Task Group 55. American Association of Physicists in Medicine. *Med Phys*. 1998;25(11):2093-2115.
21. Davidson SE, Popple RA, Ibbott GS, Followill DS. Technical note: Heterogeneity dose calculation accuracy in IMRT: study of five commercial treatment planning systems using an anthropomorphic thorax phantom. *Med Phys*. 2008;35(12):5434-5439. doi:10.1118/1.3006353
22. Kirby TH, Hanson WF, Johnston DA. Uncertainty analysis of absorbed dose calculations from thermoluminescence dosimeters. *Med Phys*. 1992;19(6):1427-1433.
23. Followill DS, Kry SF, Qin L, Leif J, Molineu A, Alvarez P, Aguirre JF, Ibbott GS. The Radiological Physics Center's standard dataset for small field size output factors. *J Appl*

*Clin Med Phys.* 2012;13(5):282-289. doi:10.1120/jacmp.v13i5.3962

24. Kerns JR, Followill DS, Lowenstein J, Molineu A, Alvarez P, Taylor PA, Kry SF. Agreement Between Institutional Measurements and Treatment Planning System Calculations for Basic Dosimetric Parameters as Measured by the Imaging and Radiation Oncology Core-Houston. *Int J Radiat Oncol Biol Phys.* 2016;95(5):1527-1534. doi:10.1016/j.ijrobp.2016.03.035
25. McKenzie EM, Balter PA, Stingo FC, Jones J, Followill DS, Kry SF. Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans. *Med Phys.* 2014;41(12):121702. doi:10.1118/1.4899177
26. Dong L, Antolak J, Salehpour M, Forster K, O'Neill L, Kendall R, Rosen I. Patient-specific point dose measurement for IMRT monitor unit verification. *Int J Radiat Oncol.* 2003;56(3):867-877. doi:10.1016/S0360-3016(03)00197-4
27. Aland T, Kairn T, Kenny J. Evaluation of a Gafchromic EBT2 film dosimetry system for radiotherapy quality assurance. *Australas Phys Eng Sci Med.* 2011;34(2):251-260. doi:10.1007/s13246-011-0072-6
28. Mohan R, Arnfield M, Tong S, Wu Q, Siebers J. The impact of fluctuations in intensity patterns on the number of monitor units and the quality and accuracy of intensity modulated radiotherapy. *Med Phys.* 2000;27(6):1226-1237. doi:10.1118/1.599000
29. McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys.* 2010;37(2):505-515. doi:10.1118/1.3276775
30. Younge KC, Matuszak MM, Moran JM, McShan DL, Fraass BA, Roberts DA. Penalization of aperture complexity in inversely planned volumetric modulated arc therapy. *Med Phys.* 2012;39(11):7160-7170. doi:10.1118/1.4762566
31. Masi L, Doro R, Favuzza V, Cipressi S, Livi L. Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy. *Med Phys.* 2013;40(7):071718. doi:10.1118/1.4810969
32. Crowe SB, Kairn T, Kenny J, Knight RT, Hill B, Langton CM, Trapp J V. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. *Australas Phys Eng Sci Med.* 2014;37(3):475-482. doi:10.1007/s13246-014-0274-9
33. Du W, Cho SH, Zhang X, Hoffman KE, Kudchadker RJ. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. *Med Phys.* 2014;41(2). doi:10.1118/1.4861821
34. Park JM, Park S-Y, Kim H, Kim JH, Carlson J, Ye S-J. Modulation indices for volumetric modulated arc therapy. *Phys Med Biol.* 2014;59(23):7315-7340. doi:10.1088/0031-9155/59/23/7315

35. Götstedt J, Hauer AK, Bäck A. Development and evaluation of aperture-based complexity metrics using film and EPID measurements of static MLC openings. *Med Phys*. 2015;42(7):3911-3921. doi:10.1118/1.4921733
36. Webb S. Use of a quantitative index of beam modulation to characterize dose conformality: illustration by a comparison of full beamlet IMRT, few-segment IMRT (fsIMRT) and conformal unmodulated radiotherapy. *Phys Med Biol*. 2003;48(14):2051-2062. doi:10.1088/0031-9155/48/14/301
37. Agnew CE, Irvine DM, McGarry CK. Correlation of phantom-based and log file patient-specific QA with complexity scores for VMAT. *J Appl Clin Med Phys*. 2014;15(6):204-216. doi:10.1120/jacmp.v15i6.4994
38. Crowe SB, Kairn T, Middlebrook N, Sutherland B, Hill B, Kenny J, Langton CM, Trapp J V. Examination of the properties of IMRT and VMAT beams and evaluation against pre-treatment quality assurance results. *Phys Med Biol*. 2015;60(6):2587-2601. doi:10.1088/0031-9155/60/6/2587
39. Hernandez V, Saez J, Pasler M, Jurado-Bruggeman D, Jornet N. Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy. *Phys Imaging Radiat Oncol*. 2018;5:37-43. doi:10.1016/j.phro.2018.02.002
40. Tonigan JR. Evaluation of intensity modulated radiation therapy (IMRT) delivery error due to IMRT treatment plan complexity and improperly matched dosimetry data. Published online 2011.
41. Rajasekaran D, Jeevanandam P, Sukumar P, Ranganathan A, Johnjothi S, Nagarajan V. A study on the correlation between plan complexity and gamma index analysis in patient specific quality assurance of volumetric modulated arc therapy. *Reports Pract Oncol Radiother*. 2015;20(1):57-65. doi:10.1016/J.RPOR.2014.08.006
42. McGarry CK, Agnew CE, Hussein M, Tsang Y, McWilliam A, Hounsell AR, Clark CH. The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. *Br J Radiol*. 2016;89(1057):S530. doi:10.1259/bjr.20150445
43. Kruse JJ. On the insensitivity of single field planar dosimetry to IMRT inaccuracies. *Med Phys*. 2010;37(6):2516-2524. doi:10.1118/1.3425781
44. Stojadinovic S, Ouyang L, Gu X, Pompoš A, Bao Q, Solberg TD. Breaking bad IMRT QA practice. *J Appl Clin Med Phys*. 2015;16(3). doi:10.1120/jacmp.v16i3.5242
45. Ezzell GA, Burmeister JW, Dogan N, Losasso TJ, Mechalakos JG, Mihailidis D, Molineu A, Palta JR, Ramsey CR, Salter BJ, Shi J, Xia P, Yue NJ, Xiao Y. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys*. 2009;36(11):5359-5373. doi:10.1118/1.3238104
46. Rangel A, Ploquin N, Kay I, Dunscombe P. Towards an objective evaluation of tolerances for beam modeling in a treatment planning system. *Phys Med Biol*. 2007;52(19):6011-



6025. doi:10.1088/0031-9155/52/19/020

47. Kron T, Clivio A, Vanetti E, Nicolini G, Cramb J, Lonski P, Cozzi L, Fogliata A. Small field segments surrounded by large areas only shielded by a multileaf collimator: comparison of experiments and dose calculation. *Med Phys*. 2012;39(12):7480-7489. doi:10.1118/1.4762564
48. Fogliata A, Lobefalo F, Reggiori G, Stravato A, Tomatis S, Scorsetti M, Cozzi L. Evaluation of the dose calculation accuracy for small fields defined by jaw or MLC for AAA and Acuros XB algorithms. *Med Phys*. 2016;43(10):5685-5694. doi:10.1118/1.4963219
49. Smilowitz JB, Das IJ, Feygelman V, Fraass BA, Kry SF, Marshall IR, Mihailidis DN, Ouhib Z, Ritter T, Snyder MG, Fairbrent L. AAPM Medical Physics Practice Guideline 5.a.: Commissioning and QA of Treatment Planning Dose Calculations - Megavoltage Photon and Electron Beams. *J Appl Clin Med Phys*. 2015;16(5):14-34. doi:10.1120/jacmp.v16i5.5768
50. IAEA. *TecDoc 1540: Specification and Acceptance Testing of Radiotherapy Treatment Planning Systems.*; 2007.
51. IAEA. *TecDoc 1583: Commissioning of Radiotherapy Treatment Planning Systems: Testing for Typical External Beam Treatment Techniques.*; 2008.
52. Das IJ, Cheng CW, Watts RJ, Ahnesjö A, Gibbons J, Li XA, Lowenstein J, Mitra RK, Simon WE, Zhu TC. Accelerator beam data commissioning equipment and procedures: Report of the TG-106 of the Therapy Physics Committee of the AAPM. *Med Phys*. 2008;35(9):4186-4215. doi:10.1118/1.2969070
53. Chen S, Yi BY, Yang X, Xu H, Prado KL, D'Souza WD. Optimizing the MLC model parameters for IMRT in the RayStation treatment planning system. *J Appl Clin Med Phys*. 2015;16(5).
54. Bedford JL, Thomas MDR, Smyth G. Beam modeling and VMAT performance with the Agility 160-leaf multileaf collimator. *J Appl Clin Med Phys*. 2013;14(2).
55. Kerns JR, Followill DS, Lowenstein J, Molineu A, Alvarez P, Taylor PA, Stingo FC, Kry SF. Technical Report: Reference photon dosimetry data for Varian accelerators based on IROC-Houston site visit data. *Med Phys*. 2016;43(5):2374-2386. doi:10.1118/1.4945697
56. Kerns JR, Followill DS, Lowenstein J, Molineu A, Alvarez P, Taylor PA, Kry SF. Reference dosimetry data and modeling challenges for Elekta accelerators based on IROC-Houston site visit data. *Med Phys*. 2018;45(5):2337-2344. doi:10.1002/mp.12865
57. Kumaraswamy LK, Schmitt JD, Bailey DW, Xu ZZ, Podgorsak MB. Spatial variation of dosimetric leaf gap and its impact on dose delivery. *Med Phys*. 2014;41(11):111711. doi:10.1118/1.4897572

58. Balasingh STP, Singh IRR, Rafic KM, Babu SES, Ravindran BP. Determination of dosimetric leaf gap using amorphous silicon electronic portal imaging device and its influence on intensity modulated radiotherapy dose delivery. *J Med Phys.* 2015;40(3):129-135. doi:10.4103/0971-6203.165072
59. Hernandez V, Vera-Sánchez JA, Vieillelveigne L, Saez J. Commissioning of the tongue-and-groove modelling in treatment planning systems: from static fields to VMAT treatments. *Phys Med Biol.* 2017;62(16):6688-6707. doi:10.1088/1361-6560/aa7b1a
60. Moran JM, Dempsey M, Eisbruch A, Fraass BA, Galvin JM, Ibbott GS, Marks LB. Safety considerations for IMRT: Executive summary. *Pract Radiat Oncol.* 2011;1(3):190-195. doi:10.1016/j.prro.2011.04.008
61. Palta JR, Deye JA, Ibbott GS, Purdy JA, Urie MM, Engler MJ, Rivard MJ. Credentialing of institutions for IMRT in clinical trials. *Int J Radiat Oncol Biol Phys.* 2004;59(4):1257-1259. doi:10.1016/j.ijrobp.2004.03.007
62. Koger B, Price R, Wang D, Toomeh D, Geneser S, Ford E. Impact of the MLC leaf-tip model in a commercial TPS: Dose calculation limitations and IROC-H phantom failures. *J Appl Clin Med Phys.* 2020;21(2):82-88. doi:10.1002/acm2.12819
63. Kim J, Han JS, Hsia AT, Li S, Xu Z, Ryu S. Relationship between dosimetric leaf gap and dose calculation errors for high definition multi-leaf collimators in radiotherapy. *Phys Imaging Radiat Oncol.* 2018;5:31-36. doi:10.1016/J.PHRO.2018.01.003
64. Kerns JR, Followill DS, Lowenstein J, Molineu A, Alvarez P, Taylor PA, Kry SF. Agreement Between Institutional Measurements and Treatment Planning System Calculations for Basic Dosimetric Parameters as Measured by the Imaging and Radiation Oncology Core-Houston. *Int J Radiat Oncol Biol Phys.* 2016;95(5):1527-1534. doi:10.1016/j.ijrobp.2016.03.035
65. Glenn MC, Peterson CB, Followill DS, Howell RM, Pollard-Larkin JM, Kry SF. Reference dataset of users' photon beam modeling parameters for the Eclipse, Pinnacle, and RayStation treatment planning systems. *Med Phys.* 2020;47(1):282-288. doi:10.1002/mp.13892
66. Glenn MC, Hernandez V, Saez J, Followill DS, Howell RM, Pollard-Larkin JM, Zhou S, Kry SF. Treatment plan complexity does not predict IROC Houston anthropomorphic head and neck phantom performance. *Phys Med Biol.* 2018;63(20):205015. doi:10.1088/1361-6560/aae29e
67. Kielar KN, Mok E, Hsu A, Wang L, Luxton G. Verification of dosimetric accuracy on the TrueBeam STx: Rounded leaf effect of the high definition MLC. *Med Phys.* 2012;39(10):6360-6371. doi:10.1118/1.4752444
68. Rangel A, Dunscombe P. Tolerances on MLC leaf position accuracy for IMRT delivery with a dynamic MLC. *Med Phys.* 2009;36(7):3304-3309. doi:10.1118/1.3134244

69. Luo W, Li J, Price RA, Chen L, Yang J, Fan J, Chen Z, McNeeley S, Xu X, Ma C-M. Monte Carlo based IMRT dose verification using MLC log files and R/V outputs. *Med Phys*. 2006;33(7):2557-2564. doi:10.1118/1.2208916
70. Lee J-W, Hong S, Kim Y-L, Choi K-S, Chung J-B, Lee D-H, Suh T-S. Effects of Static Dosimetric Leaf Gap on MLC-based Small Beam Dose Distribution for Intensity Modulated Radiosurgery. *J Appl Clin Med Phys*. 2007;8(4). doi:10.1007/978-3-540-36841-0\_456
71. Kielar KN, Mok E, Hsu A, Wang L, Luxton G. Verification of dosimetric accuracy on the TrueBeam STx: Rounded leaf effect of the high definition MLC. *Med Phys*. 2012;39(10):6360-6371. doi:10.1118/1.4752444
72. Yao W, Farr JB. Determining the optimal dosimetric leaf gap setting for rounded leaf-end multileaf collimator systems by simple test fields. *J Appl Clin Med Phys*. 2015;16(4):65-77. doi:10.1120/jacmp.v16i4.5321
73. Wasbø E, Valen H. Dosimetric discrepancies caused by differing MLC parameters for dynamic IMRT. *Phys Med Biol*. 2008;53(2):405-415. doi:10.1088/0031-9155/53/2/008
74. Edward S, Alvarez P, Taylor P, Molineu A, Peterson C, Followill D, Kry S. Differences in the patterns of failure between IROC lung and spine phantom irradiations. *Pract Radiat Oncol*. Published online 2020.
75. Izewska J, Hultqvist M, Bera P. Analysis of uncertainties in the IAEA/WHO TLD postal dose audit system. *Radiat Meas*. 2008;43(2-6):959-963. doi:10.1016/j.radmeas.2008.01.011
76. Jangda AQ, Hussein S. Validating dose rate calibration of radiotherapy photon beams through IAEA/WHO postal audit dosimetry service. *J Pak Med Assoc*. 2012;62(5). Accessed April 8, 2020. [https://ecommons.aku.edu/pakistan\\_fhs\\_mc\\_radiat\\_oncol/3](https://ecommons.aku.edu/pakistan_fhs_mc_radiat_oncol/3)

## Vita

Mallory Glenn (*née* Carson) was born in Sherman, Texas. After completing high school as salutatorian at The Colony High School, The Colony, Texas in 2011, she entered Texas A&M University in College Station, Texas. She graduated summa cum laude with a Bachelor of Science in radiological health engineering and a minor in mathematics from Texas A&M University in May, 2015. She then entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences to pursue a Ph.D. in medical physics.

Permanent address:

3805 Monterrey Circle

The Colony, Texas 75056, USA3