Robustness evaluation of RT plans and their dosimetric impact

Maria Alice Oliveira
Master in Medical Physics
Department of Physics and Astronomy
Faculty of Sciences of the University of Porto
2026

Robustness evaluation of RT plans and their dosimetric impact

Maria Alice Oliveira

Dissertation carried out as part of the Master in Medical Physics

Department of Physics and Astronomy

2026

Supervisor
Anabela Dias
Medical Physicist Expert
Instituto Português de Oncologia do Porto Francisco Gentil, E.P.E.

Co-supervisor
Susana Gonçalves,
Dosimetrist
Instituto Português de Oncologia do Porto Francisco Gentil, E.P.E.

IPOPORTO

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Acknowledgements

What a journey this has been, only made possible with the contribution of those I will always cherish and remember.

My most sincere thank you to my supervisor Dr. Anabela Dias for making sure I went the extra mile developing this work. For guiding and believing in me and for all the pragmatic advice and the empowerment words.

I am also deeply grateful to my co-supervisor Dr. Susana Gonçalves for inspiring me to work harder and smarter. For the critical insights that highly elevated this study and for always being there to motivate and help me.

A huge thank you to my third advisor Dr. Catarina Souto, a key contributor in this work. For the countless hours spent with me and the valuable feedback. For all the ideas and the unlimited dedication, for being so empathetic and for always making me feel grounded.

Forever grateful to have worked and learned so much with the most driven female trio.

Undeniably, my gratitude extends to all the Fisiquinhos and Físicos Médicos, alongside the dosimetry department in the IPOPFG. For all the energy, support, coffee breaks and lunches shared together. I would like to give a special thank you to Íris, for the precious help with the developed script and all her suggestions, besides all our conversations.

In addition, I am grateful for my classmates, especially Rafa, Bea and Carol, who kept me going through the rough times and turned this challenging journey truly chaotic and iconic in the best possible ways.

Then, for all the unconditional love and support, I have my friends and family to thank for, especially my parents and my sisters. For all the plans that made it out of the group chat and for always giving me a reason to smile at the end of the day.

Time flew by and I appreciate everyone who helped and inspired me in bringing the pieces of this document together. I really can’t thank you enough. Memories were created and everything was worth it!

Resumo

A crescente modulação e complexidade das técnicas modernas de radioterapia introduziu fontes adicionais de incerteza que devem ser consideradas durante o planeamento do tratamento. Consequentemente, a robustez, definida como a capacidade de um plano para resistir a incertezas, tornou-se numa componente fundamental do planeamento em radioterapia, particularmente em tratamentos hipofracionados, nos quais são administradas doses mais elevadas por fração. Assim, o objetivo deste trabalho é definir um conjunto de métricas de avaliação de robustez face a incertezas geométricas e mecânicas, estabelecer um índice de robustez e avaliar uma eventual correlação entre a complexidade e robustez dos planos.

Neste contexto, foi realizado um estudo retrospetivo com uma amostra de 41 doentes tratadas com Irradiação Parcial da Mama com a técnica IMRT. Tendo em conta a lateralidade da patologia, a amostra foi dividida em dois subgrupos para uma avaliação distinta da dose média recebida pelo coração. Para cada doente, foram criados sete cenários de perturbação, incluindo erros de posicionamento de ±3 mm em todas as direções translacionais e uma variação de 1 mm no DLG do MLC.

Foram utilizadas métricas baseadas nos DVH para avaliar a robustez do plano, analisando as variações nos *clinical goals* ($V_{95\%}$ do PTV, dose média no coração e $V_{30\%}$ do pulmão ipsilateral) entre a distribuição de dose planeada e a perturbada. Os planos foram considerados robustos quando os objetivos clínicos se mantiveram dentro dos limites de variação definidos em todos os cenários. Posteriormente, foi desenvolvido um *script* para avaliar a robustez através da análise dos vóxeis das imagens de planeamento, complementando a análise baseada nos DVH. Através do desvio padrão calculado para cada voxel, foram estabelecidas barreiras de robustez para as três estruturas avaliadas. Além disso, a distribuição de dose presente nos mapas de robustez permitiu a visualização das regiões mais vulneráveis às incertezas simuladas, caracterizadas por valores mais elevados de desvio padrão, localizadas predominantemente em áreas com acentuados gradientes de dose, na periferia das estruturas avaliadas.

Através de métricas baseadas em DVH e análise de vóxeis, foi definido um índice de robustez para avaliação da relação entre a complexidade e robustez dos planos. No entanto, não foi encontrada nenhuma correlação estatisticamente significativa,

sugerindo que tratamentos complexos, como IMRT de mama, podem permanecer clinicamente aceitáveis para os cenários de incerteza simulados.

De uma forma geral, os planos de tratamento revelaram-se robustos face a incertezas mecânicas e de posicionamento, indicando que o impacto dosimétrico das perturbações simuladas não comprometeu a qualidade do plano, sendo os casos não robustos justificados pela anatomia específica da doente e pela localização do leito tumoral. Estes resultados sustentam o atual nível de ação de 3 mm para a correção de desvios de posicionamento em planos de IMRT de mama, definido no protocolo do IPOPFG.

Palavras-chave: Robustez, complexidade, qualidade, mama, IMRT, radioterapia externa

Abstract

The increasing modulation and complexity of modern radiotherapy techniques introduced additional sources of uncertainty that should be considered during treatment planning. As a result, robustness, defined as a plan's capacity to handle real-world uncertainties, has become a key component of radiotherapy planning, particularly for hypofractionated treatments, where higher doses are delivered per fraction. Therefore, the aim of this work is to define a set of robust evaluation metrics against geometric and mechanical uncertainties, to establish a robustness index and to assess a potential correlation between plan complexity and robustness.

Within this context, a cohort of 41 patients treated for Partial Breast Irradiation with IMRT was selected. Right and left-sided cases were evaluated separately for mean heart dose analysis. For each patient, seven perturbation scenarios were created, including patient setup errors of ±3 mm in all translational directions and a 1 mm variation in the MLC Dosimetric Leaf Gap. A DVH-based metrics was used for assessing plan robustness, analysing variations in clinical goals (PTV $V_{95\%}$, Heart D_{mean} and Ipsilateral lung $V_{30\%}$) between planned and perturbed dose distribution. Plans were considered robust when clinical goals remained within the accepted variation limits across all scenarios. A script was then developed to assess robustness at the voxel level for the PTV, heart and ipsilateral lung, complementing the DVH-based analysis. Robustness thresholds were defined at the voxel level, using the dose standard deviation (SD) calculated for each voxel across the three evaluated structures. Additionally, robustness maps with a 3D dose distribution enabled visualization of the most vulnerable regions, those with higher standard deviation values, which were predominantly located at boundary regions with steep dose gradients. Combining both DVH- and voxel-based metrics, a robustness index was defined to evaluate the relationship between plan complexity and robustness. No significant correlation was found, suggesting that complex treatments such as breast IMRT can remain clinically acceptable under the simulated uncertainty scenarios.

Overall, treatment plans proved robust against mechanical and setup uncertainties, indicating that the simulated perturbations did not compromise the plan quality, with non-robust cases justified by patient anatomy and tumour bed location. This finding supports maintaining the current 3 mm action level for correcting setup deviations during patient positioning in breast IMRT plans, established in the IPOPFG internal protocol.

Keywords: Robustness, complexity, quality, breast, IMRT, external beam radiotherapy

Table of Contents

List of Tables

List of Graphics

List of Figures

List of Abbreviations

3D-CRT	THREE-DIMENSIONAL CONFORMAL RADIOTHERAPY
AAA	ANISOTROPIC ANALYTICAL ALGORITHM
ACROP	ADVISORY COMMITTEE IN RADIATION ONCOLOGY PRACTICE
ALARA	AS LOW AS REASONABLY ACHIEVABLE
BEV	BEAM'S EYE VIEW
CBCT	CONE BEAM COMPUTED TOMOGRAPHY
CI	CONFORMITY INDEX
CT	COMPUTED TOMOGRAPHY
CTV	CLINICAL TARGET VOLUME
DIBH	DEEP INSPIRATION BREATH HOLD
DLG	DOSIMETRIC LEAF GAP
DVH	DOSE VOLUME HISTOGRAM
DVHBW	DOSE VOLUME HISTOGRAM BAND WIDTH
EBRT	EXTERNAL BEAM RADIOTHERAPY
EM	EDGE METRIC
ESTRO	EUROPEAN SOCIETY FOR RADIOTHERAPY AND ONCOLOGY
FCUP	FACULTY OF SCIENCES OF THE UNIVERSITY OF PORTO
GBCI	GLOBAL BREAST CANCER INITIATIVE
GI	GRADIENT INDEX
GTV	GROSS TUMOUR VOLUME
HI	HOMOGENEITY INDEX
HT	HEALTHY TISSUE
ICRU	INTERNATIONAL COMMISSION ON RADIATION UNITS AND MEASUREMENTS
IGRT	IMAGE GUIDED RADIOTHERAPY
MF	MODULATION FACTOR
IMRT	INTENSITY MODULATED RADIOTHERAPY
IPOPFG	INSTITUTO PORTUGUÊS DE ONCOLOGIA DO PORTO FRANCISCO GENTIL
LTF	LEAF TRANSMISSION FACTOR
LINAC	LINEAR ACCELERATOR
MLC	MULTI-LEAF COLLIMATOR

NTCP	NORMAL TISSUE COMPLICATION PROBABILITY
OAR	ORGANS AT RISK
PBI	PARTIAL BREAST IRRADIATION
PTV	PLANNING TARGET VOLUME
QA	QUALITY ASSURANCE
RON	REGISTO ONCOLÓGICO NACIONAL
RT	RADIOTHERAPY
SD	STANDARD DEVIATION
SDVH	STANDARD DEVIATION DOSE VOLUME HISTOGRAM
SGRT	SURFACE-GUIDED RADIOTHERAPY
SIB	SIMULTANEOUS INTEGRATED BOOST
TCP	TUMOUR CONTROL PROBABILITY
TPS	TREATMENT PLANNING SYSTEM
VMAT	VOLUMETRIC MODULATED ARC THERAPY
WHO	WORLD HEALTH ORGANIZATION

Introduction

Breast cancer remains a major public health challenge, with its incidence continuing to rise worldwide. As a key component of breast cancer treatment, advanced radiotherapy techniques such as intensity modulated radiotherapy (IMRT) provide superior dose conformity and treatment precision compared with conventional 3D-CR. However, the increased complexity and modulation associated with these approaches make radiotherapy (RT) plans more vulnerable to treatment related uncertainties. Therefore, robustness is a cornerstone of modern radiotherapy, as it assesses a plan's capacity to handle real world uncertainties. Rather than relying on dose distributions optimized for static conditions, robustness ensures target coverage and OAR sparing despite motion, positioning, calibration errors and anatomical changes that may occur during treatment delivery.

This retrospective study evaluates the robustness of partial breast IMRT plans against geometric and mechanical uncertainties, with the primary objective of developing an evaluation framework that integrates dose-volume histogram (DVH) and voxel-based metrics to establish a quantitative robustness index. Since breast radiotherapy is tailored to each patient's individual clinical situation, a detailed dose distribution analysis was performed to identify which anatomical regions are most vulnerable to uncertainties arising during treatment delivery, and to determine which error scenarios may compromise treatment quality and clinical outcomes. Additionally, a robustness index was proposed, to quantitatively assess plan robustness, and a complexity analysis was conducted to investigate a potential association between plan complexity and robustness. Ultimately, this dissertation aims to provide deeper insight into the dosimetric impact of uncertainties on partial breast intensity modulated radiotherapy plans.

The present section introduces the research topic and outlines the scope of the work. Chapter 1 provides an overview of breast cancer, including a brief statistical analysis of the disease and a clinical decision guide for breast cancer treatment. The fundamental principles of the linear accelerator operation are also discussed, with particular emphasis on the multileaf collimator (MLC), intensity-modulated radiotherapy techniques, inverse treatment planning, and the evaluation of treatment plans within the treatment planning system (TPS). Then, Chapter 2 provides a literature review on the state-of-the-art regarding robustness and complexity in radiotherapy, which serves as the basis for defining the evaluation metrics adopted in this work. Chapter 3 describes the

methodology implemented in this study, followed by the analysis and discussion of the results in Chapter 4, together with future perspectives and research directions. Finally, Chapter 5 draws the main conclusions of the work. Bibliographic references cited throughout this dissertation are listed at the end of the document, followed by the attachments, including the implemented scripts, the abstract accepted for poster presentation at the 6th European Congress of Medical Physics, and a summary table with the results of this study.

1. Breast cancer

According to the World Health Organization (WHO), cancer cases are projected to reach 35 million in 2050, a 75% increase from the 20 million recorded in 2022 (World Health Organization, 2024). Among women, breast cancer remains the most prevalent malignancy. In response to this growing burden, the WHO established the Global Breast Cancer Initiative (GBCI), with the goal of reducing annual breast cancer mortality by 2.5% per year (World Health Organization, 2021). This target reflects an ambition to save 2.5 million lives over a 20-year period, from 2020 to 2040. To achieve this cumulative goal, an average of 125 000 lives would need to be saved each year (equation 1).

$$\frac{2\,500\,000\ \text{lives}}{20\ \text{years}} = 125\,000\ \text{lives per year} \qquad \text{eq. 1}$$

It is important to note that the 2.5% target refers to a reduction in the mortality rate, not a fixed absolute number, meaning the annual lives saved will vary as the baseline shifts over time.

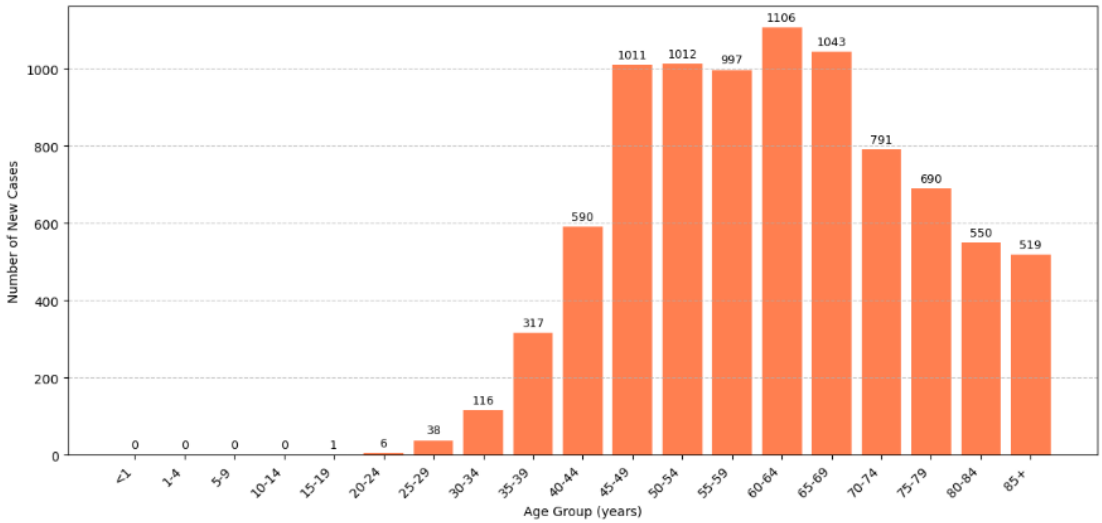
Portugal illustrates the scale of the challenge at a national level. In 2021, cancer was the second leading cause of death in the country, accounting for 27 577 deaths. According to the RON (*Registo Oncológico Nacional*), Portugal's National Cancer Registry, a total of 60 717 new cases were diagnosed among the resident population of 10 343 066 in that same year, the most recent period for which data are available (RON, 2024). Table 1 presents the distribution of new cases incidence rates (per 100 000 person years) for breast and for the total tumours combined, across all apparatus, systems, and organs.

Table 1 - Number of cases and incidence rates (/100 000 person-years) of breast cancer in Portugal.

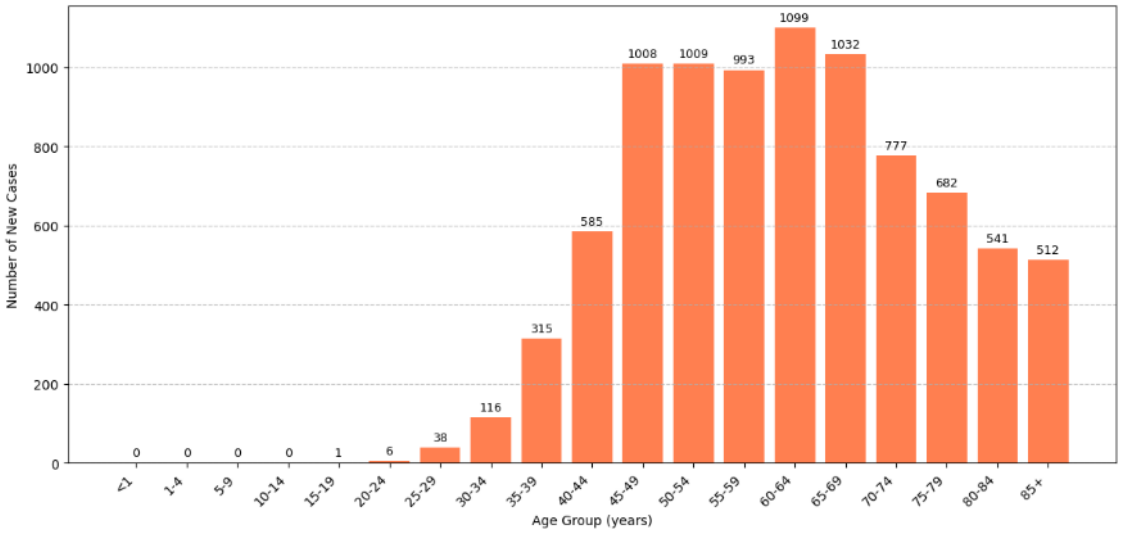
	Male		**Female**		**Total**	
	Number	Rate	Number	Rate	Number	Rate
Breast	73	1.5	8 714	160.7	8 787	85.0
Total	33 107	672.9	27 610	509.1	60 717	587

In females, breast accounted for 31.6% of all tumours, with 8 714 new cases, corresponding to an incidence rate of 160.7 per 100 000 person-years.

Breast cancer incidence shows a strong age-dependent pattern, which is illustrated in Graphs 1 and 2. Graph 1 presents the overall distribution of breast cancer incidence across all age groups, while graph 2 offers a more detailed breakdown, focusing only on female breast cancer cases over the same timeframe.



Graphic 1 - Distribution of breast cancer cases by age group – Portugal, 2021.



Graphic 2 - Distribution of female breast cancer cases by age group – Portugal, 2021.

The data confirm a strong age-dependent pattern in breast cancer incidence. New cases are negligible below the age of 30, rising sharply from 35 onwards and peaking in post-menopausal cohorts, particularly between the ages of 60 and 64.

Incidence alone, however, offers an incomplete scenario. Table 2 broadens the analysis by presenting breast cancer incidence, mortality, and the resulting mortality-to-incidence ratio across both genders, enabling a more comprehensive assessment of disease burden.

Table 2 - Breast cancer mortality, incidence, and mortality-to-incidence ratio.

	Male			Female			Total		
	Mortality	Incidence	Ratio (%)	Mortality	Incidence	Ratio (%)	Mortality	Incidence	Ratio (%)
Breast	24	73	32.9	1 793	8 714	20.6	1 817	8 787	20.7

As expected, most cases occurred in women (8 714), consistent with the preceding analysis. Male breast cancer, by contrast, represented only 0.83% of all annual diagnoses, a stark reminder that the disease, though rare, is not exclusive to women.

Taken together, the incidence and mortality data highlight the substantial burden of the disease. With 1 817 recorded deaths, the overall mortality-to-incidence ratio is 20.7%, confirming breast cancer as a major public health challenge. Addressing this burden requires sustained investment in prevention, early diagnosis, and effective multimodal treatment, within which radiotherapy plays a central role in locoregional management.

1.1 Radiotherapy

The main objective of radiotherapy is to maximize target dose while sparing healthy tissues and organs at risk (OAR), ensuring that the prescribed dose is effectively delivered to the patient. As illustrated in Figure 1, external beam radiotherapy (EBRT) is frequently used to treat breast cancer, where the breast alone, or any of the axilla, supraclavicular or internal mammary lymph node chain can be irradiated (Byrne et al., 2016). In EBRT, three main techniques are commonly used, 3D Conformal Radiotherapy (3D-CRT), Intensity Modulated Radiotherapy (IMRT) and Volumetric Modulated Arc Therapy (VMAT). All these techniques are predominantly delivered using a linear accelerator.

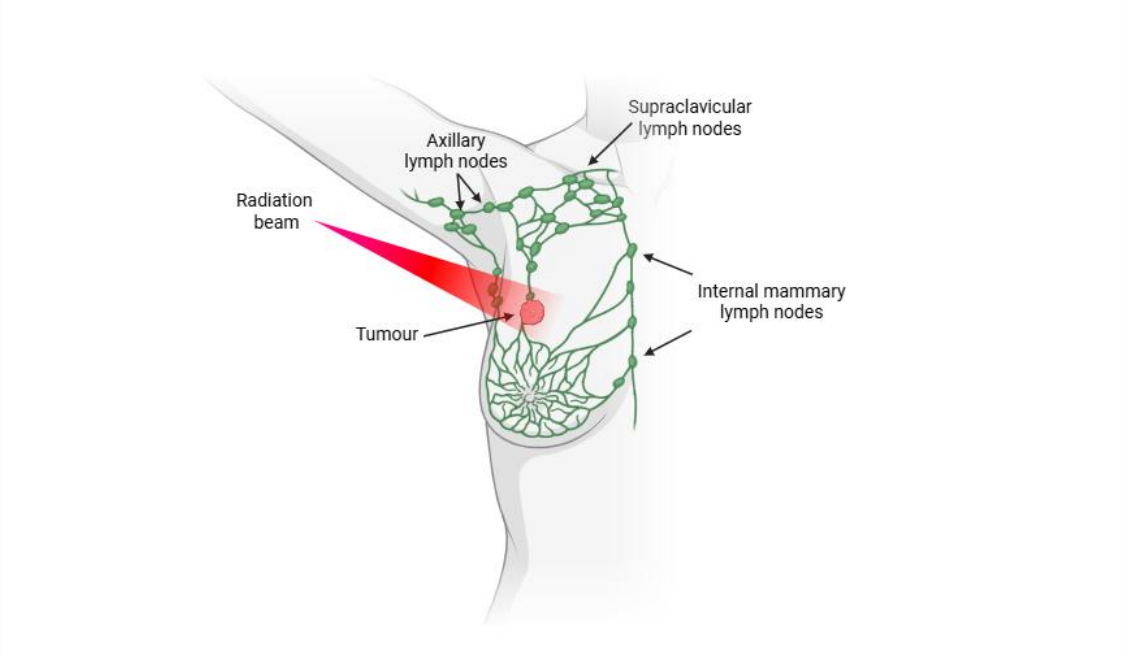


Figure 1 - Illustration of the female breast anatomy showing a tumour targeted by a radiation beam, alongside the axillary, supraclavicular, and internal mammary lymph nodes. Created in https://BioRender.com.

1.2 Breast treatment

Radiotherapy plays a crucial role in breast cancer management following both breast-conserving surgery or mastectomy, with a treatment decision guide scheme presented in Figure 2. Over the past decade, several documents endorsed by the European Society for Radiotherapy and Oncology (ESTRO) and its Advisory Committee on Radiation Oncology Practice (ACROP) have reshaped clinical practice. Treatment decisions are influenced by factors such as tumour characteristics, the type of surgery performed, lymph node status, and the patient's overall risk profile. Radiation is directed to the breast, chest wall, and/or regional lymph nodes with the aim of eliminating residual microscopic disease and lowering the likelihood of cancer recurrence.

The ESTRO-ACROP consensus on patient selection and dose/fractionation (Meattini et al., 2022) established that moderately hypofractionated whole breast irradiation (approximately 40–42.5 Gy in 15–16 fractions) should be the standard approach for the large majority of patients, despite age, tumour stage, biological subtype, margin status, or the need for a tumour bed boost. The panel also concluded that an ultra-hypofractionated regimen (26 Gy in five fractions) can be offered either as standard practice or within prospective trials for selected patients receiving breast or chest wall irradiation without nodal treatment.

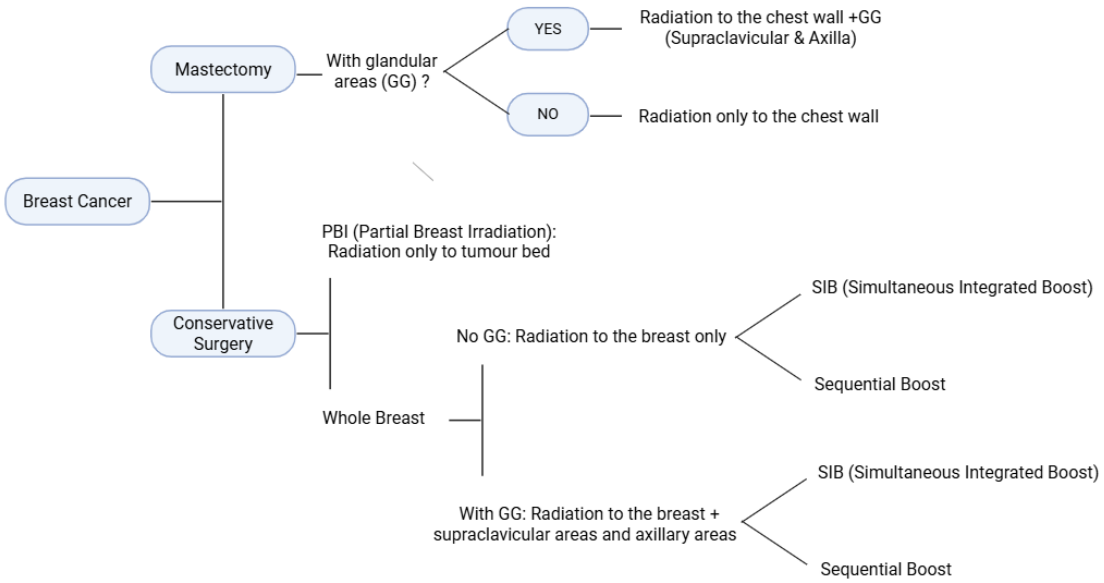


Figure 2 - Radiotherapy scheme for breast cancer treatment decision guide, created in https://BioRender.com.

For early-stage patients, partial breast irradiation (PBI) was recommended for carefully selected low-risk cases, while regional nodal irradiation and chest wall irradiation after mastectomy follow the same general hypofractionation principles, adapted to individual risk. In PBI, radiation goes only to the tumour bed, the local area where the tumour was removed. For Whole Breast Irradiation (WBI), it is important to note if glandular areas are included and how the boost is delivered. It can either be delivered after the initial treatment phase (sequential boost) or concomitantly during the daily breast radiation sessions (Simultaneous Integrated Boost).

When a mastectomy has been performed, radiation is directed solely to the chest wall, if no glandular areas are involved. Otherwise, it encompasses both the chest wall and glandular regions, with particular emphasis on the supraclavicular and axillary areas.

Regardless of the irradiated volume, accurate treatment delivery also relies on the reproducible positioning of the patient throughout the treatment course. To minimise patient movement and improve setup consistency, immobilisation devices are routinely employed according to the treatment site. Nevertheless, uncertainties remain due to setup variations and anatomical changes. Therefore, the precise quantification and mitigation of both systematic and random positioning errors is essential, as these directly influence the determination of PTV margins. To account for setup uncertainties related to patient positioning and anatomical variations, daily image guidance using orthogonal planar imaging (2D–2D) or cone-beam computed tomography (CBCT) is recommended.

Among the available image-guided radiotherapy (IGRT) techniques, surface-guided radiotherapy (SGRT) has emerged as an advanced approach, employing three-dimensional non-ionising optical systems to provide submillimetre, real-time monitoring of patient surface topology. This enables marker-free patient positioning and intrafraction motion management, provided that the system has been appropriately validated against conventional radiographic IGRT methods (Mast et al., 2023).

1.3 Linear accelerator

The linear accelerator (linac) is the bedrock of modern EBRT, comprising an ensemble of interacting components distributed across a stationary gantry stand and a rotating gantry, as illustrated in Figure 3 (Koka et al., 2022).

Microwave power is generated within the gantry stand, which houses the pulse modulator and the radiofrequency (RF) power source, which can be either a klystron or a magnetron. The choice between the two generally depends on the machine's energy specifications. Magnetrons function as RF oscillators and are typically used in lower energy linacs (4-6 MV), whereas klystrons function as RF power amplifiers and are preferred for generating the higher power required in high energy linacs (up to 25 MV). The pulse modulator supplies high-voltage pulses to the RF source. The generated microwaves are then transmitted to the gantry through an RF waveguide system. A circulator positioned along this path protects the magnetron or klystron by diverting reflected microwave energy and preventing it from propagating back toward the source.

Within the rotating gantry, an electron gun injects electrons into the accelerating waveguide, where they are accelerated by the oscillating microwave field to the required kinetic energy, typically ranging from 4 to 25 MeV. After leaving the accelerating structure, the electron beam is guided through the electron transport system. In high-energy linacs, the length of the accelerating waveguide requires the use of a bending magnet to redirect the beam toward the treatment head. The electrons then strike a high-Z target, producing a high energy bremsstrahlung photon beam, with therapeutic energies usually ranging from 4 to 25 MV.

The photon beam subsequently enters the treatment head, where it first passes through an ionization chamber that continuously monitors beam output and delivered dose rate. Before reaching the patient on the treatment couch, the beam must be shaped to match the target volume. This task is primarily performed by the multileaf collimator (MLC), shown in detail in the inset of Figure 3. The MLC consists of two opposing banks of independently movable tungsten leaves (Leaf Bank A and Leaf Bank B). The range of

beam shapes that can be produced depends on the leaf width and travel limits, which are defined by parameters such as the maximum midline overtravel and maximum field size. By adjusting the position of each leaf independently, the MLC enables the radiation field to closely conform to the tumour shape while reducing irradiation of surrounding healthy tissues.

Although not shown in Figure 3, the operation of a linac also relies on auxiliary systems such as vacuum pumps and water-cooling circuits. These systems ensure stable operating conditions and contribute to the safe and reliable delivery of radiation treatments.

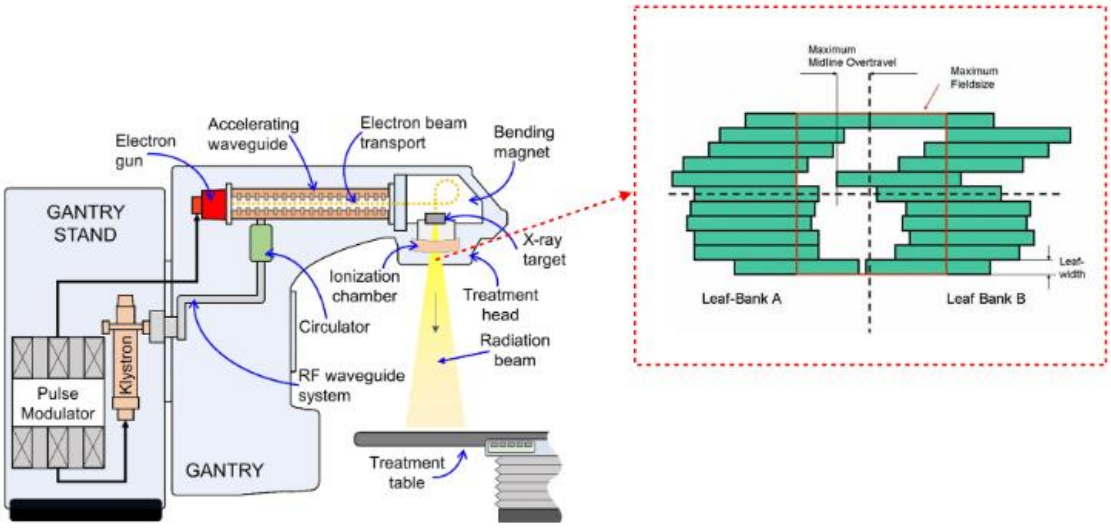


Figure 3 - Schematic representation of a linac's beam generation system, with a detailed view of the multi leaf collimator (MLC) used for beam shaping. Adapted from (Handoyo et al., 2026) and (Schlegel, Grosser, et al., 2006). Created in https://BioRender.com.

As a key component of the linac, MLC modelling within the Treatment Planning System (TPS) is critical for reliable dose calculation, particularly in dynamic delivery techniques such as IMRT and VMAT (Enomoto et al., 2024).

Depending on the TPS and its specific version, users may be restricted in which MLC properties they can configure. For instance, within Varian's Eclipse™ TPS versions prior to v18, the only editable MLC properties are the MLC transmission factor and the dosimetric leaf gap (DLG). Both parameters play a fundamental role in the accuracy of dose calculation for modulated deliveries.

The MLC leaves are arranged in opposing pairs, with each pair capable of moving independently to shape the radiation field. As depicted in Figure 4, Varian MLCs feature

rounded leaf ends. However, the dose calculation algorithms within the Eclipse™ TPS approximate these ends as rectangular. Therefore, to compensate for this geometric discrepancy, the DLG parameter is introduced, further refined through the application of the leaf transmission factor (LTF). During fluence calculation, the algorithm adjusts the position of each leaf tip by shifting it back by half the DLG value. Consequently, the effective gap between a fully closed leaf pair corresponds to the DLG value, thereby accounting for the additional radiation transmitted through the rounded leaf ends (Rostami et al., 2024).

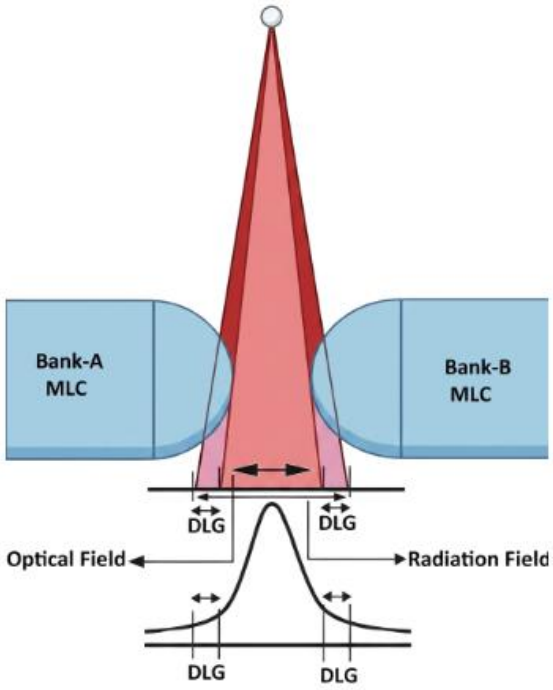


Figure 4 - Dosimetric Leaf Gap, introduced parameter to compensate for the rounded-leaf ends of the MLC leaves. Adapted from (Shende & Patel, 2017) and created in https://BioRender.com.

MLC transmission refers to the fraction of radiation penetrating through closed leaves, comprising both intraleaf transmission (through the leaf body) and interleaf leakage between adjacent leaves. Their correct characterisation is particularly relevant in breast radiotherapy, where MLC modelling inaccuracies have been shown to introduce clinically significant dose deviations to organs at risk, including the contralateral breast and the ipsilateral lung (Ju et al., 2021).

1.4 EBRT workflows

The EBRT workflow follows a sequence of interconnected steps. Regarding 3D-CRT associated with forward planning (Figure 5), a volumetric computed tomography (CT) is acquired after positioning the patient and, if needed, fitted with an immobilization device to ensure consistency across sessions. Then, target volumes and organs at risk are delineated, beam orientations and apertures are designed to conform to the target, and a three-dimensional dose distribution is computed. The plan is iteratively refined until approved by the radiation oncologist before delivery.

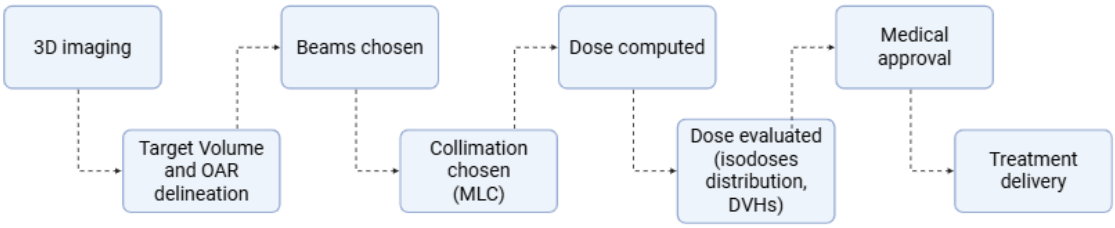


Figure 5 – EBRT forward planning workflow, created in https://BioRender.com.

In inverse planning strategies, used in IMRT and VMAT techniques, the workflow shifts, as highlighted in Figure 6. Following standard 3D imaging and contouring, clinicians define target dose prescriptions and OAR constraints. A computerized algorithm then optimizes beam fluences and MLC trajectories to meet these specific goals. Finally, the 3D dose is computed, iteratively evaluated via DVHs and isodoses distribution, approved, and verified through quality assurance (QA) tests before delivery.

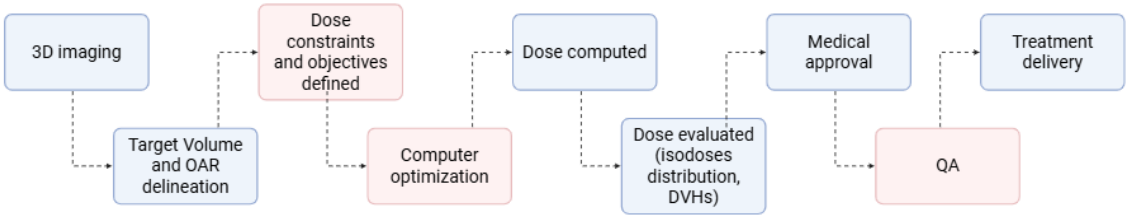


Figure 6 - EBRT inverse planning workflow, created in https://BioRender.com.

1.4.1 3D-CRT

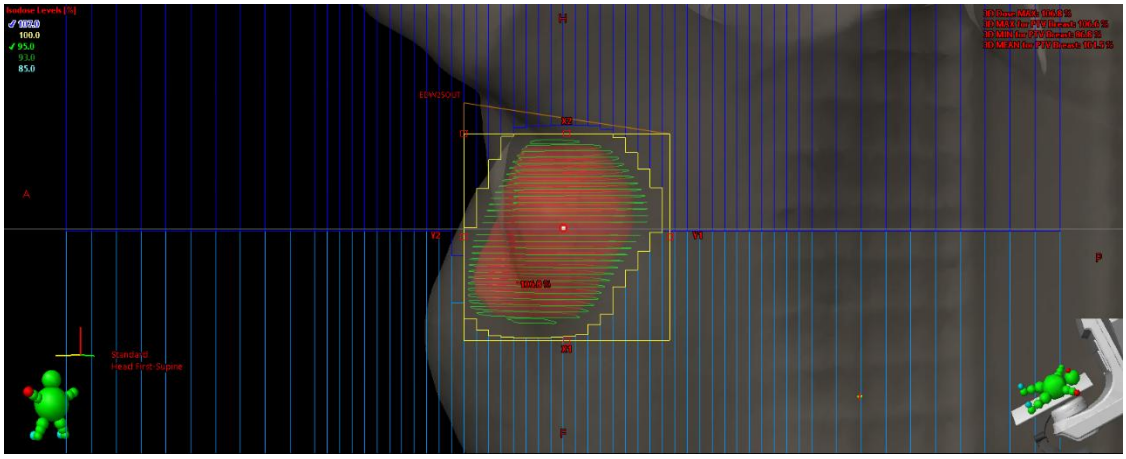
The introduction of CT in 1972 marked a pivotal turning point, enabling far more accurate tumour localization (Mayles, Philip, 2007). This technological breakthrough, alongside subsequent advances in medical imaging, laid the foundation for 3D-CRT. Its widespread implementation in modern radiotherapy facilities made it possible to deliver high radiation doses precisely conformed to the target tumour while sparing surrounding healthy

tissues. As a result, treatment margins could be reduced, facilitating strategic dose escalation and thereby increasing the probability of local tumour control (TCP) while minimising normal tissue complications (NTCP) (Schlegel et al., 2006).

The clinical effectiveness of 3D-CRT relies on both imaging, planning, and delivery. Radiation beams are shaped to match the Planning Target Volume (PTV) from the perspective of the beam's eye view (BEV), typically achieved using an MLC.

For breast cancer specifically, 3D-CRT often employs tangential fields that aim to encompass the mammary tissue while minimizing the volume of the underlying lung and heart. A common refinement is the field-in-field technique, which uses small sub-segments to improve dose homogeneity, as represented in Figure 7. By reducing "hot spots" in thinner areas of the breast, such as the apex, this method ensures a more uniform dose distribution compared to traditional physical wedges (Podgorski, E. B., 2005). However, even with this refinement, 3D-CRT remains limited by its fixed beam angles, which can result in uneven dose delivery and greater exposure to the heart, lungs, and contralateral breast.

a)



b)

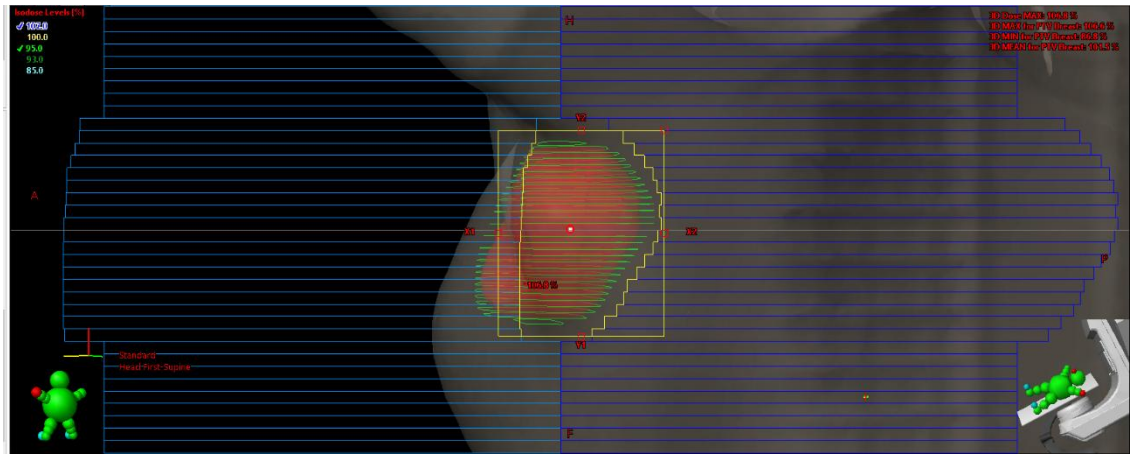


Figure 7 - Representation of the field-in-field technique for left-sided partial breast irradiation. (a) shows the primary 3D CRT treatment field, whereas (b) shows a sub-segment created within the primary field to improve dose homogeneity.

1.4.2 IMRT

IMRT represents a significant evolution from 3D-CRT, and can reduce lung and heart doses, as reported by Rastogui et al. While 3D-CRT uses beams of uniform intensity, IMRT employs nonuniform radiation fluences delivered from multiple gantry positions. This allows for the creation of high-gradient dose distributions that conform closely to the target volume, even in highly irregular or concave geometries (Rastogi et al., 2018).

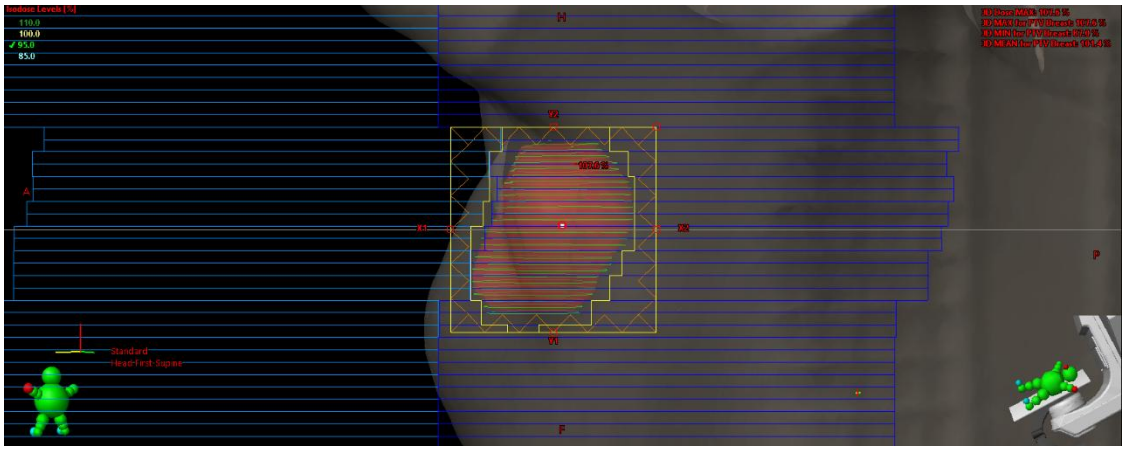
The primary clinical goal of IMRT is to improve the therapeutic ratio, escalating tumour dose to improve local control while simultaneously minimizing radiation exposure to adjacent OAR (Mayles, Philip, 2007).

A fundamental trade-off of IMRT is that the steep dose gradients responsible for its high degree of dose conformity also increase its sensitivity to geometric uncertainties. Consequently, even small setup errors or inter-fraction anatomical changes may result in discrepancies between the planned and delivered dose distributions.

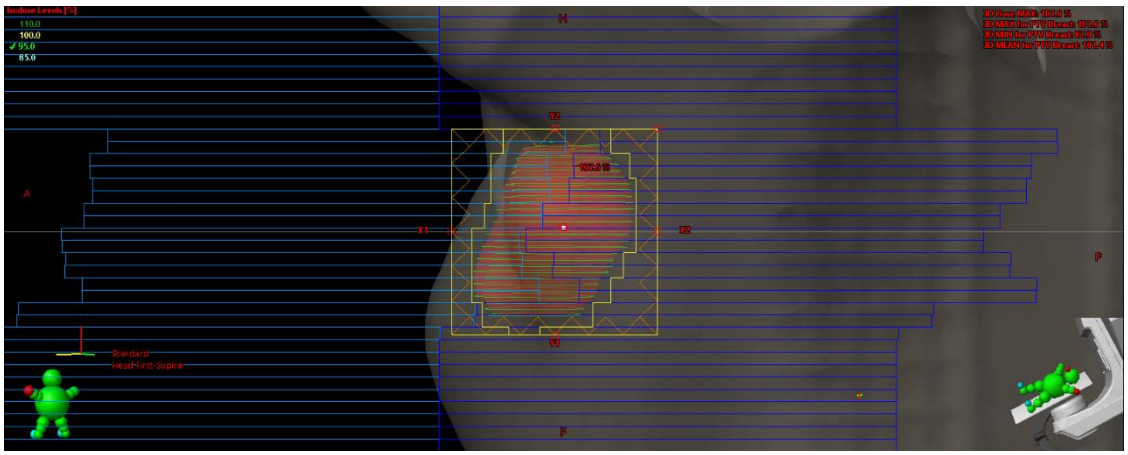
Although applying safety margins around the target volume partially mitigates this risk, relying solely on them may be insufficient in scenarios with large or unpredictable geometric variations. Robust planning strategies address this limitation directly by optimizing the treatment plan across a defined range of plausible uncertainties. Rather than relying solely on margin expansion, robust optimization ensures that the dose distribution remains clinically acceptable even under worst-case geometric scenarios, thereby providing a more reliable guarantee of target coverage and OAR sparing throughout the treatment course.

At its core, IMRT relies on the dynamic modulation provided by the MLC, as presented in the following Figure 8.

a)



b)



c)

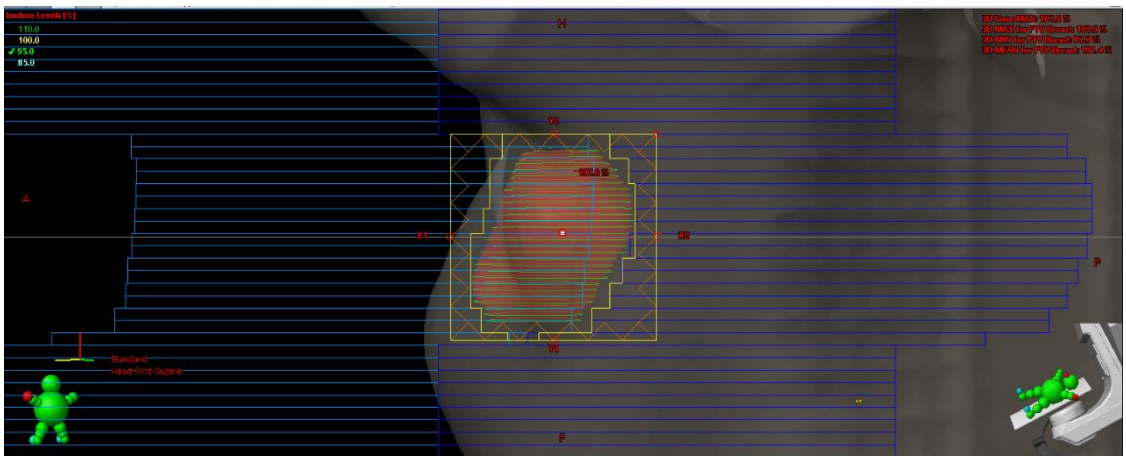


Figure 8 - Dynamic beam modulation via MLC. The sequence (a-c) illustrates the varying positions of the tungsten leaves during an IMRT delivery. This dynamic movement allows for precise control of radiation intensity, conforming the beam to the shape of the target volume while actively shielding the adjacent organs at risk.

Consequently, precise leaf positioning is perhaps the most essential MLC requirement. As IMRT uses numerous small field segments, even a 1 mm error in leaf placement can cause a dose discrepancy of approximately 10% in a 10×10 mm field (Podgorsak, E. B., 2005).

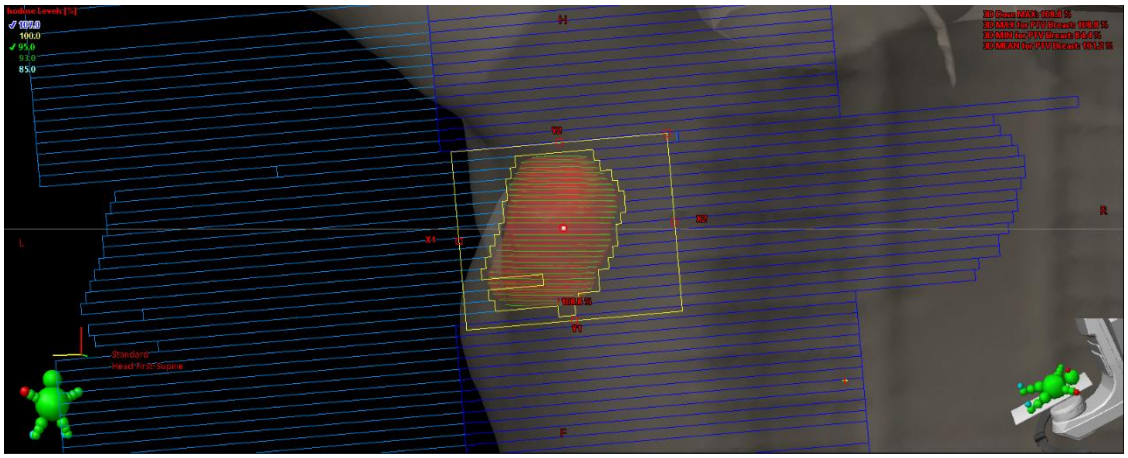
There are two distinct operational modes for MLC-based IMRT delivery, each differing in how leaf motion and beam activation are coordinated. In the step-and-shoot or static mode, the leaves move to a specific position, the beam is delivered, and the beam then turns off while the leaves move to the next position. In contrast, the dynamic mode or sliding window allows the MLC leaves to move continuously while the radiation beam is active, resulting in a more fluid modulation of intensity.

In breast radiotherapy, IMRT offers superior dose homogeneity compared to conventional techniques by modulating beam intensity across the target volume (Chang et al., 2022). This is particularly advantageous for left-sided lesions, where precise dose sculpting minimizes the risk of late cardiac toxicity by sparing the heart and lungs. Furthermore, IMRT enables advanced delivery strategies such as Accelerated Partial Breast Irradiation (APBI) and simultaneous integrated boost (SIB) technique, while optimizing OAR sparing. However, these dosimetric benefits depend on geometric accuracy. Due to high dose gradients and specific patient anatomy, even minor positioning deviations can lead to significant dosimetric errors. Consequently, rigorous patient setup and continuous image-guided verification are essential to account for inter-fraction motion and anatomical changes, ensuring that the theoretical advantages of IMRT are reliably translated into clinical outcomes (Mayles et al., 2007).

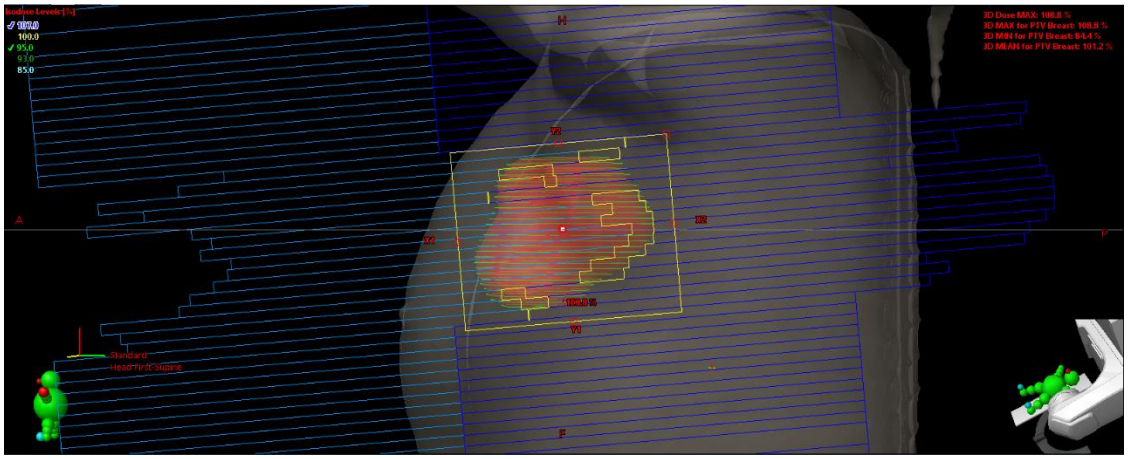
1.4.3 VMAT

VMAT has become a highly efficient technique for delivering intensity-modulated radiation. In contrast to 3D-CRT and IMRT, which relies on fixed gantry positions, VMAT administers radiation as the gantry rotates continuously around the patient. As shown in Figure 9, this approach enables simultaneous adjustments in gantry speed, dose rate, and the positions of the MLC leaves, resulting in highly conformal dose distributions that effectively shape the radiation dose around critical structures.

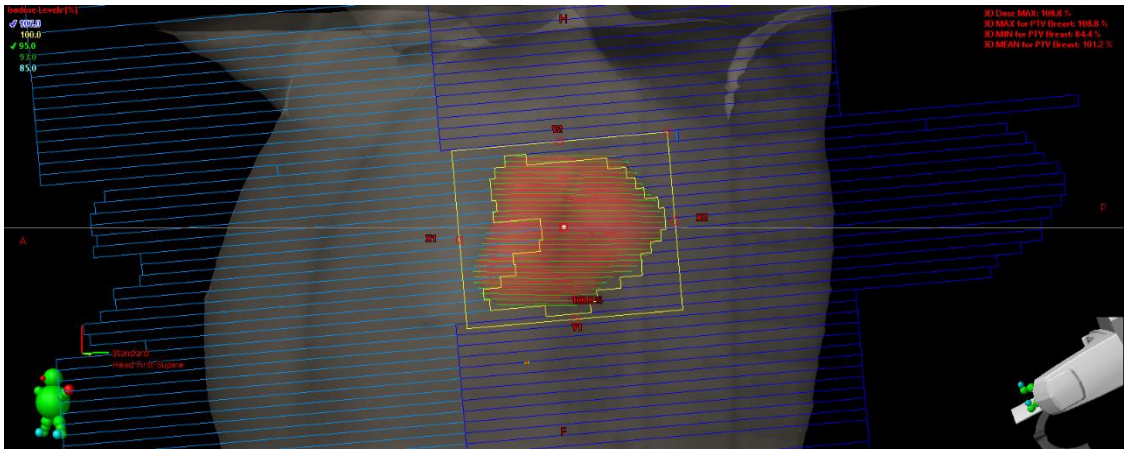
a)



b)



c)



d)

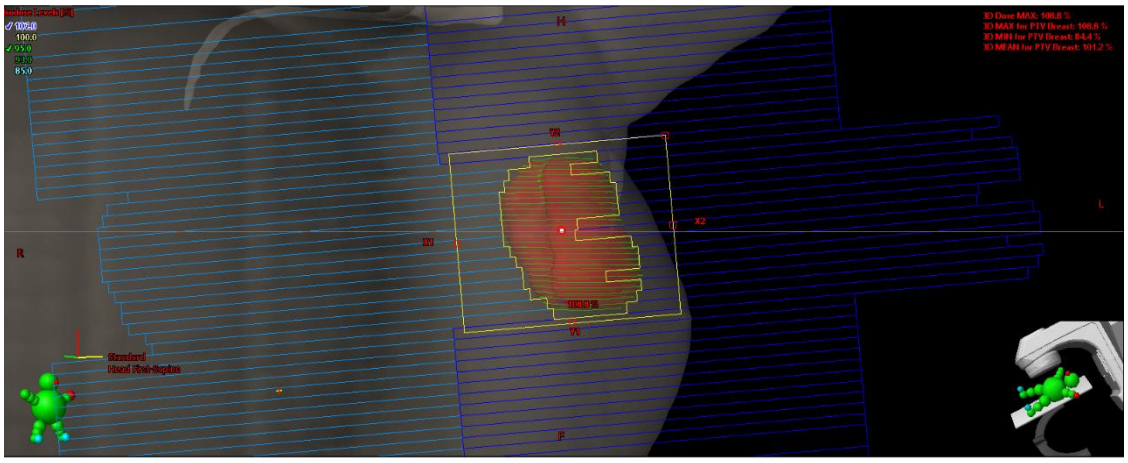


Figure 9 - VMAT delivery. The sequence (a–d) illustrates the continuous rotation of the gantry around the patient while the MLC leaves dynamically change position. This combined motion allows the radiation beam to conform to the target volume throughout the arc while reducing the dose delivered to adjacent OAR.

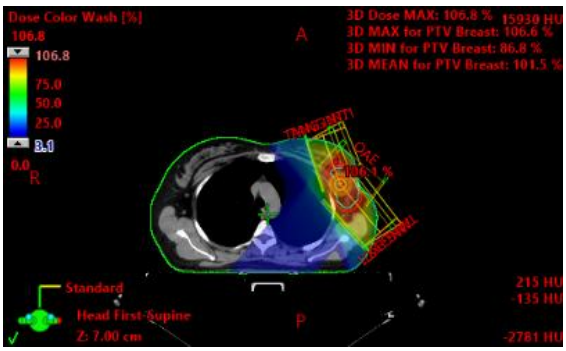
In breast treatment, VMAT is especially beneficial for complex cases, such as irradiation of the internal mammary nodes or when the target volume has a concave geometry near the heart. This technique often improves target coverage and produces a steeper dose gradient, although it can also increase the integral dose, meaning a larger volume of healthy tissue may receive very low levels of radiation (Schlegel, Bortfeld, et al., 2006).

The accuracy of these techniques depends largely on the MLC. In 3D-CRT, the leaves remain stationary to block and protect healthy tissues, whereas in VMAT, they continuously adjust their positions as the gantry rotates, allowing modulation of the beam intensity. The leaf width, typically between 5 mm and 10 mm at the isocenter, defines the level of detail achievable in beam shaping. This level of precision is essential for sparing small, sensitive structures near the breast, such as the coronary arteries (Podgorsak, E. B., 2005).

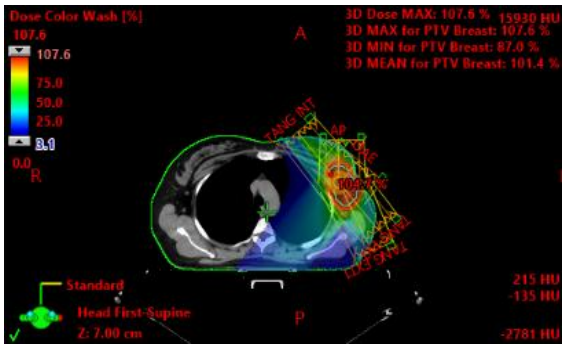
In addition to its geometric and dosimetric advantages, VMAT is intrinsically linked to inverse planning strategies. This allows for highly conformal dose distributions, although it introduces a trade-off with an increased volume of normal tissue receiving low-dose radiation, potentially raising concerns regarding secondary malignancies. Clinically, VMAT is also valued for its delivery efficiency, as treatments are typically faster than conventional fixed-field techniques, reducing patient discomfort and minimizing intra-fraction motion (Teoh et al., 2011). However, in breast radiotherapy, literature suggests that this efficiency may come at a cost. Modern techniques such as IMRT and VMAT, while improving target conformity and dose homogeneity, have been associated with

higher low dose baths to surrounding normal tissues compared to 3D-CRT (Zhang et al., 2020). In particular, Hunte et al. reported an increase in mean heart dose and low-dose volumes to the lungs and contralateral breast compared to 3D-CRT (Hunte et al., 2022), a finding corroborated by Das Majumdar et al., who similarly observed that inverse planning techniques worsened low-dose irradiation of the heart, lung, and contralateral breast relative to 3D-CRT, as highlighted in Figure 10 (Das Majumdar et al., 2022).

a)



b)



c)

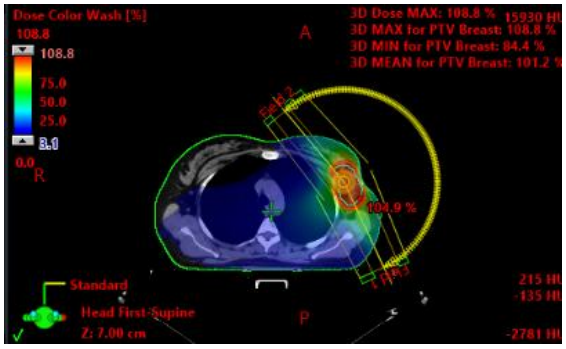


Figure 10 - Comparison of dose distributions obtained with three radiotherapy techniques for left-sided partial breast irradiation: 3D-CRT (top), IMRT (middle) and VMAT (bottom), illustrating the improved target dose conformity and the broader low-dose bath to the broader low-dose bath to the surrounding healthy tissues associated with IMRT and VMAT compared with 3D-CRT.

Hence, for left-sided cases, EBRT is frequently combined with respiratory management techniques such as deep inspiration breath hold (DIBH) to further reduce cardiac exposure (Tyran et al., 2015).

1.5 TPS parameters

While the continuous evolution of radiotherapy techniques has substantially improved delivery precision, it has concurrently increased planning complexity. This added complexity introduces new sources of uncertainty that can compromise treatment quality and clinical outcomes. As a result, the critical analysis of TPS parameters, including target volumes and structures delineation, dose volume histograms (DVHs), conformity indices, and the achievement of clinical goals has become essential to ensure treatment safety. Additionally, a thorough evaluation of both plan robustness and complexity is therefore fundamental to a safe and reliable treatment process.

1.5.1 Target Volumes

In radiotherapy planning, target volumes are defined according to the framework established by the International Commission on Radiation Units and Measurements (ICRU). Within this framework, the Gross Tumour Volume (GTV) represents the macroscopic extent of the tumour. The Clinical Target Volume (CTV) is defined by expanding the GTV to account for potential microscopic disease spread that is not detectable on imaging. The PTV encompasses the CTV with an additional margin to compensate for uncertainties in treatment delivery, such as patient setup variability and organ motion, thereby ensuring adequate coverage of the entire CTV by the prescribed dose. It is equally important to identify healthy tissues and organs that may be exposed to significant radiation doses during treatment, known as OAR. The following Figure 11 shows the Eclipse™ TPS contoured volumes for a left-sided partial breast irradiation.

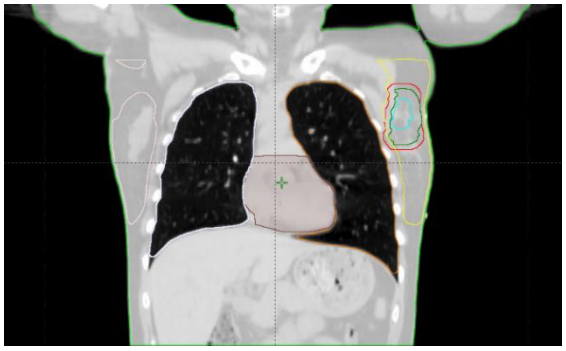


Figure 11 - Target volumes and delineated structures for a left-sided PBI treatment plan in Eclipse™. The treatment volumes are shown in blue (GTV), dark green (CTV), and red (PTV). The delineated anatomical structures include the body contour (light green), heart (brown), left lung (orange), right lung (lilac), ipsilateral breast (yellow), and contralateral breast (pink).

Once the target volumes and surrounding OAR have been clearly defined, understanding how the treatment is delivered becomes essential. Figure 12 illustrates the field arrangement for the same PBI case previously shown in Figure 11. From this spatial perspective, the integration of the anatomy with the treatment beams is observed, allowing a clear visualization of how the MLC conforms the radiation beam to the PTV while actively shielding the delineated heart.

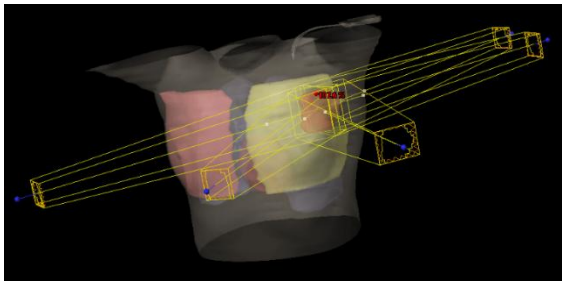


Figure 12 - Treatment field arrangement for a left-sided PBI plan in Eclipse™. The delineated anatomical structures include the PTV (red), ipsilateral breast (yellow), contralateral breast (pink), heart (brown), and lungs (blue). MLC-defined apertures are overlaid, illustrating the beam conformity to the PTV.

1.5.2 DVH

While target volumes and BEV analysis provide essential information regarding treatment geometry and beam arrangement, the evaluation of dose distribution also requires quantitative dosimetric metrics. In this context, dose volume histograms (DVHs) represent a fundamental tool for assessing treatment plan quality.

A DVH is a quantitative tool that condenses the three-dimensional dose distribution into a two-dimensional graph, indicating the percentage or volume of a given structure that receives at least a certain dose (Podgorsak, E. B., 2005).

Fundamentally, a DVH is constructed by subdividing each contoured structure into small volume elements (voxels), each assigned a specific dose value by the TPS. The most widely used format in clinical practice is the cumulative DVH, as presented in Figure 13, which plots the volume of a structure, expressed either as a percentage or in absolute terms, receiving at least a specified dose.

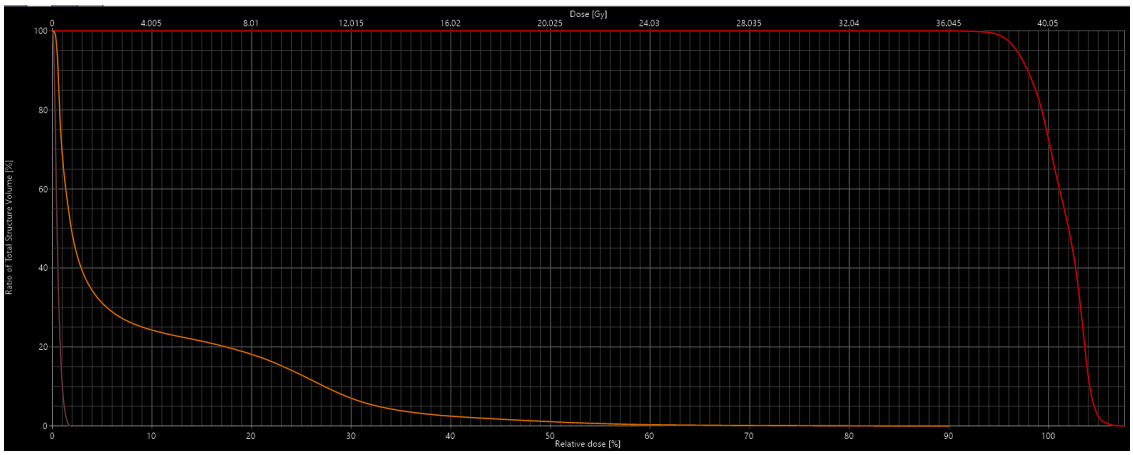


Figure 13 - Cumulative dose volume histogram (DVH). The red curve illustrates the target coverage for the PTV, while the orange and brown curves show the dose received by the organs at risk (ipsilateral lung and heart, respectively).

However, by reducing a three-dimensional dose distribution to a single curve, all spatial information is inherently lost. Consequently, DVH analysis should be complemented by conformity indices and a visual inspection of the dose distribution. Without this spatial context, a DVH cannot reveal whether hotspots within an organ at risk coincide with highly radiosensitive substructures, or if underdosed regions within a target are clustered or scattered.

1.5.3 Conformity indices

In IMRT, treatment plan evaluation extends beyond DVHs to include quantitative indices that measure dose conformity and uniformity within the target volume. Metrics such as the conformity index (CI), homogeneity index (HI), and gradient index (GI) serve as essential complements to visual isodose verification. However, their mathematical definitions have evolved over time, and a universally accepted formulation has yet to be established.

The CI expresses the ratio between the volume enclosed by the prescription isodose surface and the PTV volume, where a value of 1 indicates perfect conformity (equation 2). Values above or below 1 reflect over coverage and under coverage of the target, respectively (Feuvret et al., 2006).

$$\mathrm{CI} = \frac{\mathrm{V_{PI}}}{\mathrm{V_{PTV}}} \qquad \text{eq. } 2$$

The HI, as defined by ICRU Report 83 (ICRU, 2010), quantifies dose uniformity within the PTV according to equation 3, where $\mathrm{D_{2\%}}$, $\mathrm{D_{98\%}}$ and $\mathrm{D_{50\%}}$are the doses received by 2%, 98% and 50% of the PTV volume, respectively. A value approaching zero corresponds to a highly homogeneous dose distribution.

$$\mathrm{HI} = \frac{\mathrm{D_{2\%}} - \mathrm{D_{98\%}}}{\mathrm{D_{50\%}}} \qquad \text{eq. } 3$$

Since neither the CI nor the HI reflects the dose distribution beyond the PTV, the GI is used to assess the dose fall-off outside the target volume, defined as the ratio of the volume of tissue receiving half of the prescribed dose to the volume of tissue receiving the full prescribed dose (equation 4).

$$\mathrm{GI} = \frac{\mathrm{PIV_{half}}}{\mathrm{PIV}} \qquad \text{eq. } 4$$

A lower GI indicates a steeper dose gradient and greater sparing of surrounding HT, an important consideration in PBI given the proximity of the heart, ipsilateral lung, and contralateral breast (Correa et al., 2017).

The dosimetric indices (CI, HI, and GI) can be complemented by radiobiological metrics, such as tumour control probability (TCP) and normal tissue complication probability (NTCP), which provide a more comprehensive evaluation of treatment plan quality by incorporating the expected biological response of both tumour and surrounding tissues and are increasingly being adopted in radiotherapy plan assessment (Wang et al., 2019).

1.5.4 Clinical goals

While dosimetric and radiobiological metrics provide valuable quantitative information, treatment plans must also satisfy predefined clinical goals to be considered clinically acceptable. Clinical goals are dose constraints established to ensure adequate target coverage while limiting radiation exposure to organs at risk (OARs), thereby reducing the probability of treatment-related toxicity and long-term complications. Exceeding these constraints may increase the risk of cardiac morbidity for left-sided treatments, radiation

induced lung toxicity such as pneumonitis and pulmonary fibrosis, and unnecessary irradiation of the contralateral breast, which may increase the risk of radiation-induced secondary malignancies.

This work focused on Partial Breast Irradiation (PBI) using IMRT, with the corresponding clinical goals varying depending on tumour laterality. In this fractionation regimen, total prescribed dose is usually 40 Gy, administered in 15 fractions. The clinical goals for both left- and right-sided cases are presented in tables 3 and 4.

Table 3 - Clinical goals for left-sided PBI using IMRT.

PTV PBI	P1	D_{max} < 110%
		V 105% ≤ 5%
		V 95% ≥ 95%
Left Breast - PTV	P2	V 50% ≤ 50%
		V 95% ≤ 25%
Heart	P2	D_{mean} < 1.50 Gy
		V 15% ≤ 5%
Left Lung	P2	V 30% ≤ 10%
Healthy Tissue	P3	D_{max} < 112%
		V 110% ≤ 2 cm³
Right Lung	P3	V 10% ≤ 5%

Table 4 - Clinical goals for right-sided PBI using IMRT.

PTV PBI	P1	D_{max} < 110%
		V 105% ≤ 5%
		V 95% ≥ 95%
Right Breast - PTV	P2	V 50% ≤ 50%
		V 95% ≤ 25%
Heart	P2	D_{mean} < 0.7 Gy
		V 15% ≤ 5%
Right Lung	P2	V 30% ≤ 10%
Healthy Tissue	P3	D_{max} < 112%
		V 110% ≤ 2 cm³
Left Lung	P3	V 10% ≤ 5%

Note that P1 is the highest priority among the clinical goals, sequentially followed by P2 and P3. Additionally, D_{max} stands for maximum dose, which cannot exceed 110% of the prescribed dose for the PTV, whereas V 105% ≤ 5% indicates that the volume receiving at least 105% of the prescribed dose must not exceed 5% of the total target volume.

The principal difference resides in heart clinical goals, which are more restrictive for left-sided cases due to the increased proximity of the heart to the treatment field.

Meeting these clinical goals is crucial for ensuring both the efficacy and safety of PBI. Given the inherent challenge of simultaneously achieving adequate target coverage and maximizing OAR protection, conventional 3D-CRT techniques frequently prove insufficient. Therefore, modern dosimetric standards require advanced planning and delivery systems, which form the foundation of the state-of-the-art breast radiotherapy.

2. Robust optimization state-of-the-art

2.1 Plan robustness in radiotherapy

Robustness evaluation is a cornerstone of modern radiotherapy, as it assesses a plan's capacity to handle real-world uncertainties.

As noted by Chan et al., "recently, the notion of plan robustness, which refers to the extent of a radiotherapy plan remaining stable for maintaining the desirable radiation dose even with uncertainties, has emerged in radiotherapy. Usually, a plan is considered robust when its dosimetric changes are minimal in error scenarios" (Chan et al., 2023). We can also argue that a treatment plan is considered robust if the resulting dose distribution remains relatively insensitive to minor delivery inaccuracies (Schlegel, Bortfeld, et al., 2006).

The importance of robustness has been highlighted by several authors. Kim et al. pointed out that "advances in radiotherapy have led to increasingly conformal and complex treatment plans. The progressive reduction in safety margins around the target volume and the increased use of hypo fractionated radiotherapy further heighten their vulnerability to systematic geometric uncertainties, which may compromise target volume coverage and increase doses to normal tissues. Evaluating treatment plan robustness, therefore, is crucial to ensuring the safe and effective delivery of radiotherapy" (Kim et al., 2025). As Yock et al. define it, robustness is the resilience of the dose distribution to inherent uncertainties, quantifying how consistently clinical objectives, such as target coverage and organ sparing are maintained despite real world perturbations. While standard inverse planning algorithms often achieve ideal nominal scenarios, the modulation introduced by MLC motion can lead to highly complex treatment plans that may lack the stability required for actual delivery (Yock et al., 2019). Meanwhile, Hernandez et al. emphasize that the significance of robustness lies in its capacity to mitigate anatomical deformations, setup uncertainties, and mechanical tolerances. Since the delivered dose distribution is highly dependent on these factors, robustness is critical for treatment plan quality. A clinically robust plan ensures that the DVH maintains stability and remains within acceptable limits across a range of uncertainty scenarios, preventing minor delivery deviations from compromising the resulting dose distribution (Hernandez et al., 2020).

Furthermore, robustness is becoming a non-negotiable tool for assessing RT plans. While traditional evaluation often focuses on the nominal scenario where uncertainties

are not assessed, Lobner et al. note that patient and machine related uncertainties, such as MLC positioning, significantly influence the delivered dose. Consequently, robustness analysis establishes an essential framework, ensuring that a plan is both deliverable and resilient as required for proper treatments and better patient outcomes (Loebner et al., 2025).

2.2 TPS robustness evaluation tools

Despite its growing importance, robustness evaluation has not yet been widely integrated into commercial treatment planning systems (TPS), with most platforms offering limited or no dedicated tools for this purpose. Since its introduction by RaySearch in the treatment planning system, this feature has provided different tools for a detailed analysis of plan robustness. The conventional margin-based planning technique often fails to guarantee robustness against uncertainties, particularly for heterogeneous regions, such as the thorax. Hence, it becomes even more important to address this robustness and incorporate it into each patient's individual treatment plan, ensuring realistic treatments and better outcomes.

RayStation® TPS supports robust optimization across multiple delivery techniques, including dynamic multileaf collimation (DMLC), VMAT, TomoTherapy, and actively scanned proton and carbon ion beams (RaySearch Laboratories, 2019a). The underlying optimization algorithm follows a minimax formulation, in which the objective functions selected for robustness are optimized against the worst-case scenario across all predefined uncertainty conditions. To achieve this, the software generates a discrete set of scenarios by combining user-defined setup and density uncertainties, as illustrated in Figure 14, ensuring that the resulting treatment plan remains clinically acceptable under all considered conditions. Setup uncertainties can be specified uniformly across all beams as an alternative to conventional planning target volume (PTV) margin-based planning, or defined independently for each beam or isocenter, enabling advanced strategies such as robust field matching and robust skin flash for complex treatment geometries.

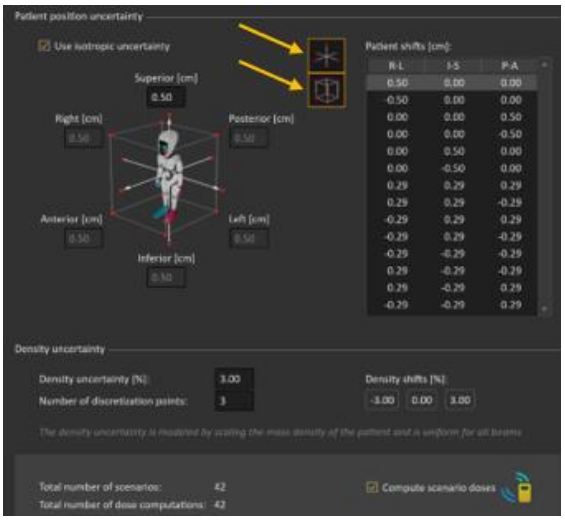
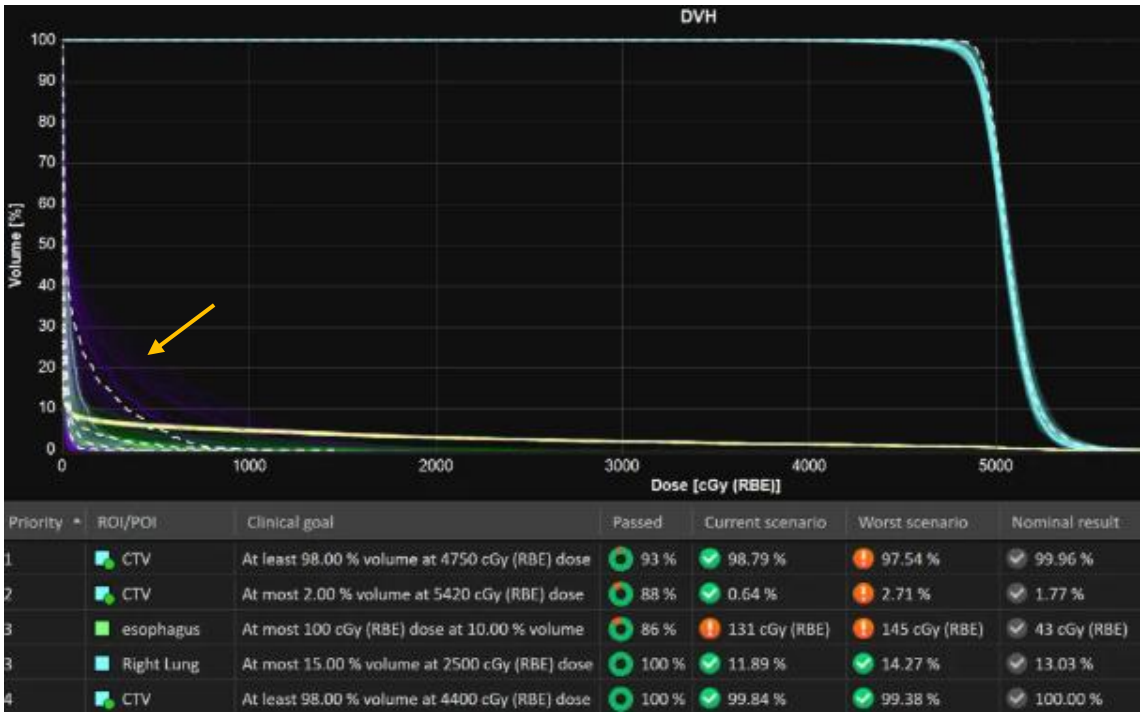


Figure 14 - Patient position and density uncertainty settings in the RayStation® robust optimization module. A 5 mm isotropic setup uncertainty is applied uniformly across all beams (upper arrow), generating 14 patient shift scenarios across the right-left, inferior-superior, and posterior-anterior directions. Alternatively, setup uncertainties can be defined independently per beam or isocenter (lower arrow). A density uncertainty of 3% is additionally considered, yielding three discrete density shift values, resulting in a total of 42 error scenarios and 42 dose computations. Adapted from (RaySearch Laboratories, 2019b).

RayStation's robust optimization tools extend beyond static geometric uncertainties by incorporating intrafractional organ motion via 4D-CT, which is highly relevant for respiratory-gated treatments. The system allows setup and density uncertainties to be simultaneously optimized alongside multiple planning CT datasets and material overrides, enabling planners to account for random anatomical variations like bowel gas. The clinical efficacy of these tools in improving dose stability and resilience has been widely demonstrated across proton and photon modalities (Byrne et al., 2016).

Plan robustness can be assessed independently of the optimization approach using the dedicated Robust Evaluation module, which evaluates treatment plans across multiple error scenarios (RaySearch Laboratories, 2019a). Three complementary metrics are available within the module, namely DVH clusters, clinical goal evaluation, and three-dimensional dose distributions, including voxel-wise minimum and maximum dose maps (Figure 15). Clinical goals established during treatment planning remain accessible throughout the analysis, with additional indicators showing the percentage of scenarios meeting each criterion. Individual scenario dose distributions can also be compared side by side with the nominal plan, and dose difference maps can be generated to facilitate the identification of deviations.

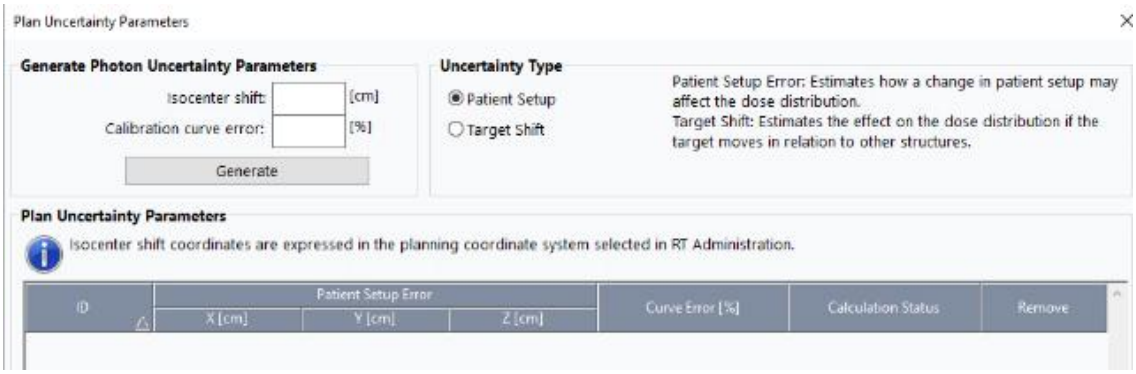


Priority	ROI/POI	Clinical goal	Passed	Current scenario	Worst scenario	Nominal result
1	CTV	At least 98.00 % volume at 4750 cGy (RBE) dose	93 %	98.79 %	97.54 %	99.96 %
2	CTV	At most 2.00 % volume at 5420 cGy (RBE) dose	88 %	0.64 %	2.71 %	1.77 %
3	esophagus	At most 100 cGy (RBE) dose at 10.00 % volume	86 %	131 cGy (RBE)	145 cGy (RBE)	43 cGy (RBE)
3	Right Lung	At most 15.00 % volume at 2500 cGy (RBE) dose	100 %	11.89 %	14.27 %	13.03 %
4	CTV	At least 98.00 % volume at 4400 cGy (RBE) dose	100 %	99.84 %	99.38 %	100.00 %

Figure 15 - DVH cluster and clinical goal evaluation in the RayStation® Robust Evaluation module for a robustly optimized IMPT plan for a lung cancer patient. The coloured curves represent individual error scenarios for the target (right cluster) and OAR indicated by the orange arrow (left cluster), with the nominal plan shown in white dash line. The table displays the clinical objectives alongside the percentage of scenarios meeting each criterion, worst-case scenario results, and nominal plan outcome. Adapted from (RaySearch Laboratories, 2019b).

At IPOPFG, treatment planning is performed in Varian Eclipse™ TPS, which currently lacks a dedicated photon robustness module. This limitation provides a strong rationale for developing a systematic robustness evaluation methodology within the present work. However, it allows evaluation of the calculated treatment plan under different scenarios involving density variations and isocenter shifts, generating DVH bands, as illustrated in Figure 16. Given that setup uncertainties can cause delivered dose distributions to deviate significantly from the initial plan, a systematic evaluation methodology is indispensable.

a)



b)

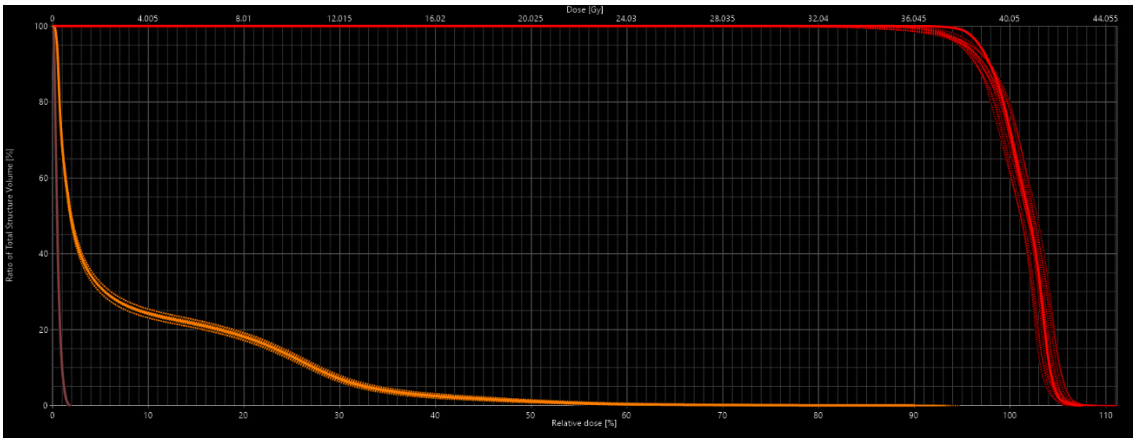


Figure 16 – Eclipse™ tool to simulate plans under density variations and isocenter shifts (a) with generated DVH bands (b).

There is no standardized framework to evaluate RT plan robustness. Nonetheless, robustness evaluation methods are increasingly relevant for modern photon techniques, particularly in hypofractionated and highly conformal treatments, due to their high complexity. In a highly exhaustive literature review analysing 225 published works on the topic, Kim et al. highlighted setup uncertainties (simulated by introducing translational isocenter shifts), range uncertainties (for charged particles) and anatomical changes (including motion, inter-fractional changes in tumour size, and delineation variability) as the most frequently evaluated uncertainties (Kim et al., 2025).

2.3 Robustness in breast radiotherapy

Robustness evaluation is particularly relevant in breast IMRT, as it is a highly modulated technique, associated with a higher number of MUs compared to conventional treatments, and increasingly used in hypofractionated regimens. These features, together with concerns regarding cardiac exposure in left-sided treatments, make robust maintenance of target coverage and OAR sparing especially important. Although techniques such as DIBH can substantially reduce cardiac dose and limit target motion in selected patients, respiratory-induced anatomical variations remain a relevant consideration, particularly in treatments delivered under free-breathing conditions. Treatment plans must therefore be sufficiently robust to maintain adequate PTV coverage and OAR sparing in the presence of these motion-related uncertainties.

This requirement for robustness must, however, be balanced against the risks of excessive fluence modulation. In breast radiotherapy, the CTV and PTV lie near the external body surface, where dose delivery is inherently more sensitive to perturbations. The skin-sparing effect of megavoltage photon beams reduces dose within the build-up region, while lateral electron disequilibrium is further accentuated in narrow beam segments, both factors increasing the susceptibility of highly modulated distributions to setup uncertainties and interfractional anatomical variations.

A further source of uncertainty is the interplay effect, arising from the temporal interaction between dynamic MLC motion and respiratory motion. Although generally less pronounced in conventionally fractionated breast IMRT than in stereotactic treatments, the interplay effect may nonetheless contribute to clinically relevant dose variations, particularly in hypofractionated regimens or highly modulated delivery sequences, and should therefore be explicitly considered when evaluating plan robustness.

To better understand how to assess the robustness of breast treatment plans, it is essential to identify the key uncertainties that affect treatment delivery and evaluate their dosimetric impact. Patient setup errors, including systematic and random translational shifts (left-right, anterior-posterior, superior-inferior) and rotational shifts (pitch, roll, yaw), can significantly affect structures near high-dose gradients. For instance, a worst-case combination of random uncertainties with a systematic shift of 3 mm or a 3° rotation can alter target and OAR dose volume endpoints by up to 9 Gy (Loebner et al., 2025). Another dominant uncertainty in breast treatment consists of organ motion and anatomical changes, caused by intra-fraction chest wall motion from breathing and inter-fraction variations. If not properly managed, chest wall motion alone can severely

degrade target coverage. During severe exhalation, target age (PTV $V_{90\%}$) in standard planning can drop to as low as 87.9%, whereas robust optimization techniques can successfully maintain this coverage at 97.6% (Miyasaka et al., 2024). Furthermore, mechanical uncertainties compromise treatment delivery particularly in highly modulated plans. Systematic MLC position errors of just 0.5 mm can cause dose variations of up to 3 Gy, while larger miscalibrations of 1 mm can lead to critical dose differences of up to 10 Gy (Loebner et al., 2025). Finally, variations in patient density, such as Hounsfield Unit (HU) changes, can significantly alter dose calculations. Yin et al. specifically included CT density uncertainty in their robustness analysis, recognizing it as a relevant variable. When evaluating a 2% CT density uncertainty combined with standard translational shifts, VMAT plans demonstrated a dose volume histogram band width (DVHBW) variation of 2.09%, compared to 2.98% for IMRT plans, highlighting the importance of explicitly including density variables in robustness analyses (Yin et al., 2024).

2.4 Metrics for robustness characterization

Although numerous studies have explored how to characterize plan robustness, it remains an open question and further investigation is required, as a universal methodology has yet to be established. While various robustness tools are already available, their clinical implementation varies significantly between oncology centers, being standard practice in some institutions while remaining absent in others.

Robustness evaluation can be divided into two broad categories: *a priori* approach based on pre-treatment simulations, and *a posteriori* approach based on delivered dose assessment using images or log files acquired during treatment, quantified by DVH-based, voxel-based or radiobiological metrics. The comparison of these three robustness evaluation methods is shown in the following table 5.

Table 5 - Comparison of robustness evaluation methods, adapted from (Kim et al., 2025).

Metric Type	Primary Advantage	Main Limitation
DVH-based	Ease of interpretation; clinician familiarity.	Lacks spatial information; heterogeneity in reporting.
Voxel-based	Provides 3D spatial insight into "hot" or "cold" spots.	Historically limited to worst-case scenarios in commercial tools
Radiobiological	Translates dose into anticipated clinical benefit.	Mathematical models may oversimplify biological response.

The following Table 6 provides a more detailed overview of the key metrics employed in robustness evaluation. Compiled from the work of Kim et al. this table expands on the three method categories outlined above by describing their specific subtypes and practical implementation.

Table 6 - Summary table of metrics used to evaluate plan robustness, adapted from (Kim et al., 2025).

Type of methods used	Subtypes and description	Details
DVH-based metrics	• Visualisation of dose-volume histograms (DVHs) Nominal DVH is shown alongside DVHs from simulated uncertainty scenarios or delivered dose distributions. • Variation of specific dosimetric parameters The variability of specific dosimetric parameters across uncertainty scenarios or between planned and delivered doses is reported using the mean or median, along with the range (minimum and maximum values) or standard deviation.	• Displayed as overlapping DVHs • Displayed as nominal DVH with surrounding bands: DVH bands corresponding to worst-case scenarios DVH bands with standard deviation DVH bands with 90% confidence interval <u>Target</u> • Maximum dose within the target Clinical target volume (CTV) D_{max}, CTV $D_{2\%}$, CTV $D_{5\%}$, $CTV_{V107\%}$ • Coverage of the target: CTV $V_{100\%}$, $V_{95\%}$, CTV $D_{99\%}$, $D_{98\%}$, $D_{50\%}$, D_{mean}, Planning target volume (PTV) $D_{95\%}$ • Dose homogeneity index • CTV conformity index <u>OAR</u> Site specific relevant dose parameter • eg Lung — mean lung dose, V_{20Gy}, V_{10Gy} • Pelvis–rectum and bladder V_{30Gy}, V_{40Gy}, V_{50Gy}, V_{60Gy} V_{70Gy} • Spinal cord — D_{max}
Voxel-based metrics	• Robustness threshold The proportion of simulated uncertainty scenarios meeting a predefined robust ness threshold. • Max-min 'error-bar' dose distribution Voxel-level dose ranges across all simulated uncertainty scenarios are	A proportion of scenarios in which CTV $D_{98\%} \geq 95\%$ or CTV $D_{95\%} \geq 95\%$ • Max-min dose distribution displayed like a dose map

Type of methods used	Subtypes and description	Details
	represented as an 'error-bar' dose distribution. • Gamma index Assesses spatial and dosimetric agreement between dose distributions through pairwise comparison. • Voxel-wise dose reconstruction Reconstructs dose distributions from nominal and uncertainty scenarios into voxel-wise (VW) minimum, mean and maximum dose distributions.	• Standard deviation volume histogram (SDVH) The standard deviation of doses at each voxel (x-axis) is plotted against the volume (y-axis). • Root mean square deviation (RMSD) volume histograms (RVH) RMSD is plotted (x-axis) against the volume (y-axis). • The area under the RVH • Gamma index analysis passing rates reported for: 2%/2 mm criteria 3%/3 mm criteria • Comparison of specific dosimetric parameters VW maximum dose distribution CTV $D_{98\%}$ VW minimum dose distribution CTV $V_{95\%}$ • Robustness threshold Plans were considered robust if CTV $V_{95\%}$ ≥ 98% in the VW minimum dose distribution
Radiobiological metrics	• Tumour control probability (TCP) for the target and normal tissue complication probability (NTCP) for the OARs. Changes in TCP or NTCP due to uncertainties or delivered vs. planned dose differences.	• Change in TCP and NTCP

According to Kim et al., most studies, exactly 179 out of 225, corresponding to approximately 80%, rely on DVH-based metrics to assess plan robustness, including DVH visualisation, variation in specific dosimetric parameters, and robustness thresholds. Among the most frequently reported parameters for evaluating CTV coverage were variations in CTV $D_{95\%}$, $D_{98\%}$ and $V_{95\%}$, as well as the proportion of scenarios in which CTV $D_{98\%}$>95%. However, assessing not only variations in DVHs between nominal and uncertainty scenarios, but also spatial information provided by maximum-minimum dose distribution or VW dose reconstruction is the recommended robust evaluation framework (Kim et al., 2025).

2.5 Plan quality and complexity in IMRT

Robustness is intrinsically related to both plan quality and complexity concepts.

Plan quality is defined by the degree to which a treatment plan satisfies the prescribed dose-volume constraints, ensuring adequate target coverage while minimizing dose to OARs (Podgorsak, E. B., 2005), and is typically evaluated through DVHs and conformity

indices. However, a high-quality plan is not necessarily a robust one. The dose distribution calculated by the TPS does not exactly correspond to the dose delivered to the patient, owing to inherent uncertainties in dose calculation and treatment delivery, including variations in patient setup and anatomy. Consequently, plan quality must also be evaluated in terms of robustness and complexity (Hernandez et al., 2020). A plan that is excessively sensitive to such perturbations may result in inadequate target coverage or the creation of unintended high-dose regions in healthy structures. Furthermore, the increased dose sculpting characteristic of modulated techniques often entails greater modulation of machine parameters, which may amplify uncertainties in dose calculation and delivery, with potential negative impact on clinical outcomes (Kaplan et al., 2022). Therefore, ensuring both high quality and high robustness is the primary objective of a modern medical physics approach to IMRT planning.

Plan complexity refers to the degree of modulation within a plan, such as the variability in aperture shape, leaf and gantry speed. Additionally, highly complex plans can be more sensitive to delivery errors, besides requiring longer beam-on times, which can increase the risk of intra-fraction movements in some treatment sites. As stated by Hernandez et al., a high degree of plan complexity may therefore compromise the overall accuracy and quality of the treatment (Hernandez et al., 2020). Consequently, high degrees of plan complexity should be avoided to maximise accuracy and robustness of radiation treatments.

Several indices have been proposed to quantify treatment plan complexity, each capturing distinct aspects of MLC modulation and aperture characteristics. Among these, the edge metric (EM) and the modulation factor (MF) stand out as particularly informative, as they offer a comprehensive characterization of plan complexity by jointly capturing geometric aperture irregularity and fluence modulation. The EM quantifies aperture irregularity through the ratio of MLC leaf-side edge length to aperture area. Therefore, larger differences between adjacent leaf positions yield higher EM values, reflecting more irregular aperture shapes (Hernandez et al., 2018). The MF, in turn, characterizes the degree of intensity modulation as the relative increase in MUs resulting from beam aperture modulation (Hernandez et al., 2025).

Additional complexity metrics include the modulation complexity score (MCS), which integrates variability in both aperture shape and area on a scale from 0 to 1, where lower values indicate greater complexity, the plan irregularity (PI) and plan modulation (PM) metrics, which assess aperture shape deviations and segment intensity modulation

respectively, and the leaf travel (LT), which indicates the average distance travelled by the MLC leaves (Hernandez et al., 2018).

3. Materials and methods

3.1 Study design

3.1.1 Patient selection

A database of all 107 breast cancer cases treated with IMRT at IPOPFG during 2024 was compiled, as the evaluation of breast IMRT, specifically regarding its robustness and complexity is a subject of significant clinical interest and the scope of this study. Within the initial cohort, inclusion and exclusion criteria were established to define the final study sample. All treatments were planned for a 2300IX linear accelerator (Varian®, Palo Alto, USA) equipped with a Millennium 120 MLC, using the anisotropic analytical algorithm (AAA) v.16.1.2 for dose calculation.

The dataset was categorized by laterality, and each patient was assigned an anonymised alphanumeric code, where "A" and "B" represented right and left breast cases, respectively. Treatment plans were individually reviewed within the Eclipse™ TPS to verify the technique used and the relevant dosimetric parameters.

According to the institution's clinical practice, which follows the protocol referenced in the literature, all left-breast patients are indicated for DIBH (Deep Inspiration Breath-Hold), while right-breast patients require specific medical indication and must be able to sustain a breath-hold of 20 seconds (Shaitelman, 2024). However, due to patient inability to sustain a stable breath-hold, logistical constraints and equipment scheduling, not all left-breast patients were able to undergo DIBH. In our dataset, only 8 of the 23 left-breast patients underwent this treatment modality, while none of the right-breast patients did so.

3.1.2 Inclusion and exclusion criteria

Palliative plans were excluded due to their distinct clinical intent, making their dosimetric parameters not directly comparable to those of curative breast irradiation. Conventionally fractionated plans were also excluded, as they follow a different radiobiological framework. Furthermore, plans involving additional nodal target volumes were removed because the additional dosimetric complexity lay beyond the scope of this analysis. Sequential boost plans were similarly excluded, as the boost is delivered in a separate phase, unlike the simultaneous integration used in SIB, rendering them technically

distinct. Additionally, SIB plans were omitted to ensure sample homogeneity and to account for the limited number of cases available within the study period. Overall, the excluded plans differed in clinical intent, fractionation strategy, target volume, or delivery technique, and their inclusion would have compromised the methodological consistency of the study.

Therefore, the analysis was narrowed to a single treatment modality, namely PBI, which involves the post-operative irradiation of a limited breast volume, targeting only the tissue surrounding the tumour bed. From a radiobiological perspective, this restricted target volume justifies the use of hypofractionated regimens, which allow for a higher dose per fraction to be delivered safely within a shortened treatment schedule.

Consequently, the sample size was reduced to 41 breast IMRT cases, including 18 right- and 23 left-sided plans.

3.2 Robustness analysis

3.2.1 Uncertainty scenarios

To assess plan robustness against geometric and mechanical uncertainties, each selected PBI case was subjected to seven perturbation scenarios.

Six setup error scenarios were created for each patient by applying ±3 mm isocenter shifts along the three translational axes: X (lateral/left-right), Y (vertical/anterior-posterior), and Z (longitudinal/cranial-caudal). For the lateral axis, a positive shift (X+3) corresponded to a 3 mm leftward displacement of the isocenter, while a negative shift (X−3) corresponded to a 3 mm rightward displacement. Similarly, for the vertical axis, a positive shift (Y+3) represented an anterior displacement, and a negative shift (Y−3) represented a posterior displacement. Finally, for the longitudinal axis, a positive shift (Z+3) indicated a cranial displacement, whereas a negative shift (Z−3) indicated a caudal displacement. The 3 mm setup deviation was chosen according to the institutional breast protocol, which states that a >3 mm deviation is the action level for special cases requiring strict target coverage.

Furthermore, given that IMRT plans delivered on the linac can be highly modulated and therefore particularly sensitive to the DLG, as discussed previously, evaluating this mechanical uncertainty is essential to ensure dosimetric accuracy. For this purpose, a virtual machine model incorporating a 1 mm variation in the MLC DLG,

designated “TRI1_Test”, was configured in Eclipse™. This 1 mm threshold was chosen because it represents a realistic worst-case scenario, accounting for normal mechanical wear of the MLC leaves over time. All generated plans were then recalculated using this model with fixed monitor units (MUs). This allowed the isolated dosimetric impact of the mechanical uncertainty to be assessed independently of plan complexity. Altogether, these seven perturbation scenarios, combined with the nominal plan, resulted in a total of eight plans per patient.

Right and left-sided cases were evaluated separately for mean heart dose analysis. Scatter plots displaying the nominal plans and perturbed scenarios were used to evaluate plan robustness. Firstly, a dedicated robustness course was created on Eclipse™, to which all the original and approved plans were copied.

Initially, scatter plots were generated for each clinically relevant structure defined in the institution's clinical goals, namely the PTV, OARs, and the remaining HT. These plots were constructed using the difference between the values obtained in each of the seven perturbation scenarios and the corresponding nominal plan value. It should be noted that for the heart, the relative error was calculated with respect to the nominal scenario value, given that mean heart dose is recorded in absolute terms (Gy) rather than as a percentage. While these scatter plots allowed general trends to be identified and the most displaced cases to be flagged, they proved insufficient for a thorough assessment of what was occurring within each uncertainty scenario and, consequently, for implementing a comprehensive robustness analysis. Therefore, an alternative robustness evaluation approach was pursued, using both DVH- and voxel-based metrics, with the aim of establishing a robustness quantitative index.

3.2.2 Metrics

3.2.2.1 DVH-based metrics

The institutional clinical goals for each structure were translated into operational robustness metrics by defining a primary threshold and an acceptable tolerance range.

The Healthy Tissue (HT) structure was removed as a standalone metric, and its objective was reassigned as the acceptable dose variation criterion for the PTV. Accordingly, PTV maximum dose values below 110% of the prescribed dose were classified as robust, while values between 110% and 112% were considered within the acceptable tolerance

range. Only plans in which PTV maximum dose exceeded the 112% threshold were classified as non-robust.

Based on institutional guidelines, PTV_$V_{95\%}$ was assessed to ensure adequate coverage. Values below 95% breached the first robustness barrier, falling into an acceptable tolerance range, while coverage dropping below 90% was classified as non-robust. Similarly, mean heart dose thresholds were established according to the department's clinical goals to mitigate long-term cardiac toxicity. For left-sided cases, mean heart doses exceeding 2 Gy were considered non-robust, whereas values between 1.5 Gy and 2 Gy fell within the acceptable tolerance. For right-sided cases, mean heart doses exceeding 1 Gy were classified as non-robust, with values between 0.7 Gy and 1 Gy deemed acceptable. Regarding the ipsilateral lung, $V_{30\%}$ values exceeding 15% of the prescribed dose were considered non-robust, whereas values between 10 and 15% fell within the acceptable tolerance. The specific quantitative limits defining the first barrier and the ultimate non-robustness threshold for all evaluated structures are summarized in table 7.

These DVH-based metrics were then employed to assess plan robustness by analysing variations in clinical goals between the nominal and perturbed dose distributions. Conformity, gradient and homogeneity indices, although reported for baseline plan approval, were excluded from this phase. Robustness evaluation prioritises absolute clinical safety thresholds over the spatial characteristics of the nominal plan, as geometric uncertainties inherently degrade conformity but do not necessarily compromise patient safety if target and OAR constraints are maintained. Consequently, the analysis focused exclusively on PTV $V_{95\%}$, Heart D_{mean} and Ipsilateral lung $V_{30\%}$ as they represent the most critical parameters.

Table 7 - Quantitative thresholds for the first robustness barrier and the ultimate non-robustness threshold applied in the DVH-based evaluation of the PTV, heart, and ipsilateral lung.

Structure parameter	First barrier	Threshold
PTV $V_{95\%}$	>95%	>90%
Heart D_{mean} (left-sided cases)	<1.5Gy	<2Gy
Heart D_{mean} (right-sided cases)	<0.7Gy	<1Gy
Ipsilateral lung $V_{30\%}$	<10%	<15%

Other OARs such as the contralateral breast and contralateral lung were not evaluated for robustness because their distance from the steep dose gradient makes them far less susceptible to clinically significant variations during standard setup uncertainties. The final selection prioritised the dosimetrically most relevant endpoints for target coverage and nearby OAR protection. The PTV reflects target coverage adequacy, the heart acts as the most critical dose-limiting OAR due to its proximity to the target, and the ipsilateral lung is highly sensitive to volumetric dose variations arising from geometric shifts.

Plans were considered robust when clinical goals remained within the accepted variation limits across all perturbation scenarios. Failure to meet the minimum target coverage threshold or exceeding the acceptable tolerance for either mean heart dose or ipsilateral lung $V_{30\%}$ in any scenario, classified the plan as non-robust.

3.2.2.2 Voxel-based metrics

To complement the DVH-based framework, an Eclipse™ Scripting API (Figure 17) was developed to process voxel-level spatial information derived from maximum–minimum dose distributions, following the general approach described by Kim et al. The full code is provided in Attachment 2 – Eclipse™ scripting API code.

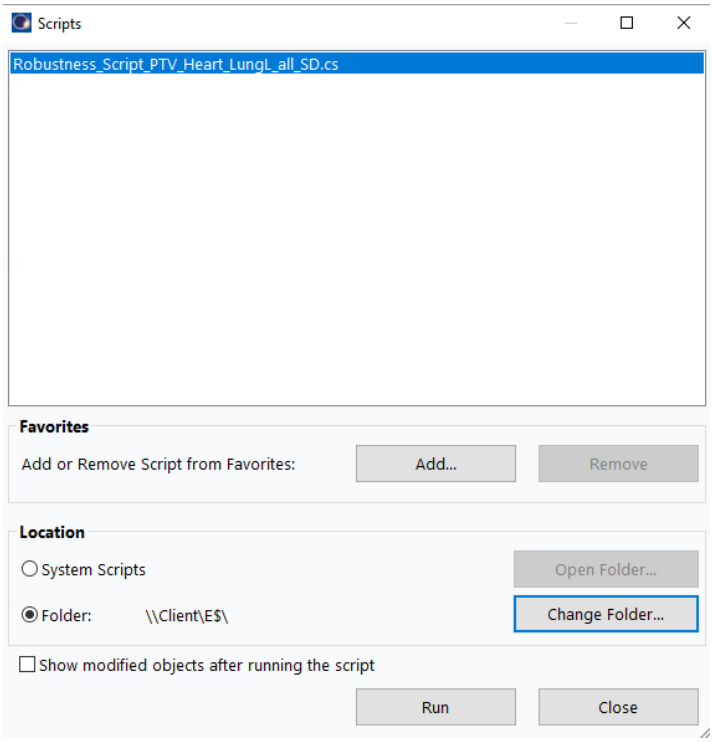


Figure 17 – Eclipse™ scripting API ready for execution.

Figure 18 outlines the methodology implemented in the script, which loops through the 8 plans, including the nominal and the 7 perturbation scenarios, to extract the exact dose at the voxel coordinates.

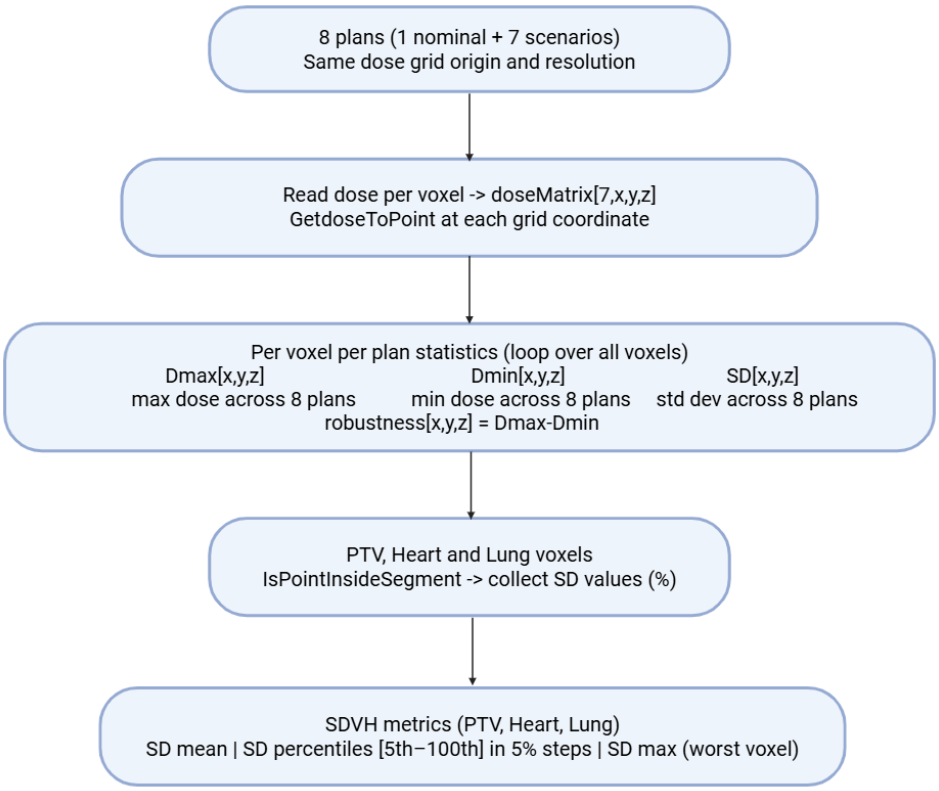


Figure 18 - Methodology implemented by the script for the voxel-based metrics analysis, created in https://BioRender.com.

This approach returned the voxel-wise standard deviation (SD) of dose values, expressed as a percentage of the prescribed dose, across the eight scenarios for the PTV and OARs. In this study, the heart and the ipsilateral lung were selected as the OARs of interest, with right and left-sided cases analysed separately. The script processed voxel-wise SD values into an ascending cumulative distribution. To reconstruct the Standard Deviation Volume Histogram (SDVH), SD thresholds were sampled at 5% volume increments chosen to provide sufficient data resolution, ranging from the 5th to the 100th percentile, the latter corresponding to the maximum SD. The total number of voxels per structure was also returned in the output window, enabling rapid critical review and preliminary verification of the extracted values. It is worth noting that in Eclipse™, the CT image voxel dimensions are defined by the image properties, whereas dose voxel dimensions are determined by the calculation grid settings. Consequently, when calculating dose, Eclipse™ readjusts the voxel dimensions to align with the calculation grid.

A second output of the algorithm was a .csv file compiling the SD values metrics described above, alongside a 3D voxel-based robustness map exported as a .mhd file. For each patient, a dedicated folder, identified by the patient ID, was generated, containing both the extracted statistical parameters and the robustness analysis files for each structure under evaluation. The robustness of each voxel was defined as the difference between the maximum and minimum dose calculated across all 8 scenarios. The resulting robustness maps were subsequently imported into 3D Slicer alongside the CT images and Structure Sets exported from Eclipse™, enabling a comprehensive visual and spatial analysis of dose uncertainty distribution throughout the volume of interest. The algorithm's output was manually validated by temporarily incorporating a print statement to confirm that the SD extracted values were correctly expressed as percentages rather than absolute values, with the corresponding validation code provided in Attachment 3 - Validation code for SD percentage output.

Voxel-based metrics were then employed to assess plan robustness by establishing robustness thresholds, based on the clinical goals for the heart, PTV and ipsilateral lung, and by analysing the corresponding SDVH for the best and worst cases regarding the three evaluated structures.

For the PTV, $V_{95\%} \geq 95{,}0\%$ represents the clinical goal. Therefore, the first barrier was defined as $SD_{95} < 5\%\ D_p$, meaning that for patients where the 95th percentile of the voxel-wise standard deviation distribution across all scenarios did not exceed 5% of the prescribed dose, all PTV voxels were considered robust. It is worth noting that SD_{max} was not used, since voxels at the PTV are in regions of steep dose gradient at the boundary with healthy tissue. As a result, they are highly sensitive to setup deviations, as the $\pm 3\ mm$ introduced isocenter shifts, and would systematically inflate the maximum SD regardless of the plan quality. The second barrier required $SD_{mean} < 2\%\ D_p$, ensuring that PTV coverage remained robust under the simulated uncertainties. Voxels satisfying the second barrier but not the first were classified as predominantly robust, while the ones failing both barriers were considered non-robust.

Based on the clinical goals stating that the mean heart dose should not exceed 1.7Gy (approximately 4.25% of 40Gy prescribed dose), the first robustness threshold was defined as $SD_{95} < 4\%\ D_p$. All voxels meeting this condition were considered robust. A second threshold was defined by $SD_{mean} < 2\%\ D_p$, ensuring that the mean heart dose remained stable across all simulated scenarios, thereby limiting uncertainty in the parameter most critical to long-term cardiac toxicity.

Finally, consistent with parameters established for the heart, the first barrier for the ipsilateral lung, another OAR, was set at $\mathrm{SD}_{95} < 4\%\mathrm{D}_{\mathrm{p}}$ and the second at $\mathrm{SD}_{\mathrm{mean}} < 2\%\mathrm{D}_{\mathrm{p}}$. The parallel nature of the lung makes mean dose variability the most clinically appropriate robustness criterion. Therefore, keeping it below 2% of the prescribed dose directly helps protect against a significantly higher risk of radiation-induced side effects. As with the PTV, the maximum SD was excluded from this criterion, as it was not considered a representative metric, since voxels at the lung boundary are similarly exposed to steep dose gradients, and the maximum SD would therefore reflect local peripheral voxel variability rather than the overall robustness of the voxels across the structure. Based on this consideration, the specific quantitative thresholds established for the voxel-wise robustness analysis of each structure are summarized in table 8.

Table 8 - Quantitative thresholds for the first and second robustness barriers applied in the voxel-wise evaluation of the PTV, heart, and ipsilateral lung.

Structure	First barrier	Second barrier
PTV	$\mathrm{SD}_{95} < 5\%\mathrm{D}_{\mathrm{p}}$	$\mathrm{SD}_{\mathrm{mean}} < 2\%\mathrm{D}_{\mathrm{p}}$
Heart	$\mathrm{SD}_{95} < 4\%\mathrm{D}_{\mathrm{p}}$	$\mathrm{SD}_{\mathrm{mean}} < 2\%\mathrm{D}_{\mathrm{p}}$
Ipsilateral lung	$\mathrm{SD}_{95} < 4\%\mathrm{D}_{\mathrm{p}}$	$\mathrm{SD}_{\mathrm{mean}} < 2\%\mathrm{D}_{\mathrm{p}}$

To further visualise the voxel robustness of each structure across the cohort, a scatter plot was generated for all 41 PBI patients, simultaneously displaying the $SD_{95}(\%)$ for the PTV, heart, and ipsilateral lung. This allowed straightforward identification of the most and least robust cases for each structure. For both cases, the SDVH was plotted as a histogram in which voxel SD values were binned into 5% volume intervals, representing the percentage of voxels within the structure as extracted by the implemented code. This approach provided a clear visualisation of dose variability across the sampled perturbation scenarios.

Subsequently, the worst-case patients identified through the voxel-wise analysis were further examined in Eclipse™ and 3D Slicer. This step aimed to determine whether the observed lack of robustness was due to patient-specific geometry or anatomy, and to verify whether these findings corroborated the plan degradation previously identified through DVH analysis.

By integrating DVH-based robustness metrics with voxel-wise dose analysis and patient-specific geometric assessments, this methodology enabled a comprehensive evaluation of how robustly optimised plans respond dosimetrically to setup and mechanical uncertainties, accounting for both systematic trends and individual outliers within a clinically meaningful context. Additionally, a quantitative robustness index incorporating both DVH and voxel-wise metrics was proposed to provide a consolidated measure of plan resilience.

3.2.2.3 Robustness index (RI)

Following the establishment of both DVH- and voxel-based robustness metrics, a quantitative robustness index (RI) was developed. The primary objective of this index is to assign a single robustness score to each patient's plan, thereby enabling a direct evaluation of a potential correlation between plan robustness and complexity. To achieve this, the robustness index presented in equation 5 was proposed.

$$\mathrm{RI} = 1 - \sqrt{\mathrm{w_{PTV}}\left(\frac{\mathrm{SD_{95,PTV}}}{\mathrm{CG_{PTV}}}\right)^2 + \mathrm{w_{Heart}}\left(\frac{\mathrm{SD_{95,Heart}}}{\mathrm{CG_{Heart}}}\right)^2 + \mathrm{w_{I\_Lung}}\left(\frac{\mathrm{SD_{95,I\_Lung}}}{\mathrm{CG_{I\_Lung}}}\right)^2} \qquad \text{eq. 5}$$

Where $\mathrm{w_{PTV}}$, $\mathrm{w_{Heart}}$ and $\mathrm{w_{I_{Lung}}}$ represent the weighting factors assigned to the PTV, heart, and ipsilateral lung, respectively. These factors were determined based on established institutional priorities, and their specific values are summarized in table 9. Additionally, $\mathrm{SD_{95}}$ denotes the 95th percentile SD of the dose for each structure, whereas CG represents the strictest clinical goal threshold defined for the corresponding structure.

Table 9 - Clinical weighting factors w applied in the robustness index evaluation for left- and right-sided cases.

Structure	*w* for left-sided cases	*w* for right-sided cases
PTV	0.5	0.5
Heart	0.35	0.25
Ipsilateral lung	0.15	0.25

It should be noted that the sum of the weighting factors for all structures must equal 1. For left-sided cases, a higher weighting factor was assigned to the heart compared to the ipsilateral lung, reflecting its clinical priority as the primary dose-limiting organ at risk.

Finally, within the proposed robustness index scale, a value of 1 represents a perfectly robust plan, whereas a value of 0 denotes a highly unstable or non-robust plan.

3.3 Complexity analysis

This study further investigated the association between plan complexity and plan robustness. Although various methods exist to quantify complexity, there is no consensus on the optimal approach. Therefore, two specific metrics were adopted to evaluate different aspects of plan complexity, namely the modulation factor and the edge metric. The modulation factor, based on the metric proposed by Webb et al., was defined as the ratio of the total monitor units summed across all treatment fields to the fraction dose in cGy, therefore expressed in $\mathrm{MU/cGy}$. To calculate this, total MUs were manually extracted and divided by the fraction dose (Webb, 2003).

The edge metric, expressed in $\mathrm{mm^{-1}}$, was based on Younge et al. study and assesses plan deliverability through aperture complexity analysis, given by Equation 6 (Younge et al., 2016).

$$\mathrm{M} = \frac{1}{\mathrm{MU}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{MU_i} \times \frac{\mathrm{y_i}}{\mathrm{A_i}} \qquad \mathrm{eq.}\,6$$

Where N is the number of control points, MU the total number of MU in the plan, $\mathrm{MU_i}$ the number of MU delivered through aperture i, $\mathrm{A_i}$ the open area of aperture i and $\mathrm{y_i}$ the aperture perimeter excluding the MLC leaf ends, represented by the green dash lines in figure 19.

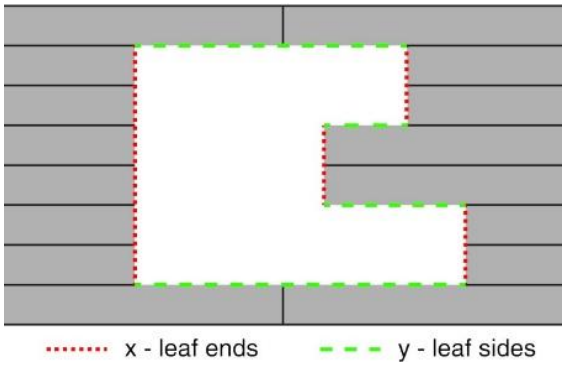


Figure 19 - Schematic representation of an MLC aperture used for the complexity metric calculation. The aperture perimeter is divided into contributions from the leaf ends (vertical red dotted lines) and the leaf sides (horizontal green dashed lines). Reproduced from (Younge et al., 2016). Copyright © 2012 American Association of Physicists in Medicine.

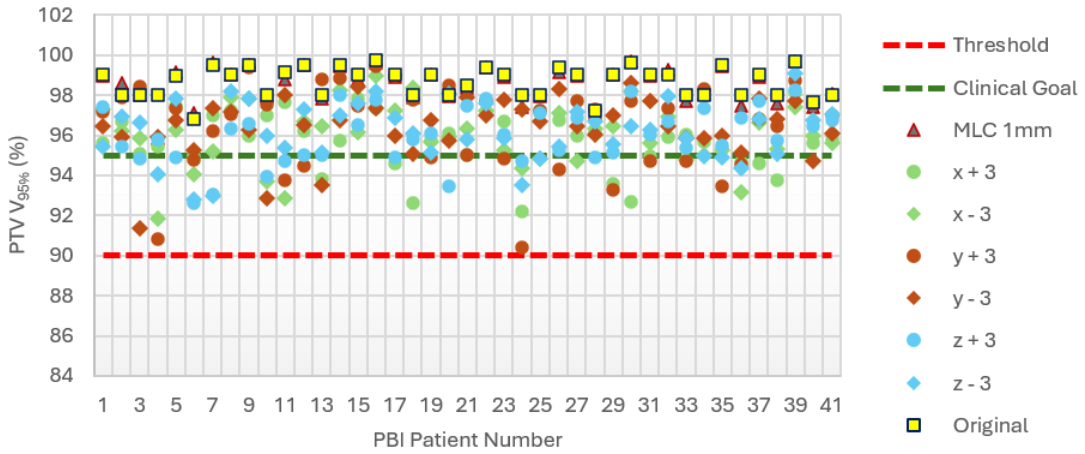
To automate the evaluation of the metric proposed by Younge et al., a custom script was developed and executed in Eclipse™ to determine the total complexity of each nominal plan, which, as previously noted, remained representative of all eight scenarios per patient given the fixed-MU dose recalculation. Subsequently, a statistical analysis was performed to investigate the correlation between plan complexity and the calculated robustness index. Prior to the correlation analysis, the normality of the data distribution was assessed with the Shapiro-Wilk test, confirming that it did not follow a normal distribution. Consequently, the non-parametric Spearman's rank correlation coefficient was used to investigate this relationship. Statistical significance was defined by a $\mathrm{p-value} < 0.05$. All statistical evaluations were performed using Python software.

4. Results and discussion

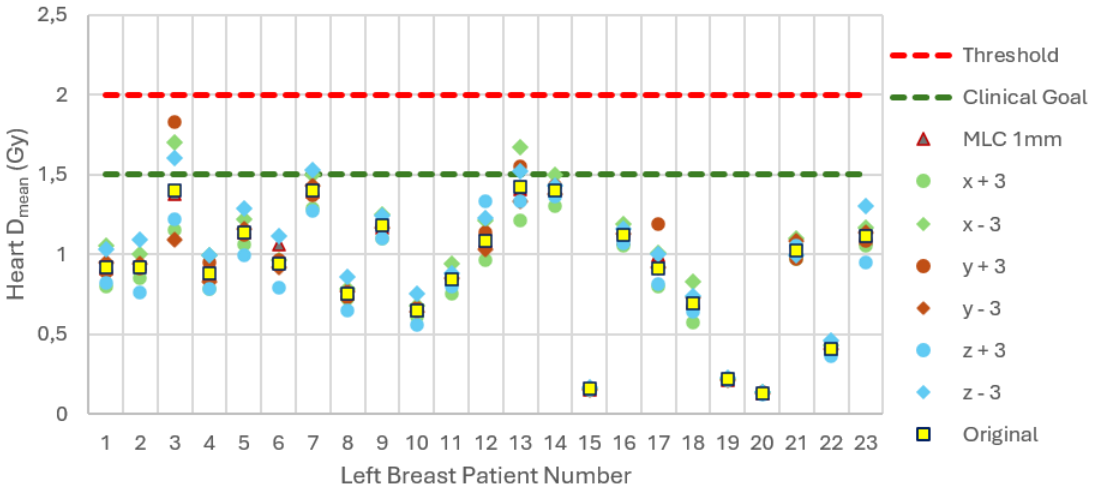
4.1 DVH-based evaluation

Plan robustness was evaluated using scatter plots (Figure 20 a-d), displaying the dosimetric variations between the nominal plans and the simulated perturbed scenarios. For this analysis, $\mathrm{PTV\ V_{95\%}}$, $\mathrm{Heart\ D_{mean}}$ and $\mathrm{Ipsilateral\ lung\ V_{30\%}}$ were prioritized as the primary clinical goals. $\mathrm{PTV\ V_{95\%}}$ represents target volume coverage and is a critical parameter in robust analysis, ensuring that setup and mechanical uncertainties do not compromise treatment efficacy. He et al. used this parameter to evaluate target coverage degradation in the presence of $\pm 3\ \mathrm{mm}$ setup errors, consistent with those implemented in the retrospective study under consideration (He et al., 2023). Furthermore, maintaining the robustness of mean heart dose is essential to minimize the risk of late radiation-induced cardiovascular complications. Darby et al. demonstrated that in breast patients undergoing radiotherapy, the rate of major coronary events increases by approximately 7.4% per Gy of mean heart dose, with no apparent threshold, highlighting $\mathrm{Heart\ D_{mean}}$ as a critical OAR metric (Darby et al., 2013) . Additionally, the ipsilateral lung was incorporated into the robustness analysis using the $V_{30\%}$ metric to minimize the risk of radiation-induced lung injury in breast cancer patients undergoing radiotherapy. Gueiderikh et al. demonstrated that $\mathrm{ipsilateral\ lung\ V_{30\%}} < 15\%\mathrm{Dp}$ is a key dosimetric constraint (Gueiderikh et al., 2023).

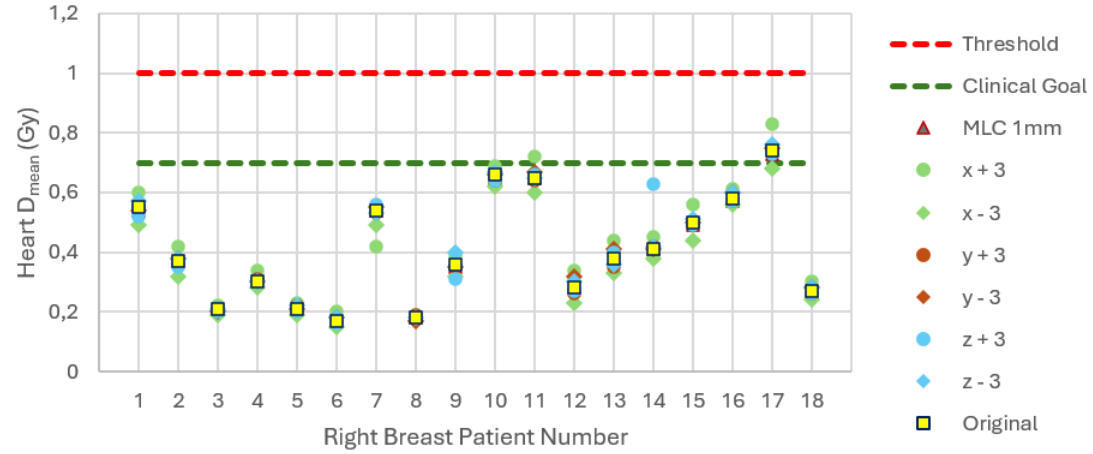
a)



b)


2,5
2
1,5
1
0,5
0
Heart D_{mean} (Gy)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Left Breast Patient Number
Threshold
Clinical Goal
MLC 1mm
x + 3
x - 3
y + 3
y - 3
z + 3
z - 3
Original

c)


1,2
1
0,8
0,6
0,4
0,2
0
Heart D_{mean} (Gy)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Right Breast Patient Number
Threshold
Clinical Goal
MLC 1mm
x + 3
x - 3
y + 3
y - 3
z + 3
z - 3
Original

d)

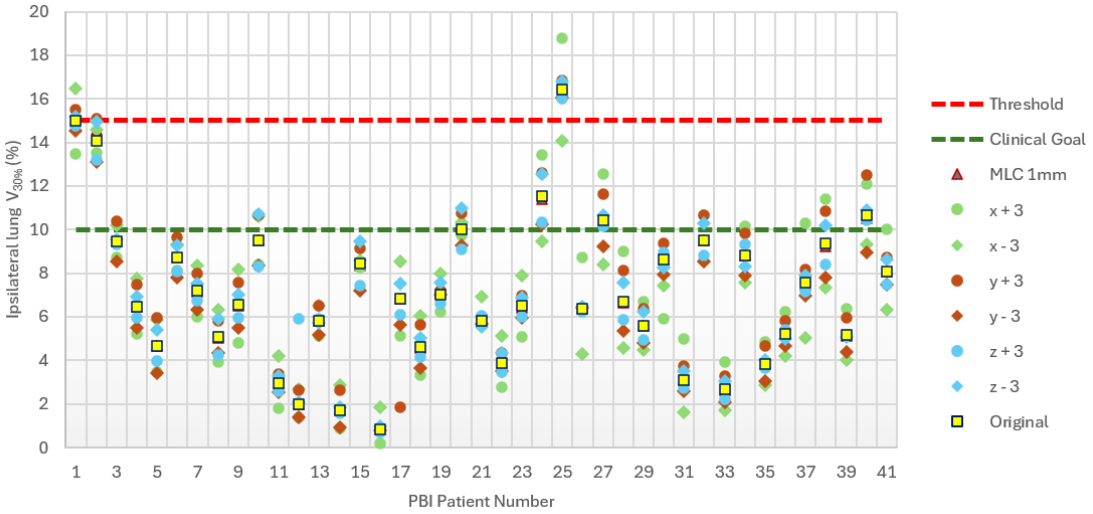


Figure 20 - Scatter plots of PTV $V_{95\%}$ (a), Heart Dmean (b-c) and Ipsilateral lung $V_{30\%}$ (d). Yellow squares represent the nominal plans, while the remaining markers correspond to simulated setup uncertainties (±3 mm shifts in the x, y, and z directions) and MLC variations (1 mm in the DLG). The green and red dashed lines represent the clinical goal and the robustness threshold, respectively. Points falling below this threshold in target coverage or exceeding its value in the mean heart dose or ipsilateral lung $V_{30\%}$ indicate lack of robustness.

Based on the results presented in Figure 20 a), all plans were considered robust with respect to PTV coverage, as none fell below the established robustness threshold (indicated by the dashed red line). However, three cases were identified as notably less robust (patients 3, 4, and 24). For patient 24 specifically, the Y+3 scenario resulted in target coverage extremely close to the threshold. This highlighted a critical situation that warranted a more detailed investigation of the dose distribution, namely through a voxel-wise analysis.

Regarding cardiac robustness, Figures 20 b) and c) revealed a clear laterality-dependent trend, as left-sided cases consistently exhibited lower robustness than right-sided ones, with their perturbed scenarios approaching tolerance limits more closely. Within the left-sided cohort, Figure 20 b), patient 3 emerged as the least robust case, with three uncertainty scenarios (Y+3, X-3, and Z-3) approaching the maximum acceptable dose limit. Although all plans remained within clinically acceptable tolerances, a careful evaluation was still required. Notably, even the nominal plan for this patient, represented by the yellow square, already approached the robustness boundary, further underscoring the clinical relevance of a careful evaluation of this case. On the other hand, patient 20 represented the most robust case across the cohort. The perturbed scenarios for this patient showed minimal deviation from the nominal plan, with all eight uncertainty

scenarios comfortably satisfying both the clinical goals and the established robustness threshold.

For right-sided cases, Figure 20 c), the X+3 plan (a leftward isocenter shift) proved to be the worst-case scenario in terms of Heart D_{mean} for almost all patients, as this shift displaces the treatment fields toward the chest midline and mediastinum, increasing the overlap with the cardiac structures, consistent with recent robustness analyses identifying mediastinum-directed isocenter shifts as the most detrimental for mean heart dose (Ding et al., 2022).

For the ipsilateral lung, presented in Figure 20 d), patients 1, 2 and 25 were considered non-robust, as perturbation scenarios exceeded the 15% threshold. Patient 25 stood out as the most critical case, with the X+3 scenario reaching approximately 19%, well above the threshold. Across the remaining patients, most scenarios fell below the 10% clinical goal, reflecting generally acceptable ipsilateral lung sparing under the studied perturbations.

Overall, as physically expected, increasing the MLC DLG relative to the original beam model resulted in a slight dose increment to the heart. However, the dosimetric impact of variations in the MLC DLG was minimal, consistent with recent studies reporting that MLC-related uncertainties produce comparatively small dose changes relative to setup-related uncertainties (Enomoto et al., 2024). Variations along the Y and Z directions produced an intermediate effect, ranking between the MLC and X-direction shifts, which is in line with the moderate OAR dose increases reported for anterior-posterior and cranio-caudal isocenter shifts in similar robustness evaluations (He et al., 2023).

It should be noted that the presented DVH-based analysis may not fully capture spatial dose coverage limitations at the PTV boundary, particularly near the skin surface, where steep dose gradients can obscure localized underdosage. Therefore, a voxel-based evaluation will further clarify these spatial coverage aspects under the studied perturbation scenarios.

4.2 Voxel-based evaluation

To complement the DVH-based analysis, a voxel-wise analysis was developed to assess spatial dose variability distribution. This approach provided a visual representation of the most susceptible regions to the simulated mechanical and geometric uncertainties, particularly relevant for the least robust cases.

As described in the chapter 3, a dedicated script was implemented and subsequently executed in the Eclipse™ TPS with the purpose of segmenting the voxels belonging to the PTV, heart, and ipsilateral lung. For each voxel coordinate within these structures, the standard deviation (SD) of the dose across the eight scenarios was recorded. These voxel-wise SD values, expressed as a percentage of the prescribed dose, were then used to identify and characterise the most and least robust cases.

The magnitude of SD values is influenced by both the individual anatomy of the PTV and the surrounding patient geometry. Since isocenter shifts were systematically applied, certain voxels will inevitably receive higher doses under specific scenarios, resulting in greater SD values for the affected structures. Therefore, elevated SD values are expected to arise predominantly at the periphery of each structure, as these regions are inherently more susceptible to steep dose gradients, an effect further amplified by the 3 mm isocenter shifts introduced for the robustness evaluation. Moreover, a PTV positioned in closer proximity to the sternum is expected to yield higher SD values for both the heart and the ipsilateral lung, relative to a more laterally situated PTV.

Neither the maximum nor the mean SD value alone is sufficient to fully characterise the dosimetric behaviour of a given structure. The maximum SD may be driven by a single highly sensitive voxel, a localised dosimetric hot spot that is not representative of the structure's overall response, whereas the mean SD may dilute the contribution of the most and least sensitive voxels, thereby masking clinically relevant dosimetric heterogeneity within the structure. For this reason, robustness thresholds were employed to assess plan robustness.

The scatter plot presented in Figure 21 illustrates the 95th percentile SD values for each of the three evaluated structures, enabling a rapid and intuitive identification of the most and least robust cases across the patient cohort.

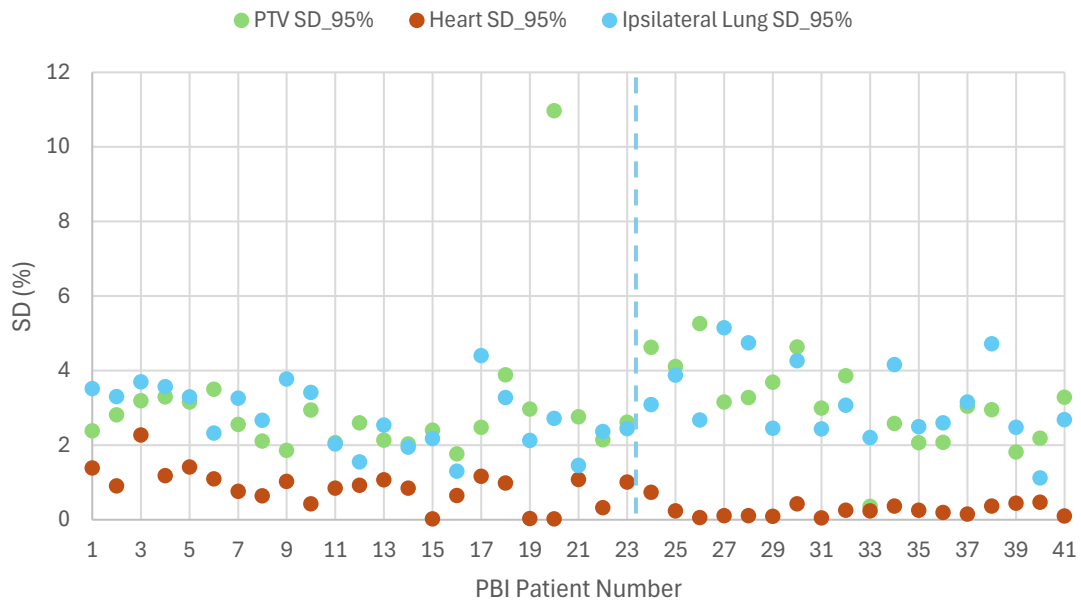


Figure 21 - Scatter plot of the 95th percentile dose standard deviation (SD_{95}%, expressed as a percentage of the prescribed dose) for the PTV, heart, and ipsilateral lung across all patients. Each data point represents an individual patient, colour-coded by structure - PTV (green), heart (brick red), and ipsilateral lung (blue). The vertical blue dashed line separates left-sided (left of the line) from right-sided (right of the line) cases.

As illustrated in Figure 21, a clear laterality-dependent pattern emerges when comparing cardiac SD_{95}% values across the two groups. Right-sided cases consistently present lower SD values for the heart, indicating greater cardiac robustness relative to left-sided cases, whose perturbed scenarios exhibit larger dosimetric variability in this structure. This trend is expected, as already demonstrated by the DVH analysis, given that the left breast lies in closer proximity to the heart, making it inherently more susceptible to dose variations arising from geometric uncertainties. This laterality-dependent behaviour will be further explored in the dedicated mean heart dose voxel-wise evaluation.

4.2.1 PTV evaluation

Regarding the PTV, patients 20 and 26 failed the voxel-wise robustness criteria, presenting SD_{95} values of 10.97% and 5.264%, respectively, thereby failing the first robustness barrier defined as $\mathrm{SD}_{95} < 5\%\mathrm{D}_{\mathrm{p}}$. Furthermore, their mean SD values were 2.861% and 2.353%, respectively, both exceeding the established robustness limit of $\mathrm{SD}_{\mathrm{mean}} < 2\%\mathrm{D}_{\mathrm{p}}$, classifying both cases as non-robust. A detailed inspection performed in Eclipse™ allowed the identification of the underlying factors contributing to this lack of robustness. In the case of patient 26, as evidenced in Figure 22, the 95% isodose line failed to consistently encompass the PTV across all perturbed scenarios, most critically

in the Y+3 scenario, illustrated in the right panel of Figure 22. As a result, certain regions of the PTV were subject to pronounced dose variability, manifesting as elevated SD values in those voxels.

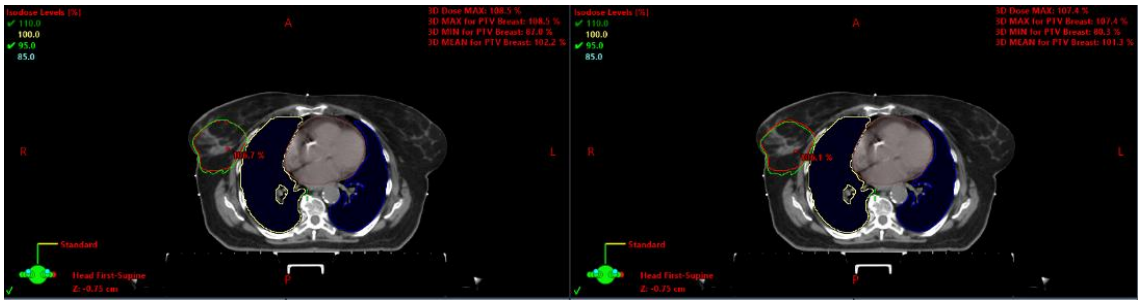
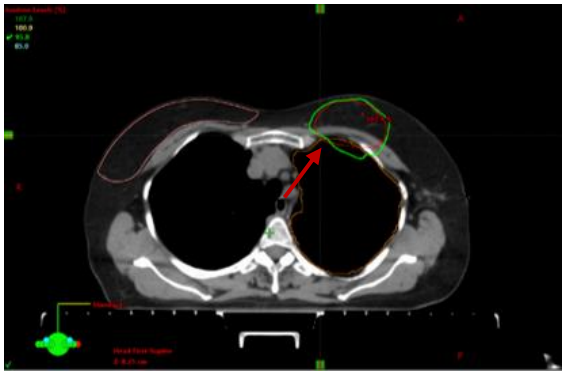


Figure 22 - Dose distribution maps for case 26. The nominal plan (left panel) and the worst-case scenario, Y+3 (right panel). In the latter, the PTV is not entirely encompassed by the 95% isodose line (shown in green), resulting in elevated SD values across the affected voxels and consequently classifying this case as non-robust.

Case 20 was non-robust with respect to PTV coverage, primarily due to the anatomical location of the tumour, with the PTV located in the upper inner quadrant of the breast. Given the clinical requirement to minimise lung dose, which was already at its tolerance limit, and since the field angulation could not be modified to avoid irradiation of the contralateral breast, the resulting treatment plan relied on a highly modulated fluence pattern, with a 3.716 MU/cGy modulation factor and an edge metric of 0.093 mm^{-1}, values based on the metrics proposed by Webb et al. and Younge et al., respectively. This extreme modulation rendered the PTV particularly sensitive to geometric uncertainties, as the 3 mm isocenter shifts. The Z+3 scenario proved to be by far the worst-case perturbation for this patient, as it directly affected the region of highest dosimetric vulnerability. As shown in the figures 23 a) and b), the 95% isodose line (shown in green) failed to consistently encompass the entirety of the PTV (delineated in red) across all perturbed scenarios, further corroborating the non-robust classification of this case.

a)



b)

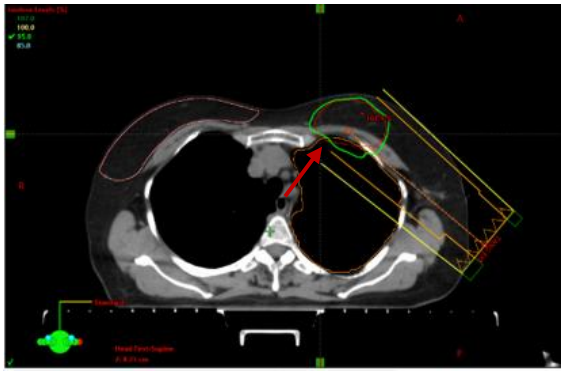
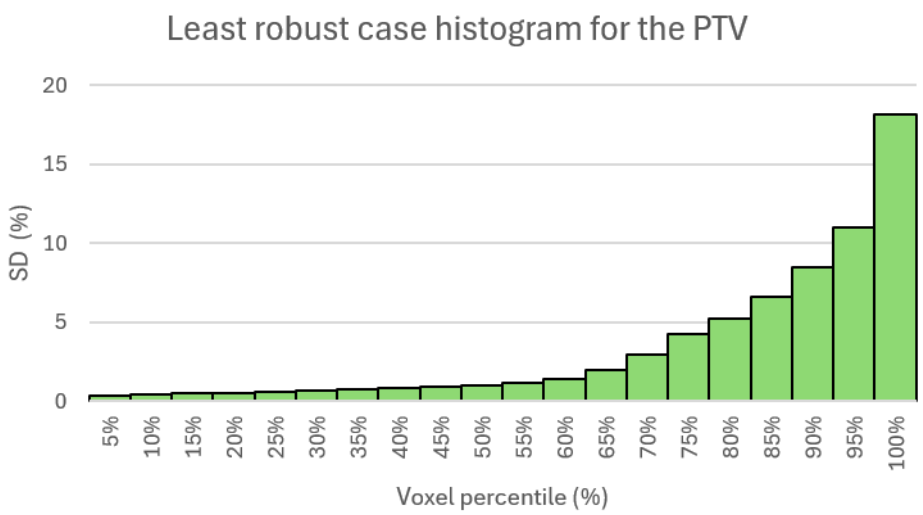


Figure 23 - Dose distribution maps for case 20 in Eclipse™, without (a) and with (b) the treatment field overlay. In both panels, the 95% isodose line (green) fails to fully encompass the PTV (red contour), with the red arrow indicating where target coverage is most deficient. This anatomical configuration explains the outlier behaviour of case 20 in the PTV scatter plot, as well as its classification as non-robust with respect to target volume coverage.

Following the robustness evaluation of the cohort, case 20 was clearly identified as the least robust with respect to PTV coverage, while case 33 emerged as the most robust. Figures 24 a) and b) present voxel-wise dosimetric variability histograms for these two cases. Although both share a similar distributional shape, with SD values remaining low across most voxels and rising sharply beyond the 90th percentile, the difference in magnitude is considerable. Case 20 reaches SD values of approximately 18% at the 100th percentile, roughly 40 times those observed in case 33 (~0.45%), underscoring the substantial disparity in robustness between the two extremes of the cohort. This is further reflected in the SD_{95} metric, which remains markedly lower for case 33, consistent with its superior PTV coverage robustness.

a)



b)

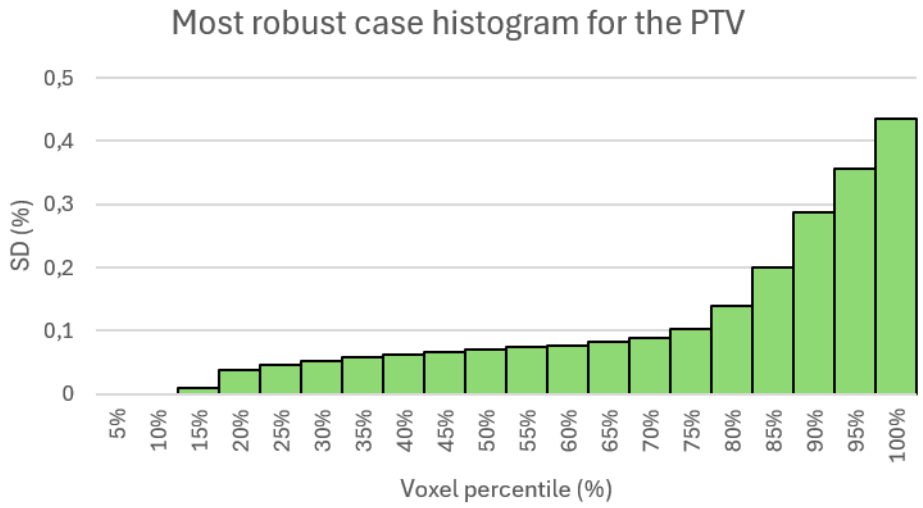


Figure 24 - Voxel-wise dose standard deviation histograms for the least robust (20, a) and most robust (33, b) PTV cases. The markedly lower SD values observed for case 33, in terms of SD_{95}%, highlight the substantial difference in dosimetric robustness between the two cases.

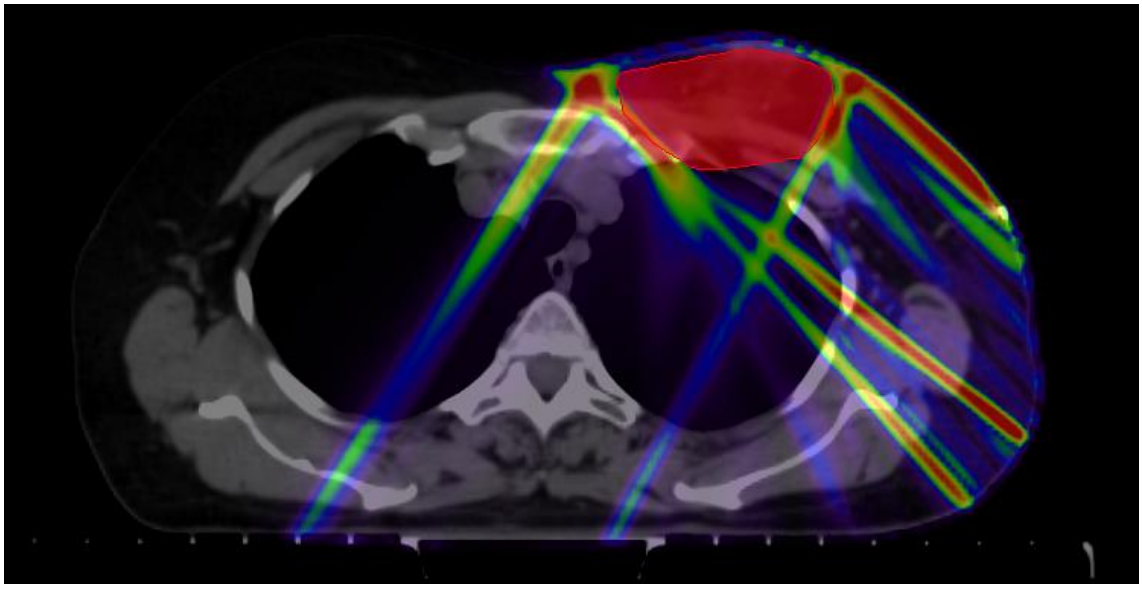
Of note, patient 25 was the second least robust case for PTV coverage, having failed the SD_mean criterion with a value of 3.311% of the prescribed dose, exceeding the tolerance threshold and confirming increased sensitivity to the uncertainties considered.

In contrast, although patients 3 and 4 were close to the robustness threshold identified by the DVH-based analysis presented in Figure 20 a), they showed no concerning values at the voxel level. It was therefore concluded that these cases remained robust against the mechanical and geometric uncertainties simulated in this study, ensuring adequate

target coverage without compromising the overall quality of the treatment plans. Similarly, patient 24, whose Y+3 scenario was extremely close to the robustness limit for PTV coverage, presented an SD_{95} of 4.626% and an SD_{mean} of 1.415% of the prescribed dose, once again metrics bordering on the acceptable tolerance values. Nevertheless, the limits were met in both analyses, confirming that this case is also robust.

To complement the voxel-level analysis, the robustness maps generated by the script were exported to 3D Slicer, along with the CT scans and structure sets (with the PTV delineated in red) for each patient. The subsequent Figure 25 presents the least (a) and most (b) robust cases for the PTV.

a)



b)

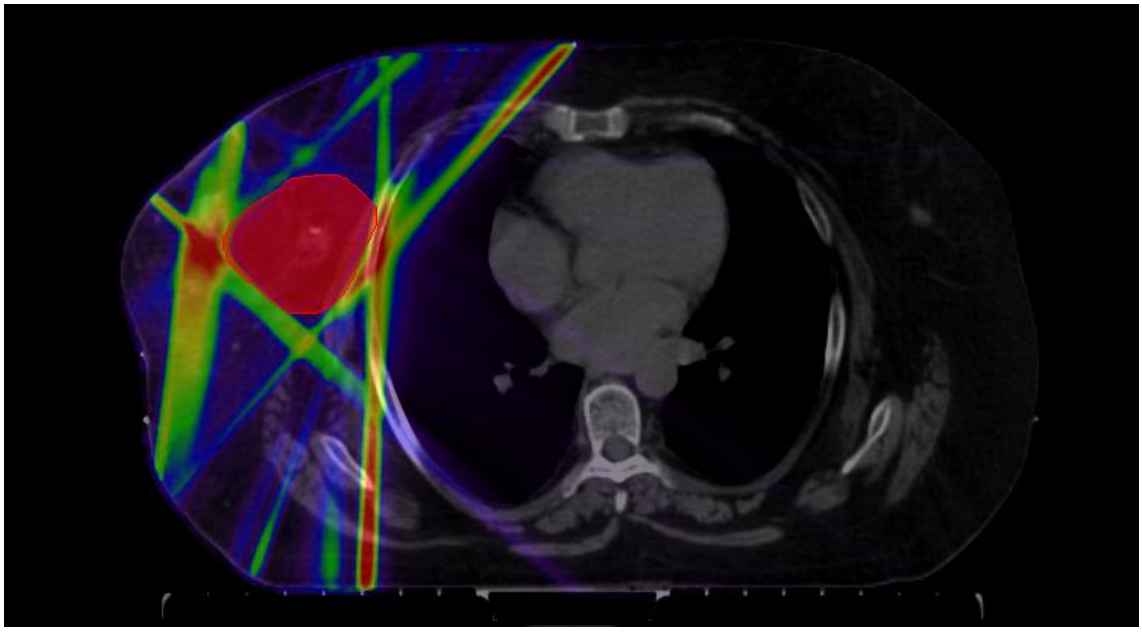


Figure 25 - Robustness maps for the least (a) and most (b) robust cases regarding the PTV.

4.2.2 Heart evaluation

Regarding cardiac robustness, all cases met the established OAR thresholds, namely $\mathrm{SD}_{95} < 4\%\mathrm{D}_{\mathrm{p}}$ and $\mathrm{SD}_{\mathrm{mean}} < 2\%\mathrm{D}_{\mathrm{p}}$, and were classified as robust. Nonetheless, case 3 stood out as the least robust in the voxel-wise heart evaluation, exhibiting the highest maximum SD value across the cohort (21.241%), indicative of a highly localised dosimetric hot spot. It should be noted, however, that this metric alone does not fully characterise overall structural robustness, and the case remained within acceptable limits. The observed behaviour is nonetheless explained by the specific tumour anatomy of this patient, as illustrated in Figure 26. The tumour cavity lies in close anatomical proximity to the heart (delineated in pink), and under the X-3 perturbation scenario, as shown in the right panel of Figure 26, the 95% isodose line advances notably towards the delineated heart.

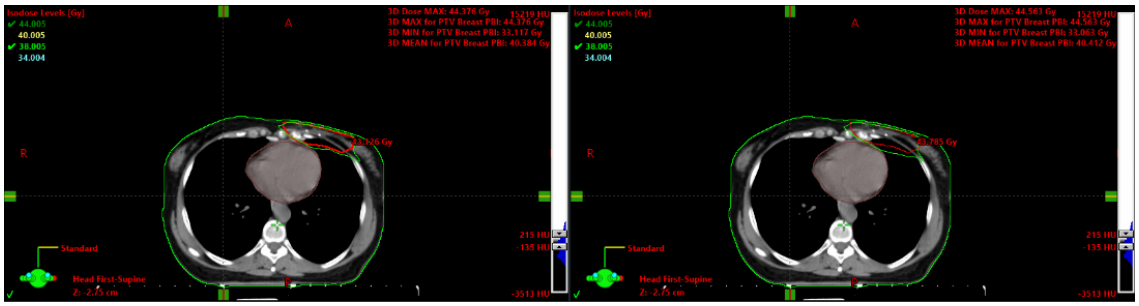
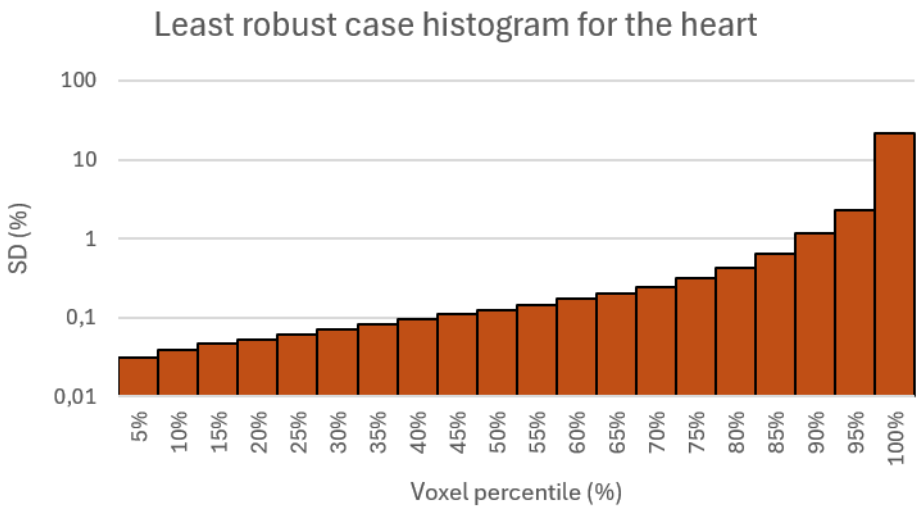


Figure 26 - Dose distribution maps for case 3 in Eclipse™, shown for the X-3 perturbation scenario (right panel) and the nominal plan (left panel). The tumour cavity is situated in close proximity to the heart (delineated in pink), and in the worst-case scenario (X-3), the 95% isodose line is seen to approach the cardiac contour considerably, accounting for the elevated SD values observed for this structure and identifying patient 3 as the least robust cardiac case in the voxel-wise analysis.

As previously discussed, this finding was anatomically predictable as a rightward shift (X-3) increases cardiac exposure in left-sided cases like patient 3, worsening OAR dosimetry.

To further illustrate the differences in heart dose analysis, the voxel-wise SD histograms for the least and most robust cases are shown in Figure 27 (a–b). The y-axis was plotted on a logarithmic scale to ensure the low-SD bins remained visible, as they would have been nearly invisible on a linear scale.

a)



b)

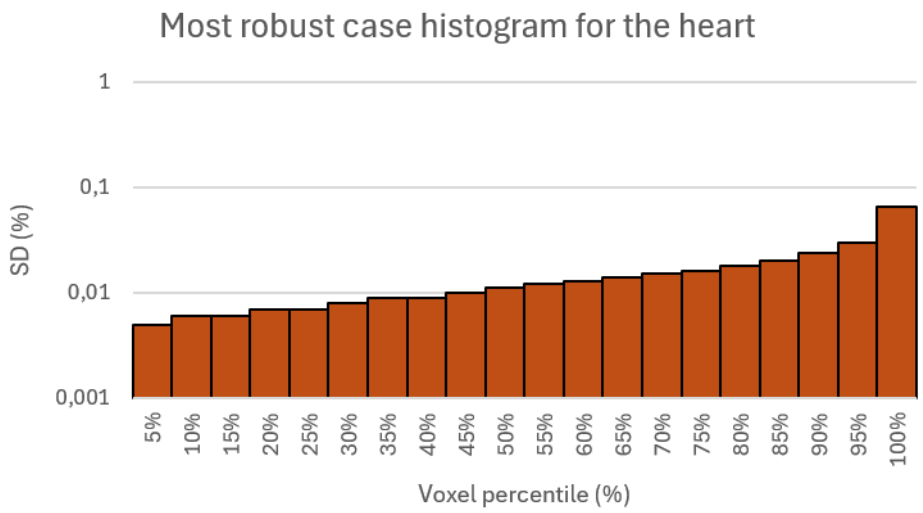


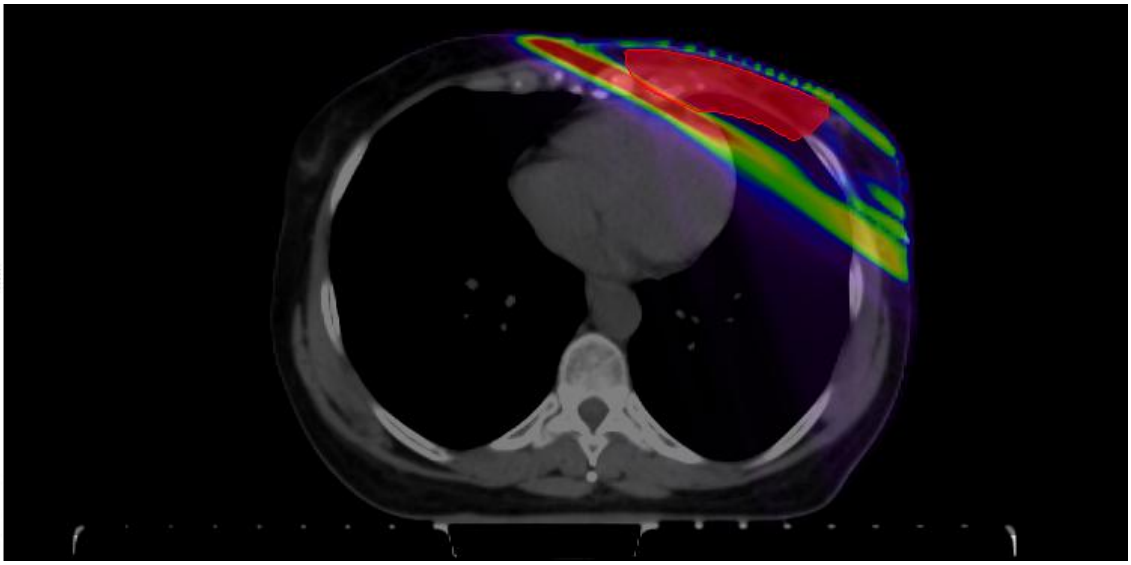
Figure 27 - Voxel-wise dose standard deviation histograms for the least robust (3, a) and most robust (19, b) cardiac cases, displayed on a logarithmic y-axis scale. The markedly broader distribution and higher SD values observed for case 3 reflect its greater sensitivity to geometric uncertainties, in contrast to case 19, the most robust, whose distribution is concentrated at substantially lower SD values.

As expected, the least robust case (patient 3) showed a broader distribution extending towards larger standard deviations, reflecting greater dosimetric variability due to geometric and mechanical uncertainties. The most robust case, in contrast, had a narrower distribution concentrated at lower SD values, suggesting minimal sensitivity to the perturbation scenarios evaluated.

Overall, these findings highlighted the complementary value of voxel-wise analysis relative to DVH-based robustness metrics. This was most evident for the X-3 scenario, the worst case in this study, which already bordered the DVH tolerance limit and, when assessed at the voxel level, further revealed localised hot spots within the cardiac volume. Although patient 3 remained within acceptable robustness limits, the combined assessment underscores the value of spatially resolved evaluation in identifying the cardiac subregions most susceptible to dosimetric uncertainty.

The following Figure 28 shows the least and most robust cases regarding the heart at the voxel level, with PTV represented in red.

a)



b)

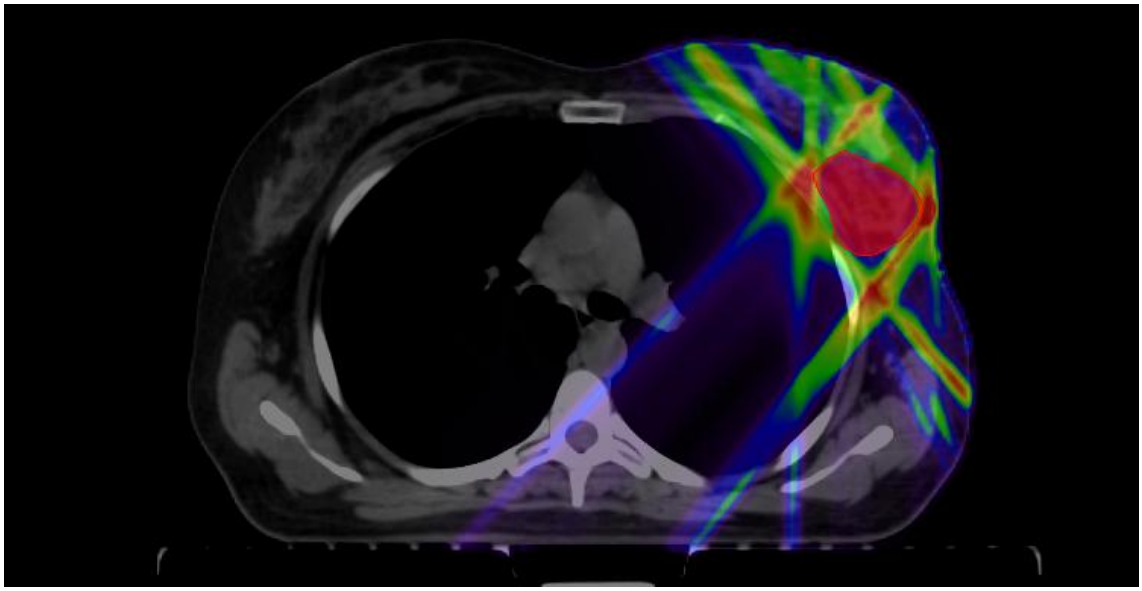


Figure 28 - Robustness maps for the least (a) and most (b) robust cases regarding the heart.

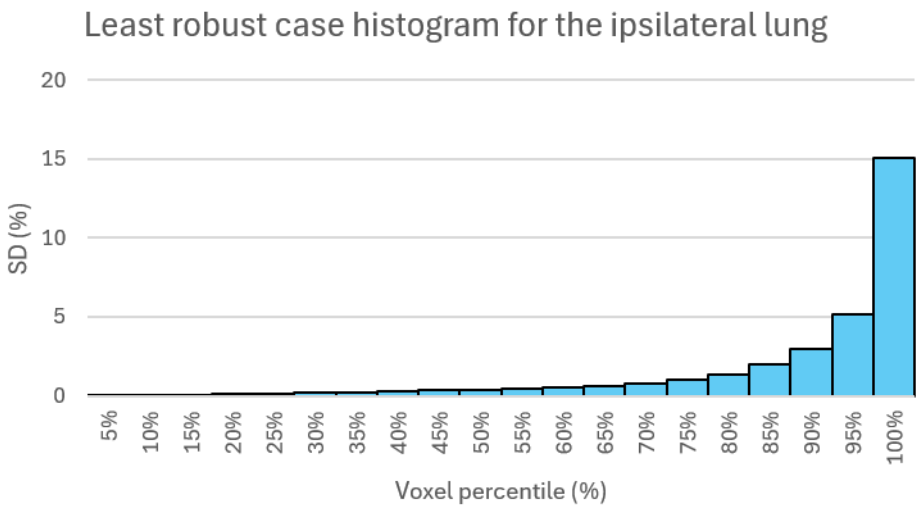
4.2.3 Ipsilateral lung evaluation

Regarding the ipsilateral lung, its large volume, approximately ten times that of the PTV in most cases would be expected to attenuate the dosimetric impact of 3 mm isocenter shifts through a volume dilution effect. Nevertheless, six cases failed the second robustness barrier, presenting SD mean values exceeding 4% of the prescribed dose and being classified as partially robust. This can be explained by the non-uniform dose distribution within the lung. In breast radiotherapy, irradiation is concentrated in the

anterior-medial region adjacent to the treatment field, and setup shifts that alter the extent of this irradiated zone can produce meaningful changes in the global mean dose, even in large lung volumes.

The least robust case was patient 27, whose tumour cavity lies near the costal wall, resulting in partial inclusion of the lung within the PTV. Given the relatively small lung volume of the patient, a Deep Inspiration Breath Hold (DIBH) technique would likely be beneficial, as the pulmonary expansion achieved during breath-holding increases the distance between the chest wall and the lung parenchyma, thereby reducing the dose delivered to this structure. The most robust case was patient 21, whose dosimetric variability across all perturbation scenarios remained minimal. The voxel-wise SD histograms for both cases are shown in the Figures 29 (a-b). Patient 27 exhibited a broader distribution extending towards higher SD values, while patient 40 showed a markedly narrower distribution concentrated at lower SD values, a visual contrast that clearly reflects the difference in dosimetric robustness between the two cases.

a)



b)

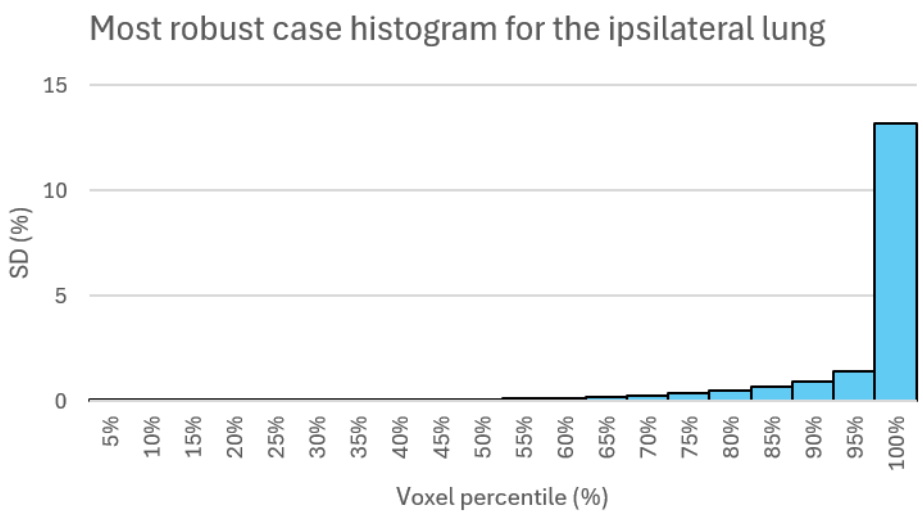
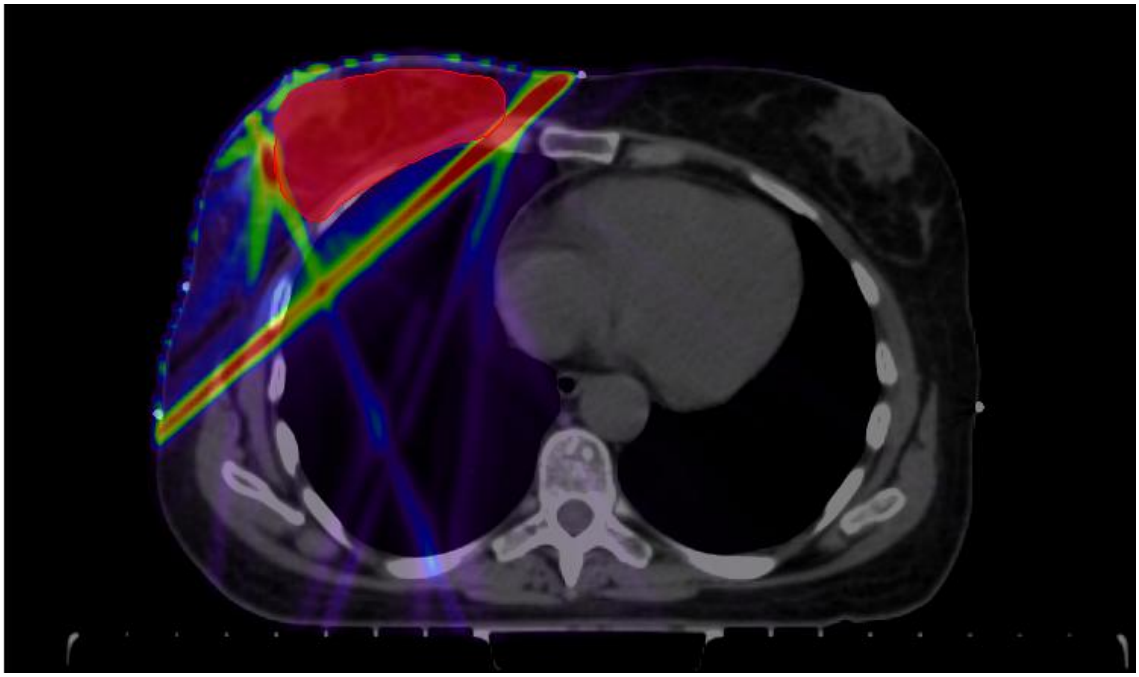


Figure 29 - Voxel-wise dose standard deviation histograms for the least robust (27, a) and most robust (21, b) ipsilateral lung cases. Patient 27 exhibited a broader distribution extending towards higher SD values, consistent with the partial inclusion of the lung within the PTV and its proximity to the costal wall. Patient 21, in contrast, showed a markedly narrower distribution concentrated at lower SD values, reflecting greater dosimetric stability across the evaluated perturbation scenarios.

Figure 30 shows the least and most robust cases regarding the ipsilateral lung.

a)



b)

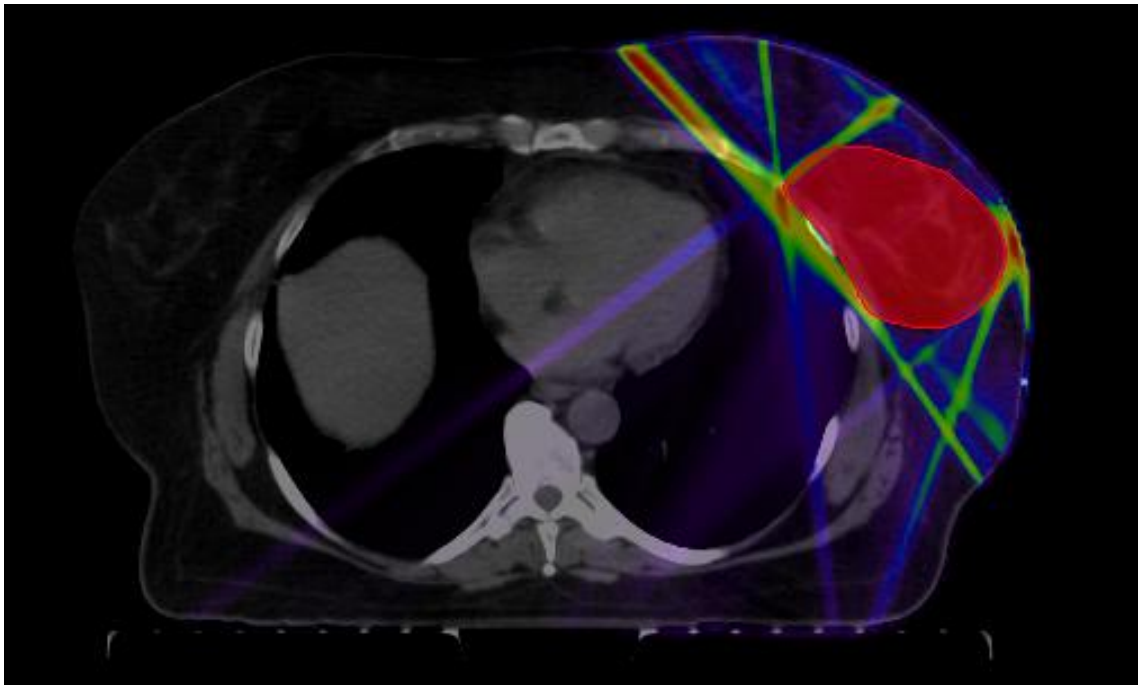


Figure 30 - Robustness maps for the least (a) and most (b) robust cases regarding the ipsilateral lung.

4.3 Complexity Analysis

Plan complexity is a fundamental dimension of IMRT treatment plan quality, complementary to dosimetric coverage and robustness (Hernandez et al., 2020). In IMRT, complexity arises from the degree of modulation of machine parameters, particularly MLC aperture shaping and MU delivery. Higher complexity is associated with increased uncertainty in dose calculation and plan deliverability, with potential consequences for patient-specific quality assurance outcomes (Duan et al., 2023).

Across the cohort, the MF ranged from 2.611 to 5.155 MU/cGy (mean 3.48 ± 0.63) and the EM ranged from 0.058 to 0.174 mm^{-1} (mean 0.085 ± 0.024), while the (RI) ranged from 0.091 to 0.914 (mean 0.63 ± 0.16). Patient 3, the least robust case regarding the heart, revealing a RI of 0.091, also showed the highest complexity in the cohort, with MF = 5.155 MU/cGy and EM = 0.144 mm^{-1}, both the highest complexity registered across the dataset.

A table with all the robustness indices and the complexity metrics is provided in Attachment 4 - Patient-specific complexity metrics and robustness indices.

Shapiro-Wilk tests showed that none of the three variables followed a normal distribution (RI: $W = 0.933$, $p = 0.018$; MF: $W = 0.924$, $p = 0.009$; EM: $W = 0.782$, $p < 0.001$),

warranting non-parametric analysis, as the Spearman statistical analysis performed by Hernandez et. al. (Hernandez et al., 2018). Spearman's rank correlation was therefore applied to examine the relationship between plan robustness and complexity, and between the two complexity indices themselves.

No statistically significant correlation was found between robustness and complexity, as the RI correlated with neither MF (rho = 0.093, p = 0.565) nor EM (rho = −0.173, p = 0.280). Within the dataset, plan complexity was therefore not a reliable predictor of overall robustness. This is illustrated by patient 26, who had the second-lowest RI of the cohort (0.219) despite MF and EM values close to the group averages (3.476 MU/cGy and 0.088 mm^{-1}, respectively), a clear case of low robustness driven by patient-specific anatomical factors rather than plan complexity. Case 3 represented the opposite extreme. The only patient combining maximal complexity with the lowest RI in the cohort, but as an isolated point it does not establish a general trend across the remaining 40 cases.

In contrast, the two complexity indices were significantly and moderately correlated (rho = 0.577, p < 0.001), indicating that MF and EM are complementary, each capturing a distinct physical aspect of treatment delivery, thus offering a more multidimensional characterization of plan complexity.

The absence of a significant correlation between plan robustness and complexity reflects the quality of the planning workflow itself. As all plans were optimized and the comprehensive DVH-based and voxel-wise analyses proved they were indeed robust. Even comparatively complex plans were rendered sufficiently robust to dose perturbations, effectively decoupling complexity from clinical robustness outcomes in this cohort. This can be explained by the fact that clinical radiotherapy practice does not aim merely to meet the clinical goals, but rather to apply the as low as reasonably achievable (ALARA) principle.

This study presented limitations that should be acknowledged. First, the 41 cohort is a modest sample size for robust non-parametric correlation analysis. Therefore, the absence of a significant relationship between robustness and complexity should be interpreted with caution, as it may reflect limited statistical power rather than a true absence of association in a larger population. Second, the analysis relied on two specific complexity metrics, the modulation factor and the edge metric, which, although complementary, do not capture the full multidimensional nature of plan complexity. Other established metrics such as the modulation complexity score, plan irregularity or leaf travel indices might reveal different associations with robustness. Third, robustness was

assessed using a fixed set of dose perturbation scenarios, which, while representative of common sources of uncertainty (setup and range errors), may not fully reflect the complete spectrum of clinical uncertainties encountered during treatment delivery, such as inter-fractional anatomical changes or organ motion beyond the modelled scenarios. Finally, this analysis was conducted within a single institution and treatment planning workflow, which may limit the generalizability of the findings to centers with different planning protocols, optimization algorithms, or delivery systems.

5. Conclusion

The main objective of this dissertation was to evaluate robustness of partial breast IMRT plans considering geometric and mechanical uncertainties and define a robustness index. As an universally accepted framework for robustness analysis is yet to be established, DVH- and voxel-based metrics were employed to access plan robustness, as recommended by Kim et al. (Kim et al., 2025).

Plans were considered robust against positioning shifts of $\pm$3 mm and 1 mm DLG variations, with outliers being explained by patient specific anatomy and tumour bed location. Therefore, action levels described in the institutional protocol for positioning, image acquisition, and verification of breast cancer radiotherapy treatments are consistent with the conducted study, thus not requiring an update.

In the DVH-based analysis, the X+3 scenarios were the least robust, followed by variations in the Y and Z directions. MLC-related uncertainties revealed minimal impact compared to the simulated setup deviations. Voxel-wise analysis further complemented these results, adding spatial information, which allowed precise localization of the most vulnerable regions, mainly in the high gradient dose areas. Based on DVH and voxel metrics, a robustness index was created. Left and right-sided cases revealed similar robustness indexes, with mean values of 0.627 and 0.635, respectively.

Finally, plan complexity evaluation, using the modulation factor and the edge metric, was introduced to assess whether robustness and complexity were correlated. However, the statistical analysis between plan complexity and robustness was not significant, except for specific cases where the clinical goals for the nominal plan were extremely close to the target coverage and OAR sparing thresholds.

Future work should aim to validate these findings in larger, institutional cohorts to strengthen the statistical power of the robustness-complexity relationship. Additionally, extending the robustness analysis to include a wider range of perturbation scenarios, such as inter-fractional anatomical variation and respiratory motion, would provide a more comprehensive characterization of plan robustness under real clinical conditions.

Furthermore, including a probabilistic robustness evaluation, instead of a discrete analysis, would allow the incorporation of machine learning–based models, which could predict a complexity threshold balancing plan quality and robustness.

Finally, integrating complexity and robustness metrics directly into the treatment planning optimization, rather than evaluating them retrospectively, may support the decision-making process, ensuring consistency between planned and administrated dose.

References

Byrne, M., Hu, Y., & Archibald-Heeren, B. (2016). Evaluation of RayStation robust optimisation for superficial target coverage with setup variation in breast IMRT. *Australasian Physical & Engineering Sciences in Medicine*, *39*(3), 705–716. https://doi.org/10.1007/s13246-016-0466-6

Chan, R. C. K., Ng, C. K. C., Hung, R. H. M., Li, Y. T. Y., Tam, Y. T. Y., Wong, B. Y. L., Yu, J. C. K., & Leung, V. W. S. (2023). Comparative Study of Plan Robustness for Breast Radiotherapy: Volumetric Modulated Arc Therapy Plans with Robust Optimization versus Manual Flash Approach. *Diagnostics*, *13*(22), 3395. https://doi.org/10.3390/diagnostics13223395

Chang, J. S., Chang, J. H., Kim, N., Kim, Y. B., Shin, K. H., & Kim, K. (2022). Intensity Modulated Radiotherapy and Volumetric Modulated Arc Therapy in the Treatment of Breast Cancer: An Updated Review. *Journal of Breast Cancer*, *25*(5), 349. https://doi.org/10.4048/jbc.2022.25.e37

Correa, C., Harris, E. E., Leonardi, M. C., Smith, B. D., Taghian, A. G., Thompson, A. M., White, J., & Harris, J. R. (2017). Accelerated Partial Breast Irradiation: Executive summary for the update of an ASTRO Evidence-Based Consensus Statement. *Practical Radiation Oncology*, *7*(2), 73–79. https://doi.org/10.1016/j.prro.2016.09.007

Darby, S. C., Ewertz, M., McGale, P., Bennet, A. M., Blom-Goldman, U., Brønnum, D., Correa, C., Cutter, D., Gagliardi, G., Gigante, B., Jensen, M.-B., Nisbet, A., Peto, R., Rahimi, K., Taylor, C., & Hall, P. (2013). Risk of Ischemic Heart Disease in Women after Radiotherapy for Breast Cancer. *New England Journal of Medicine*, *368*(11), 987–998. https://doi.org/10.1056/NEJMoa1209825

Das Majumdar, S. K., Amritt, A., Dhar, S. S., Barik, S., Beura, S. S., Mishra, T., Muduly, D. K., Dash, A., & Parida, D. K. (2022). A Dosimetric Study Comparing 3D-CRT vs. IMRT vs. VMAT in Left-Sided Breast Cancer Patients After Mastectomy at a

Tertiary Care Centre in Eastern India. *Cureus*. https://doi.org/10.7759/cureus.23568

Ding, Z., Zeng, Q., Kang, K., Xu, M., Xiang, X., & Liu, C. (2022). Evaluation of Plan Robustness Using Hybrid Intensity-Modulated Radiotherapy (IMRT) and Volumetric Arc Modulation Radiotherapy (VMAT) for Left-Sided Breast Cancer. *Bioengineering*, *9*(4), 131. https://doi.org/10.3390/bioengineering9040131

Duan, L., Qi, W., Chen, Y., Cao, L., Chen, J., Zhang, Y., & Xu, C. (2023). Evaluation of complexity and deliverability of IMRT treatment plans for breast cancer. *Scientific Reports*, *13*(1), 21474. https://doi.org/10.1038/s41598-023-48331-x

Enomoto, H., Fujita, Y., Matsumoto, S., Nakajima, Y., Nagai, M., Tonari, A., & Ebara, T. (2024). Dosimetric impact of MLC positional errors on dose distribution in IMRT. *Journal of Applied Clinical Medical Physics*, *25*(2), e14158. https://doi.org/10.1002/acm2.14158

Feuvret, L., Noël, G., Mazeron, J.-J., & Bey, P. (2006). Conformity index: A review. *International Journal of Radiation Oncology*Biology*Physics*, *64*(2), 333–342. https://doi.org/10.1016/j.ijrobp.2005.09.028

Gueiderikh, A., Sarrade, T., Kirova, Y., De La Lande, B., De Vathaire, F., Auzac, G., Martin, A. L., Everhard, S., Meillan, N., Bourgier, C., Benyoucef, A., Lacornerie, T., Pasquier, D., Racadot, S., Moignier, A., Paris, F., André, F., Deutsch, E., Duchemann, B., … Rivera, S. (2023). Radiation-induced lung injury after breast cancer treatment: Incidence in the CANTO-RT cohort and associated clinical and dosimetric risk factors. *Frontiers in Oncology*, *13*, 1199043. https://doi.org/10.3389/fonc.2023.1199043

Handoyo, T., Palupi, I., Setiaji, G., & Firmansyah, T. (2026). Simulation-based design of a pulse modulator for a 6 MeV medical LINAC using LTspice. *Discover Electronics*, *3*(1), 56. https://doi.org/10.1007/s44291-026-00208-9

He, Y., Chen, S., Gao, X., Fu, L., Kang, Z., Liu, J., Shi, L., & Li, Y. (2023). Robustness of VMAT to setup errors in postmastectomy radiotherapy of left-sided breast cancer: Impact of bolus thickness. *PLOS ONE*, *18*(1), e0280456. https://doi.org/10.1371/journal.pone.0280456

Hernandez, V., Hansen, C. R., Widesott, L., Bäck, A., Canters, R., Fusella, M., Götstedt, J., Jurado-Bruggeman, D., Mukumoto, N., Kaplan, L. P., Koniarová, I., Piotrowski, T., Placidi, L., Vaniqui, A., & Jornet, N. (2020). What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans. *Radiotherapy and Oncology*, *153*, 26–33. https://doi.org/10.1016/j.radonc.2020.09.038

Hernandez, V., Lara-Aristimuño, I., Abella, R., & Saez, J. (2025). Quantification of the aperture modulation in radiotherapy treatment plans. *Medical Physics*, *52*(12), e70144. https://doi.org/10.1002/mp.70144

Hernandez, V., Saez, J., Pasler, M., Jurado-Bruggeman, D., & Jornet, N. (2018). Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy. *Physics and Imaging in Radiation Oncology*, *5*, 37–43. https://doi.org/10.1016/j.phro.2018.02.002

Hunte, S. O., Clark, C. H., Zyuzikov, N., & Nisbet, A. (2022). Volumetric modulated arc therapy (VMAT): A review of clinical outcomes—what is the clinical evidence for the most effective implementation? *The British Journal of Radiology*, *95*(1136), 20201289. https://doi.org/10.1259/bjr.20201289

ICRU. (2010). Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT): Contents. *Journal of the ICRU*, *10*(1), NP.3-NP. https://doi.org/10.1093/jicru/ndq002

Ju, E., Heo, E. J., Park, C. G., Kim, M., Kim, K. H., Shim, J. B., Park, Y. J., Lee, N. K., Kim, C. Y., & Lee, S. (2021). Dosimetric comparison of VitalBeam® and Halcyon™ 2.0 for hypofractionated VMAT with simultaneous integrated boost

treatment of early-stage left-sided breast cancer. *Journal of Applied Clinical Medical Physics*, *22*(10), 232–238. https://doi.org/10.1002/acm2.13428

Kaplan, L. P., Placidi, L., Bäck, A., Canters, R., Hussein, M., Vaniqui, A., Fusella, M., Piotrowski, T., Hernandez, V., Jornet, N., Hansen, C. R., & Widesott, L. (2022). Plan quality assessment in clinical practice: Results of the 2020 ESTRO survey on plan complexity and robustness. *Radiotherapy and Oncology*, *173*, 254–261. https://doi.org/10.1016/j.radonc.2022.06.005

Kim, S., Bernstein, D., & Taylor, A. (2025). Evaluation of Radiotherapy Plan Robustness: A Systematic Literature Review. *Clinical Oncology*, *48*, 103938. https://doi.org/10.1016/j.clon.2025.103938

Koka, K., Verma, A., Dwarakanath, B. S., & Papineni, R. V. (2022). Technological Advancements in External Beam Radiation Therapy (EBRT): An Indispensable Tool for Cancer Treatment. *Cancer Management and Research*, *Volume 14*, 1421–1429. https://doi.org/10.2147/CMAR.S351744

Loebner, H. A., Bertholet, J., Mackeprang, P.-H., Volken, W., Fix, M. K., & Manser, P. (2025). Robustness assessment of radiotherapy treatment plans in Switzerland. *Zeitschrift Für Medizinische Physik*, *36*(1), 47–59. https://doi.org/10.1016/j.zemedi.2025.03.002

Mast, M. E., Leong, A., Korreman, S. S., Lee, G., Probst, H., Scherer, P., & Tsang, Y. (2023). ESTRO-ACROP guideline for positioning, immobilisation and setup verification for local and loco-regional photon breast cancer irradiation. *Technical Innovations & Patient Support in Radiation Oncology*, *28*, 100219. https://doi.org/10.1016/j.tipsro.2023.100219

Meattini, I., Becherini, C., Boersma, L., Kaidar-Person, O., Marta, G. N., Montero, A., Offersen, B. V., Aznar, M. C., Belka, C., Brunt, A. M., Dicuonzo, S., Franco, P., Krause, M., MacKenzie, M., Marinko, T., Marrazzo, L., Ratosa, I., Scholten, A., Senkus, E., … Coles, C. E. (2022). European Society for Radiotherapy and

Oncology Advisory Committee in Radiation Oncology Practice consensus recommendations on patient selection and dose and fractionation for external beam radiotherapy in early breast cancer. *The Lancet Oncology*, *23*(1), e21–e31. https://doi.org/10.1016/S1470-2045(21)00539-8

Miyasaka, Y., Ono, T., Chai, H., Souda, H., Lee, S. H., Ishizawa, M., Akamatsu, H., Sato, H., & Iwai, T. (2024). A robust treatment planning approach for chest motion in postmastectomy chest wall intensity modulated radiation therapy. *Journal of Applied Clinical Medical Physics*, *25*(1), e14217. https://doi.org/10.1002/acm2.14217

Podgorsak, E. B. (Ed.). (2005). *Radiation Oncology Physics: A Handbook for Teachers and Students*. International Atomic Energy Agency (IAEA).

Rastogi, K., Sharma, S., Gupta, S., Agarwal, N., Bhaskar, S., & Jain, S. (2018). Dosimetric comparison of IMRT versus 3DCRT for post-mastectomy chest wall irradiation. *Radiation Oncology Journal*, *36*(1), 71–78. https://doi.org/10.3857/roj.2017.00381

RaySearch Laboratories. (2019a). *Evaluation of Robustness in RayStation*. https://www.raysearchlabs.com

RaySearch Laboratories. (2019b). *Evaluation of Robustness in RayStation*. RaySearch Laboratories. https://www.raysearchlabs.com/media/publications/white-papers/evaluation-of-robustness-in-raystation/

RON. (2024). *Registo Oncológico Nacional de Todos os Tumores na População Residente em Portugal, em 2021* (No. 2021). Instituto Português de Oncologia do Porto Francisco Gentil - EPE.

Rostami, A., Barzegar, M., Usman, M., Paloor, S. P., Mkanna, A. Y., Al-Sabahi, A. F., & Hammoud, R. W. (2024). Technical Note: Investigating of dosimetric leaf gap and leaf transmission factor variations across gantry and collimator angles in

volumetric modulation arc therapy. *Journal of Applied Clinical Medical Physics*, *25*(12), e14523. https://doi.org/10.1002/acm2.14523

Schlegel, W., Bortfeld, T., & Grosu, A.-L. (Eds). (2006). *New Technologies in Radiation Oncology*. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/3-540-29999-8

Schlegel, W., Grosser, K. H., Häring, P., & Rhein, B. (2006). Beam Delivery in 3D Conformal Radiotherapy Using Multi-Leaf Collimators. In W. Schlegel, T. Bortfeld, & A.-L. Grosu (Eds), *New Technologies in Radiation Oncology* (pp. 257–266). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-29999-8_20

Shaitelman, S. F. (2024). *Partial Breast Irradiation for Patients With Early-Stage Invasive Breast Cancer or Ductal Carcinoma In Situ: An ASTRO Clinical Practice Guideline*.

Shende, R., & Patel, G. (2017). Validation of Dosimetric Leaf Gap (DLG) prior to its implementation in Treatment Planning System (TPS): TrueBeam™ millennium 120 leaf MLC. *Reports of Practical Oncology & Radiotherapy*, *22*(6), 485–494. https://doi.org/10.1016/j.rpor.2017.09.001

Teoh, M., Clark, C. H., Wood, K., Whitaker, S., & Nisbet, A. (2011). Volumetric modulated arc therapy: A review of current literature and clinical use in practice. *The British Journal of Radiology*, *84*(1007), 967–996. https://doi.org/10.1259/bjr/22373346

Tyran, M., Mailleux, H., Tallet, A., Fau, P., Gonzague, L., Minsat, M., Moureau-Zabotto, L., & Resbeut, M. (2015). Volumetric-modulated arc therapy for left-sided breast cancer and all regional nodes improves target volumes coverage and reduces treatment time and doses to the heart and left coronary artery, compared with a field-in-field technique. *Journal of Radiation Research*, *56*(6), 927–937. https://doi.org/10.1093/jrr/rrv052

Wang, H., Cooper, B. T., Schiff, P., Sanfilippo, N. J., Wu, S. P., Hu, K. S., Das, I. J., & Xue, J. (2019). Dosimetric assessment of tumor control probability in intensity

and volumetric modulated radiotherapy plans. *The British Journal of Radiology*, *92*(1094), 20180471. https://doi.org/10.1259/bjr.20180471

Webb, S. (2003). 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. *Physics in Medicine and Biology*, *48*(14), 2051–2062. https://doi.org/10.1088/0031-9155/48/14/301

World Health Organization. (2021, March 8). *New global breast cancer initiative highlights renewed commitment to improve survival*. World Health Organization. https://www.who.int/news/item/08-03-2021-new-global-breast-cancer-initiative-highlights-renewed-commitment-to-improve-survival

World Health Organization. (2024, February 1). *Global cancer burden growing, amidst mounting need for services*. World Health Organization. https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services

Yin, C., Deng, J., Mei, G., Cheng, H., He, Y., & Liu, J. (2024). Comparison of plan quality and robustness using VMAT and IMRT for breast cancer. *Open Physics*, *22*(1), 20240026. https://doi.org/10.1515/phys-2024-0026

Yock, A. D., Mohan, R., Flampouri, S., Bosch, W., Taylor, P. A., Gladstone, D., Kim, S., Sohn, J., Wallace, R., Xiao, Y., & Buchsbaum, J. (2019). Robustness Analysis for External Beam Radiation Therapy Treatment Plans: Describing Uncertainty Scenarios and Reporting Their Dosimetric Consequences. *Practical Radiation Oncology*, *9*(4), 200–207. https://doi.org/10.1016/j.prro.2018.12.002

Younge, K. C., Roberts, D., Janes, L. A., Anderson, C., Moran, J. M., & Matuszak, M. M. (2016). Predicting deliverability of volumetric-modulated arc therapy (VMAT) plans using aperture complexity analysis. *Journal of Applied Clinical Medical Physics*, *17*(4), 124–131. https://doi.org/10.1120/jacmp.v17i4.6241

Zhang, Q., Liu, J., Ao, N., Yu, H., Peng, Y., Ou, L., & Zhang, S. (2020). Secondary cancer risk after radiation therapy for breast cancer with different radiotherapy techniques. *Scientific Reports*, *10*(1), 1220. https://doi.org/10.1038/s41598-020-58134-z

Attachment 1 - Abstract accepted for poster presentation at the 6th European Congress of Medical Physics

ROBUSTNESS EVALUATION OF BREAST IMRT PLANS AGAINST GEOMETRIC AND MECHANICAL UNCERTAINTIES

A. Oliveira[1,2], C. Souto[1,3], S. Gonçalves[1,3], I. Almeida[1,3], A.G. Dias[1,3]

[1]Medical Physics, Radiobiology and Radiation Protection Group, IPO Porto Research Center (CI-IPO), Portuguese Oncology Institute of Porto (IPO/Porto) / Porto Comprehensive Cancer Centre (Porto.CCC) & RISE@CI-IPO (Health Research Network), Porto, Portugal
[2]Physics and Astronomy Department, Universidade do Porto, Faculdade de Ciências, Porto, Portugal
[3]Medical Physics Department, Instituto Português de Oncologia do Porto FG, EPE, Porto, Portugal

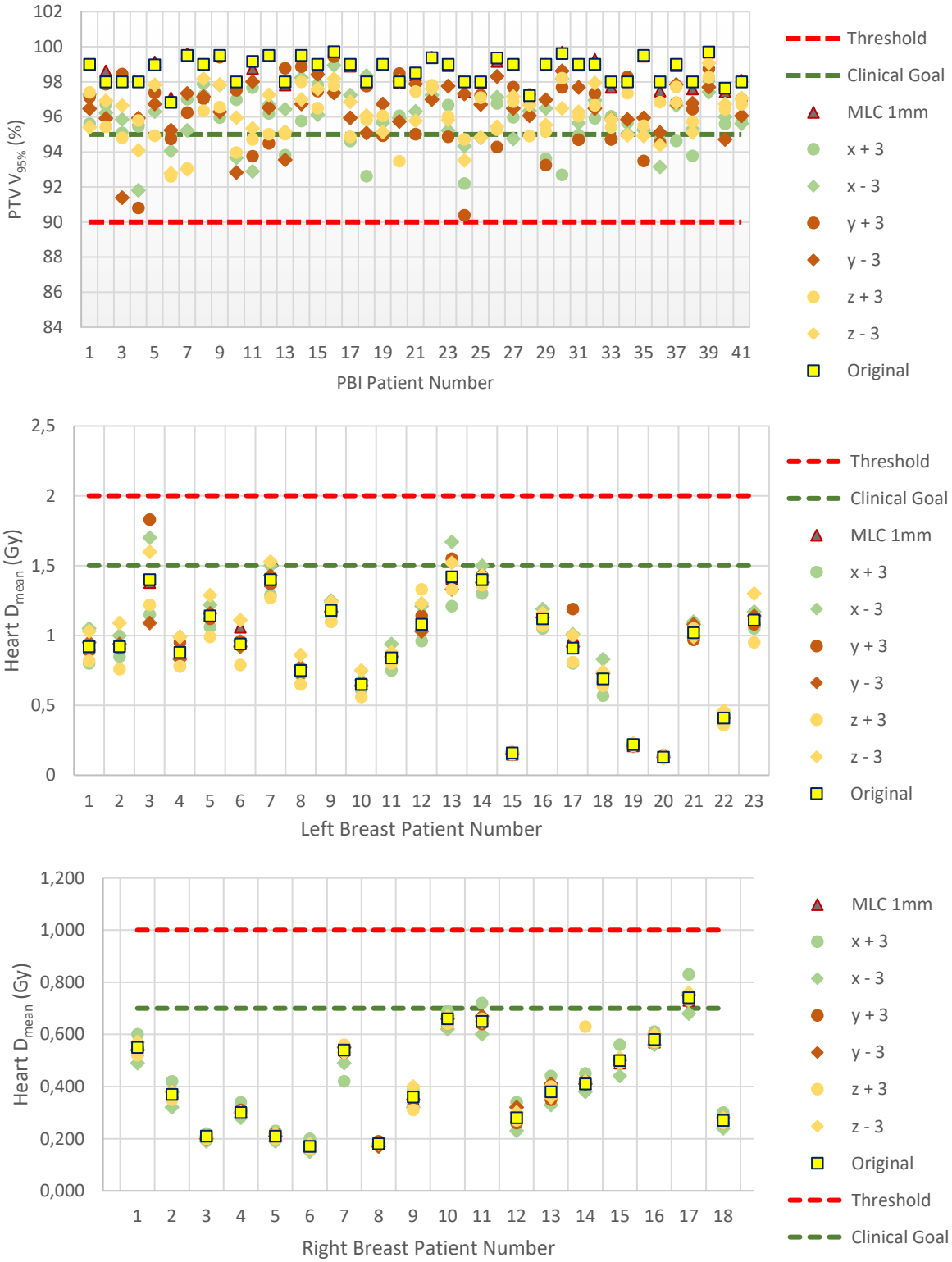
Background and Aims
Robustness is a cornerstone of modern Radiotherapy, as it assesses a plan´s capacity to handle real-world uncertainties. Rather than relying on dose distributions optimized for static conditions, robustness ensures target coverage and OAR sparing despite motion, positioning and calibration errors that may occur during treatment delivery. This retrospective study evaluates the robustness of partial breast IMRT plans against geometric and mechanical uncertainties.

Methods and Materials
A cohort of 41 patients treated for Partial Breast Irradiation with IMRT was selected. Right and left-sided cases were evaluated separately for mean heart dose analysis. For each patient, seven perturbation scenarios were created, including patient setup errors of ±3 mm in all translational directions and a ±1 mm variation in the MLC Dosimetric Leaf Gap. A DVH-based metrics was used for assessing plan robustness, analysing variations in clinical goals (PTV_V95% and Heart_Dmean) between planned and perturbated dose distribution. Plans were considered robust when clinical goals remained within the accepted variation limits across all scenarios.

Resuts
Scatter plots (Figures 1-3) displaying the nominal plans and perturbed scenarios were used to evaluate plan robustness. For right-sided breast cases, X+3 plan revealed to be the worst in terms of Heart Dmean. Overall, variations in the MLC have minimal impacts, followed by variations in Y and Z directions. Future work includes more patient setup simulations to reach a level of action, a spatial dose variability distribution, for visual representation of areas most vulnerable to uncertainties, and the introduction of a gamma passing rate analysis.

Keywords
Robustness, IMRT, Breast

[1]



Figures 1-3: Scatter plots of PTV_V95% and Heart_Dmean. Yellow squares represent the nominal plans, while the remaining markers correspond to simulated setup uncertainties (±3 mm shifts in the x, y, and z directions) and MLC variations (1 mm in the Dosimetric Leaf Gap). The green and red dashed lines represent the clinical goal and

the robustness threshold, respectively. Points falling below this threshold in target coverage or exceeding its value in the mean heart dose indicate lack of robustness.

Attachment 2 - Eclipse™ scripting API code

```
using System;
using System.IO;
using System.Windows;
using System.Linq;
using System.Collections.Generic;
using VMS.TPS.Common.Model.API;
using VMS.TPS.Common.Model.Types;

namespace VMS.TPS
{
  public class Script
  {
    public void Execute(ScriptContext context)
    {
      Patient patient = context.Patient;
      Course course = context.Course;

      if (patient == null)
        throw new ApplicationException("Patient not loaded");

      //--------------------------------------
      // GET THE 8 PLANS
      //--------------------------------------
      // Takes the first 8 plans from the current course
```

```
var plans = course.PlanSetups.Take(8).ToList();
if (plans.Count < 8)
  throw new ApplicationException("8 plans are required");

// Reference plan: defines the dose grid geometry shared by all scenarios
PlanSetup refPlan = plans.First();
var structureSet = refPlan.StructureSet;
var dose = refPlan.Dose;

//--------------------------------------
// DOSE GRID
//--------------------------------------
VVector origin = dose.Origin;
int sizeX = dose.XSize;
int sizeY = dose.YSize;
int sizeZ = dose.ZSize;
double resX = dose.XRes;
double resY = dose.YRes;
double resZ = dose.ZRes;

//--------------------------------------
// MATRICES
//--------------------------------------
// doseMatrix: dose value for each of the 8 scenarios, at every voxel (x,y,z)
double[,,,] doseMatrix = new double[8, sizeX, sizeY, sizeZ];
```

```
double[,,] Dmax = new double[sizeX, sizeY, sizeZ];       // max dose across scenarios per voxel

double[,,] Dmin = new double[sizeX, sizeY, sizeZ];       // min dose across scenarios per voxel

double[,,] SD = new double[sizeX, sizeY, sizeZ];          // standard deviation across scenarios per voxel

double[,,] robustness = new double[sizeX, sizeY, sizeZ];  // Dmax - Dmin (dose range) per voxel


//---------------------------------------

// READ DOSE FROM EACH PLAN

//---------------------------------------

// Loops over the 8 plans and samples the dose at every voxel position

for (int s = 0; s < 8; s++)

{

  var d = plans[s].Dose;

  for (int k = 0; k < sizeZ; k++)

  {

    for (int j = 0; j < sizeY; j++)

    {

      for (int i = 0; i < sizeX; i++)

      {

        VVector voxel = new VVector(

          origin.x + i * resX,

          origin.y + j * resY,

          origin.z + k * resZ);

        doseMatrix[s, i, j, k] = d.GetDoseToPoint(voxel).Dose;
```

```
        }
       }
      }
     }

     //--------------------------------------
     // CALCULATE ROBUSTNESS
     //--------------------------------------
     for (int i1 = 0; i1 < sizeX; i1++)
     {
      for (int j1 = 0; j1 < sizeY; j1++)
      {
       for (int k1 = 0; k1 < sizeZ; k1++)
       {
        double max = double.MinValue;
        double min = double.MaxValue;
        double sum = 0;

        // Loop through the scenario plans and check the dose at this voxel (i1, j1, k1)
        for (int s1 = 0; s1 < 8; s1++)
        {
         double d = doseMatrix[s1, i1, j1, k1];
         if (d > max) max = d; // check if this scenario's dose is greater than the current
max
         if (d < min) min = d; // check if this scenario's dose is smaller than the current
min
```

```
        sum += d;
      }

      double mean = sum / 8;
      double variance = 0;
      for (int s1 = 0; s1 < 8; s1++)
      {
        double diff = doseMatrix[s1, i1, j1, k1] - mean;
        variance += diff * diff;
      }
      variance /= 8;

      Dmax[i1, j1, k1] = max;
      Dmin[i1, j1, k1] = min;
      SD[i1, j1, k1] = Math.Sqrt(variance);
      robustness[i1, j1, k1] = max - min;
    }
  }
}

//---------------------------------------
// EXTRACT PTV
//---------------------------------------
var ptv = structureSet.Structures
  .FirstOrDefault(s => !s.IsEmpty && s.DicomType == "PTV");
```

```
// Collect the SD value of every voxel that falls inside the PTV
List<double> sdPTV = new List<double>();
for (int i2 = 0; i2 < sizeX; i2++)
{
  for (int j2 = 0; j2 < sizeY; j2++)
  {
    for (int k2 = 0; k2 < sizeZ; k2++)
    {
      VVector voxel2 = new VVector(
        origin.x + i2 * resX,
        origin.y + j2 * resY,
        origin.z + k2 * resZ);
      if (ptv != null && ptv.IsPointInsideSegment(voxel2))
        sdPTV.Add(SD[i2, j2, k2]);
    }
  }
}

//---------------------------------------
// SDVH METRICS (PTV)
//---------------------------------------
// Sort SD values to compute percentile-based metrics (SDVH = SD-Volume Histogram)
sdPTV.Sort();
int N = sdPTV.Count;
double SDmean = sdPTV.Average();
```

```
double SDmax = sdPTV.Max();
double SD95 = sdPTV[(int)(0.95 * N)];
double SD90 = sdPTV[(int)(0.90 * N)];
double SD85 = sdPTV[(int)(0.85 * N)];
double SD80 = sdPTV[(int)(0.80 * N)];
double SD75 = sdPTV[(int)(0.75 * N)];
double SD70 = sdPTV[(int)(0.70 * N)];
double SD65 = sdPTV[(int)(0.65 * N)];
double SD60 = sdPTV[(int)(0.60 * N)];
double SD55 = sdPTV[(int)(0.55 * N)];
double SD50 = sdPTV[(int)(0.50 * N)];
double SD45 = sdPTV[(int)(0.45 * N)];
double SD40 = sdPTV[(int)(0.40 * N)];
double SD35 = sdPTV[(int)(0.35 * N)];
double SD30 = sdPTV[(int)(0.30 * N)];
double SD25 = sdPTV[(int)(0.25 * N)];
double SD20 = sdPTV[(int)(0.20 * N)];
double SD15 = sdPTV[(int)(0.15 * N)];
double SD10 = sdPTV[(int)(0.10 * N)];
double SD5 = sdPTV[(int)(0.05 * N)];

//---------------------------------------
// EXPORT RESULTS (PTV)
//---------------------------------------
// TODO: set the output folder path before running the script
string path = @"<SET_OUTPUT_PATH_HERE>\" + patient.Id + @"\";
```

```
Directory.CreateDirectory(path);


File.WriteAllText(path + "metrics.csv",
  "Metric,Value\n" +
  "SDmean," + SDmean.ToString("F3") + "\n" +
  "SD95," + SD95.ToString("F3") + "\n" +
  "SD90," + SD90.ToString("F3") + "\n" +
  "SD85," + SD85.ToString("F3") + "\n" +
  "SD80," + SD80.ToString("F3") + "\n" +
  "SD75," + SD75.ToString("F3") + "\n" +
  "SD70," + SD70.ToString("F3") + "\n" +
  "SD65," + SD65.ToString("F3") + "\n" +
  "SD60," + SD60.ToString("F3") + "\n" +
  "SD55," + SD55.ToString("F3") + "\n" +
  "SD50," + SD50.ToString("F3") + "\n" +
  "SD45," + SD45.ToString("F3") + "\n" +
  "SD40," + SD40.ToString("F3") + "\n" +
  "SD35," + SD35.ToString("F3") + "\n" +
  "SD30," + SD30.ToString("F3") + "\n" +
  "SD25," + SD25.ToString("F3") + "\n" +
  "SD20," + SD20.ToString("F3") + "\n" +
  "SD15," + SD15.ToString("F3") + "\n" +
  "SD10," + SD10.ToString("F3") + "\n" +
  "SD5," + SD5.ToString("F3") + "\n" +
  "SDmax," + SDmax.ToString("F3"));
```

```
//---------------------------------------
// EXTRACT Heart
//---------------------------------------
var heart = structureSet.Structures
  .FirstOrDefault(s => !s.IsEmpty && s.DicomType == "ORGAN" &&
  s.Id.ToLower().Contains("heart"));

List<double> sdHeart = new List<double>();
for (int i4 = 0; i4 < sizeX; i4++)
{
  for (int j4 = 0; j4 < sizeY; j4++)
  {
    for (int k4 = 0; k4 < sizeZ; k4++)
    {
      VVector voxel4 = new VVector(
        origin.x + i4 * resX,
        origin.y + j4 * resY,
        origin.z + k4 * resZ);
      if (heart != null && heart.IsPointInsideSegment(voxel4))
        sdHeart.Add(SD[i4, j4, k4]);
    }
  }
}

//---------------------------------------
```

```
// SDVH METRICS (Heart)
//---------------------------------------
sdHeart.Sort();
int NHeart = sdHeart.Count;
double SDmeanHeart = sdHeart.Average();
double SDmaxHeart = sdHeart.Max();
double SD95Heart = sdHeart[(int)(0.95 * NHeart)];
double SD90Heart = sdHeart[(int)(0.9 * NHeart)];
double SD85Heart = sdHeart[(int)(0.85 * NHeart)];
double SD80Heart = sdHeart[(int)(0.80 * NHeart)];
double SD75Heart = sdHeart[(int)(0.75 * NHeart)];
double SD70Heart = sdHeart[(int)(0.70 * NHeart)];
double SD65Heart = sdHeart[(int)(0.65 * NHeart)];
double SD60Heart = sdHeart[(int)(0.60 * NHeart)];
double SD55Heart = sdHeart[(int)(0.55 * NHeart)];
double SD50Heart = sdHeart[(int)(0.50 * NHeart)];
double SD45Heart = sdHeart[(int)(0.45 * NHeart)];
double SD40Heart = sdHeart[(int)(0.40 * NHeart)];
double SD35Heart = sdHeart[(int)(0.35 * NHeart)];
double SD30Heart = sdHeart[(int)(0.30 * NHeart)];
double SD25Heart = sdHeart[(int)(0.25 * NHeart)];
double SD20Heart = sdHeart[(int)(0.20 * NHeart)];
double SD15Heart = sdHeart[(int)(0.15 * NHeart)];
double SD10Heart = sdHeart[(int)(0.10 * NHeart)];
double SD5Heart = sdHeart[(int)(0.05 * NHeart)];
```

```
//---------------------------------------

// EXPORT RESULTS (Heart)

//---------------------------------------


File.WriteAllText(path + "metricsHeart.csv",

  "Metric,Value\n" +

  "SDmeanHeart," + SDmeanHeart.ToString("F3") + "\n" +

  "SD95Heart," + SD95Heart.ToString("F3") + "\n" +

  "SD90Heart," + SD90Heart.ToString("F3") + "\n" +

  "SD85Heart," + SD85Heart.ToString("F3") + "\n" +

  "SD80Heart," + SD80Heart.ToString("F3") + "\n" +

  "SD75Heart," + SD75Heart.ToString("F3") + "\n" +

  "SD70Heart," + SD70Heart.ToString("F3") + "\n" +

  "SD65Heart," + SD65Heart.ToString("F3") + "\n" +

  "SD60Heart," + SD60Heart.ToString("F3") + "\n" +

  "SD55Heart," + SD55Heart.ToString("F3") + "\n" +

  "SD50Heart," + SD50Heart.ToString("F3") + "\n" +

  "SD45Heart," + SD45Heart.ToString("F3") + "\n" +

  "SD40Heart," + SD40Heart.ToString("F3") + "\n" +

  "SD35Heart," + SD35Heart.ToString("F3") + "\n" +

  "SD30Heart," + SD30Heart.ToString("F3") + "\n" +

  "SD25Heart," + SD25Heart.ToString("F3") + "\n" +

  "SD20Heart," + SD20Heart.ToString("F3") + "\n" +

  "SD15Heart," + SD15Heart.ToString("F3") + "\n" +

  "SD10Heart," + SD10Heart.ToString("F3") + "\n" +

  "SD5Heart," + SD5Heart.ToString("F3") + "\n" +
```

```
  "SDmaxHeart," + SDmaxHeart.ToString("F3"));


//-------------------------------------------------------------
// EXTRACT Left Lung (matches structure IDs such as "lung_l" or "lung left")
//-------------------------------------------------------------
var lung = structureSet.Structures
  .FirstOrDefault(s => !s.IsEmpty && s.DicomType == "ORGAN" &&
  s.Id.ToLower().Contains("lung left"));


List<double> sdLung = new List<double>();
for (int i5 = 0; i5 < sizeX; i5++)
{
  for (int j5 = 0; j5 < sizeY; j5++)
  {
    for (int k5 = 0; k5 < sizeZ; k5++)
    {
      VVector voxel5 = new VVector(
        origin.x + i5 * resX,
        origin.y + j5 * resY,
        origin.z + k5 * resZ);
      if (lung != null && lung.IsPointInsideSegment(voxel5))
        sdLung.Add(SD[i5, j5, k5]);
    }
  }
}
```

```
//---------------------------------------
// SDVH METRICS (Lung)
//---------------------------------------
sdLung.Sort();
int NLung = sdLung.Count;
double SDmeanLung = sdLung.Average();
double SDmaxLung = sdLung.Max();
double SD95Lung = sdLung[(int)(0.95 * NLung)];
double SD90Lung = sdLung[(int)(0.90 * NLung)];
double SD85Lung = sdLung[(int)(0.85 * NLung)];
double SD80Lung = sdLung[(int)(0.80 * NLung)];
double SD75Lung = sdLung[(int)(0.75 * NLung)];
double SD70Lung = sdLung[(int)(0.70 * NLung)];
double SD65Lung = sdLung[(int)(0.65 * NLung)];
double SD60Lung = sdLung[(int)(0.60 * NLung)];
double SD55Lung = sdLung[(int)(0.55 * NLung)];
double SD50Lung = sdLung[(int)(0.50 * NLung)];
double SD45Lung = sdLung[(int)(0.45 * NLung)];
double SD40Lung = sdLung[(int)(0.40 * NLung)];
double SD35Lung = sdLung[(int)(0.35 * NLung)];
double SD30Lung = sdLung[(int)(0.30 * NLung)];
double SD25Lung = sdLung[(int)(0.25 * NLung)];
double SD20Lung = sdLung[(int)(0.20 * NLung)];
double SD15Lung = sdLung[(int)(0.15 * NLung)];
double SD10Lung = sdLung[(int)(0.10 * NLung)];
double SD5Lung = sdLung[(int)(0.05 * NLung)];
```

```
//---------------------------------------
// EXPORT RESULTS (Lung)
//---------------------------------------

File.WriteAllText(path + "metricsLung.csv",
  "Metric,Value\n" +
  "SDmeanLung," + SDmeanLung.ToString("F3") + "\n" +
  "SD95Lung," + SD95Lung.ToString("F3") + "\n" +
  "SD90Lung," + SD90Lung.ToString("F3") + "\n" +
  "SD85Lung," + SD85Lung.ToString("F3") + "\n" +
  "SD80Lung," + SD80Lung.ToString("F3") + "\n" +
  "SD75Lung," + SD75Lung.ToString("F3") + "\n" +
  "SD70Lung," + SD70Lung.ToString("F3") + "\n" +
  "SD65Lung," + SD65Lung.ToString("F3") + "\n" +
  "SD60Lung," + SD60Lung.ToString("F3") + "\n" +
  "SD55Lung," + SD55Lung.ToString("F3") + "\n" +
  "SD50Lung," + SD50Lung.ToString("F3") + "\n" +
  "SD45Lung," + SD45Lung.ToString("F3") + "\n" +
  "SD40Lung," + SD40Lung.ToString("F3") + "\n" +
  "SD35Lung," + SD35Lung.ToString("F3") + "\n" +
  "SD30Lung," + SD30Lung.ToString("F3") + "\n" +
  "SD25Lung," + SD25Lung.ToString("F3") + "\n" +
  "SD20Lung," + SD20Lung.ToString("F3") + "\n" +
  "SD15Lung," + SD15Lung.ToString("F3") + "\n" +
  "SD10Lung," + SD10Lung.ToString("F3") + "\n" +
```

```
    "SD5Lung," + SD5Lung.ToString("F3") + "\n" +
    "SDmaxLung," + SDmaxLung.ToString("F3"));



//---------------------------------------
// EXPORT ROBUSTNESS MAP
//---------------------------------------
// Write the raw binary voxel data (one double per voxel, X varying fastest)
using (BinaryWriter bw = new BinaryWriter(File.Open(path + "robustness.raw", FileMode.Create)))
{
  for (int k3 = 0; k3 < sizeZ; k3++)
    for (int j3 = 0; j3 < sizeY; j3++)
      for (int i3 = 0; i3 < sizeX; i3++)
        bw.Write(robustness[i3, j3, k3]);
}

// Write a MetaImage header so the .raw file can be opened in 3D Slicer / ITK-SNAP / ParaView
File.WriteAllText(path + "robustness.mhd",
    "ObjectType = Image\n" +
    "NDims = 3\n" +
    "DimSize = " + sizeX + " " + sizeY + " " + sizeZ + "\n" +
    "ElementSpacing = " + resX.ToString("F4") + " " + resY.ToString("F4") + " " + resZ.ToString("F4") + "\n" +
```

```
        "Offset = " + origin.x.ToString("F4") + " " + origin.y.ToString("F4") + " " +
origin.z.ToString("F4") + "\n" +
        "ElementType = MET_DOUBLE\n" +
        "ElementByteOrderMSB = False\n" +
        "ElementDataFile = robustness.raw\n");



    // FINAL SUMMARY POPUP
    //---------------------------------------
    MessageBox.Show(
      "Robustness analysis completed\n\n" +
      "SDmean = " + SDmean.ToString("F3") + " %\n" +
      "SD95 = " + SD95.ToString("F3") + " %\n" +
      "SDmax = " + SDmax.ToString("F3") + " %\n" +
      "PTV voxels = " + N + "\n" +
      "SDmeanHeart = " + SDmeanHeart.ToString("F3") + " %\n" +
      "SD95Heart = " + SD95Heart.ToString("F3") + " %\n" +
      "SDmaxHeart = " + SDmaxHeart.ToString("F3") + " %\n" +
      "Heart voxels = " + NHeart + "\n" +
      "SDmeanLung = " + SDmeanLung.ToString("F3") + " %\n" +
      "SD95Lung = " + SD95Lung.ToString("F3") + " %\n" +
      "SDmaxLung = " + SDmaxLung.ToString("F3") + " %\n" +
      "Lung voxels = " + NLung);
  }
 }
}
```

Attachment 3 - Validation code for SD percentage output

```
//---------------------------------------
    // EXTRACT PTV AND CONTROL VALUES
    //---------------------------------------
    var ptv = structureSet.Structures
      .FirstOrDefault(s => !s.IsEmpty && s.DicomType == "PTV");

    List<double> sdPTV = new List<double>();
    double maxiPTV = 0;   // highest SD value found inside the PTV
    int imax = 0;         // voxel indices (i, j, k) where that max SD occurs
    int jmax = 0;
    int kmax = 0;

    for (int i2 = 0; i2 < sizeX; i2++)
    {
      for (int j2 = 0; j2 < sizeY; j2++)
      {
        for (int k2 = 0; k2 < sizeZ; k2++)
        {
          VVector voxel2 = new VVector(
            origin.x + i2 * resX,
            origin.y + j2 * resY,
            origin.z + k2 * resZ);
```

```
            if (ptv != null && ptv.IsPointInsideSegment(voxel2))
              sdPTV.Add(SD[i2, j2, k2]);

            // Track the voxel with the maximum SD value within the PTV
            if (SD[i2, j2, k2] > maxiPTV && ptv.IsPointInsideSegment(voxel2)) maxiPTV = SD[i2, j2, k2];
            if (maxiPTV == SD[i2, j2, k2] && ptv.IsPointInsideSegment(voxel2)) imax = i2;
            if (maxiPTV == SD[i2, j2, k2] && ptv.IsPointInsideSegment(voxel2)) jmax = j2;
            if (maxiPTV == SD[i2, j2, k2] && ptv.IsPointInsideSegment(voxel2)) kmax = k2;
          }
        }
      }

      //-----------------------------------
      // ITERATIVE PRINT OF VOXEL DOSE
      //-----------------------------------
      // For the voxel with the highest SD, show the dose delivered by each of the 8 scenarios
      for (int splano = 0; splano < 8; splano++)
      {
        double doseteste = doseMatrix[splano, imax, jmax, kmax];
        MessageBox.Show(
          "Voxel dose at max SD location, scenario " + splano.ToString("F3") + " = " + doseteste);
      }
```

Attachment 4 - Patient-specific complexity metrics and robustness indices

Patient_number	Complexity		Robustness_Index
	Modulation_Factor (MU/cGy)	Edge_Metric (mm^{-1})	
1	3,020	0,077	0,435
2	3,497	0,086	0,620
3	5,155	0,144	0,091
4	3,171	0,082	0,512
5	2,990	0,087	0,428
6	3,002	0,081	0,555
7	2,681	0,087	0,673
8	2,936	0,069	0,725
9	3,117	0,063	0,569
10	3,592	0,090	0,783
11	3,040	0,073	0,654
12	3,239	0,065	0,628
13	2,698	0,075	0,566
14	4,140	0,083	0,656
15	3,414	0,085	0,913
16	4,076	0,080	0,739
17	2,824	0,066	0,509
18	3,857	0,085	0,588
19	4,478	0,076	0,914
20	3,716	0,093	0,866
21	3,541	0,072	0,571
22	3,225	0,075	0,842
23	3,805	0,084	0,590
24	3,643	0,119	0,504
25	4,501	0,115	0,655
26	3,476	0,088	0,219
27	2,922	0,071	0,566
28	4,469	0,093	0,500
29	2,802	0,071	0,614
30	2,816	0,077	0,697
31	3,091	0,058	0,747
32	3,151	0,174	0,608
33	3,846	0,073	0,775

34	3,461	0,088	0,698
35	2,969	0,064	0,680
36	2,611	0,065	0,620
37	3,391	0,062	0,700
38	3,582	0,089	0,889
39	3,291	0,076	0,774
40	4,717	0,101	0,553
41	4,696	0,143	0,630