Evaluation of absorbed dose uncertainty in modulated radiotherapy plans

Julia Götstedt

Department of Radiation Physics
Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2022

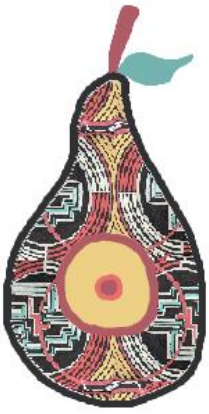
Cover illustration by Jeanette Segerström

Evaluation of absorbed dose uncertainty in modulated radiotherapy plans

julia.gotstedt@vgregion.se

ISBN 978-91-8009-941-7 (PRINT)
ISBN 978-91-8009-942-4 (PDF)

Printed in Borås, Sweden 2022
Printed by Stema Specialtryck AB



The pear of knowledge
has got a new leaf.

Evaluation of absorbed dose uncertainty in modulated radiotherapy plans

Julia Götstedt
Department of Radiation Physics, Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ABSTRACT

The purpose of this work was to develop and evaluate methods to meet the challenges of quality control (QC) for modulated radiotherapy plans. It was shown that nine of 15 intensity modulated radiotherapy (IMRT) plans, with deliberately introduced dose errors larger than 5% in at least one evaluated dose volume histogram (DVH) metric, were not detected with a QC method which combined Delta^4 (ScandiDos) measurements and internationally recommended criteria for evaluation (**Paper I**). The dose difference between calculation and high spatial resolution measurements, using EBT3 film and electronic portal imaging device (EPID), for 30 static beam apertures of varying size and shape was used as a measure of beam aperture complexity (**Paper II**). The linear correlation to the beam aperture complexity was evaluated for three aperture-based complexity metrics developed in this study and five other metrics suggested in the literature. The strongest correlation, with a Pearson's r-value of -0.94, was found for the developed edge area metric (EAM). EAM was further evaluated for 18 static beam openings originating from control points of clinically used volumetric modulated arc therapy (VMAT) plans and for 200 full VMAT plans planned for different treatment sites (**Paper III**). The results indicated that the EAM must be interpreted differently for different diagnoses. Evaluation of beam aperture shape, modulation variations, measurements, and delivery simulations, as methods for assessment of the dosimetric uncertainty for VMAT plans, showed that the dosimetric uncertainty could differ even though the plans appeared to be equal based solely on dosimetric comparisons of the dose distributions, e.g., DVH metric evaluations (**Paper IV**).

In conclusion, it is possible to decrease dosimetric uncertainties in modulated radiotherapy plans to enable a higher treatment quality. The dosimetric uncertainties can be assessed by different methods but it is important to define the purpose of the method, and to validate that the method fulfils the defined purpose.

Keywords: Modulated radiotherapy, IMRT, VMAT, dosimetric uncertainty, quality control, quality assurance, complexity metric, edge area metric

ISBN 978-91-8009-941-7 (PRINT)
ISBN 978-91-8009-942-4 (PDF)

SAMMANFATTNING PÅ SVENSKA

Ungefär hälften av alla cancerpatienter i Sverige får strålbehandling, som enda behandlingsform eller i kombination med andra behandlingsformer. För varje enskild patient utformas strålbehandlingen individuellt med hjälp av ett dosplaneringssystem. Hur strålning ges anpassas för att skapa en optimal behandling, dvs val av form, vinklar och energier för strålfälten, baserat på hur stråldosen beräknas att fördelas i en tredimensionell röntgenbild av patienten. Denna beräknade dosfördelning bedöms utifrån kriterier att nå tillräckligt hög stråldos i behandlingsområdet för att nå tumörkontroll men samtidigt så låg dos som möjligt i kringliggande frisk vävnad för att minska risken för biverkningar. Strålbehandlingen ges oftast med en linjäraccelerator. Behandlingen kan ske med statiska strålfält vilket innebär ett strålfält med en bestämd form från en bestämd infallsvinkel, eller dynamiska strålfält då både strålfältets form och/eller infallsvinkel kan variera under bestrålningen. Dynamisk strålbehandling innebär att behandlingsplanen med större möjlighet kan anpassas för att skapa en mer fördelaktig dosfördelning med avseende på de kliniska målen för en god behandling. Samtidigt innebär den dynamiska behandlingstekniken fler parametrar som kan leda till en ökad osäkerhet i beräkning av stråldosfördelningen och också mekaniska osäkerheter vid leveransen av behandlingen.

Studier inom denna avhandling visar att vanligt förekommande metoder för att uppskatta osäkerhetsnivån hos dynamiska behandlingsplaner inte uppfyller de uppsatta kraven i alla situationer (kanske finns ett bättre ord). Vidare påvisas att det finns osäkerheter vid beräkning, leverans och mätning av dosfördelning för dynamiska strålbehandlingsplaner och storleken för dessa olika typer av osäkerheter kan påverkas av behandlingsplanens utformning. Avhandlingen presenterar också framtagande av olika typer av metoder för att uppskatta osäkerhetsnivån i den dosfördelning som faktiskt ges till patienten. Fortsatt arbete med att implementera dessa metoder i det kliniska arbetet skulle innebära möjligheten att inkludera en uppskattning av osäkerhetsnivån i den totala bedömningen av kvalitén av dynamiska strålbehandlingar, utöver bedömningen av de kliniska målen för den beräknade dosfördelningen, vilket skulle kunna påverka valet av plan som används för behandling.

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. **IMRT patient-specific QA using Delta^4 dosimetry system and evaluation based on ICRU 83 recommendations**
Julia Nilsson, Anna Karlsson Hauer A and Anna Bäck
J. Phys.: Conf. Ser. **2013**;444:012048
https://doi.org/10.1088/1742-6596/444/1/012048

II. **Development and evaluation of aperture-based complexity metrics using film and EPID measurements of static MLC openings**
Julia Götstedt, Anna Karlsson Hauer and Anna Bäck
Med. Phys. **2015**;42(7):3911-3921.
http://dx.doi.org/10.1118/1.4921733

III. **Edge area metric complexity scoring of volumetric modulated arc therapy plans**
Julia Götstedt and Anna Bäck
Phys Imaging Radiat Oncol. **2021**;17:124-129
https://doi.org/10.1016/j.phro.2021.02.002

IV. **Evaluation of methods for dosimetric uncertainty assessment of VMAT plans**
Julia Götstedt, Anna Karlsson and Anna Bäck
Manuscript **2022**

RELATED PAPERS AND CONFERENCE CONTRIBUTIONS

Quality index to predict differences between planned and measured dose due to limitations in TPS
J Götstedt, A Karlsson Hauer, and A Bäck
EP-1512. ESTRO33 Vienna, Austria **(2014)**
Radiotherapy and Oncology 111(1):S166-S167
https://doi.org/10.1016/S0167-8140(15)31630-3

Complexity metric as a complement to measurement based IMRT/VMAT patient-specific QA
J Götstedt, A Karlsson Hauer and A Bäck
J. Phys.: Conf. Ser. 573 012016 **(2015)**
https://doi.org/10.1088/1742-6596/573/1/012016

Impact of different dose calculation algorithms on aperture-based complexity metric evaluations
A Bäck, A Larsson, J Götstedt, and A Karlsson Hauer
EP-1595. ESTRO35 Turin, Italy **(2016)**
Radiotherapy and Oncology 119:S741-S742
https://doi.org/10.1016/S0167-8140(16)32845-6

Complexity metric based on fraction of penumbra dose – initial study
A Bäck, F Nordström, M Gustafsson, J Götstedt and A Karlsson Hauer
J. Phys.: Conf. Ser. 847 012002 **(2017)**
https://doi.org/10.1088/1742-6596/847/1/012002

Introducing the fraction of penumbra dose in the evaluation of VMAT treatment plans
A Bäck, F Nordström, M Gustafsson, J Götstedt, and Anna Karlsson Hauer
EP- 1457. ESTRO36 Vienna, Austria **(2017)**
https://www.postersessiononline.eu/173580348_eu/congresos/ESTRO36/aula/-EP_1457_ESTRO36.pdf

Further developments of two complexity metrics to consider clinical aspects of VMAT treatment plans
J Götstedt, A Bäck, and A Karlsson Hauer
EP-1486. ESTRO36 Vienna, Austria **(2017)**
Radiotherapy and Oncology 123:S795-S796
https://doi.org/10.1016/S0167-8140(17)31921-7

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

Different aspects of plan complexity in prostate VMAT plans
E Terzidis, F Nordström, J Götstedt, and A Bäck
12th ICDDose Quebec City, Canada **(2022)**
J. Phys.: Conf. Ser. (to be published)

CONTENT

ABBREVIATIONS

AAA	Anisotropic Analytical Algorithm
AA	Aperture Area
AI	Aperture Irregularity
ASC	Aperture Shape Controller
CAM	Converted Aperture Metric
CC	Collapsed Cone
CRT	Conventional Radiation Therapy
CT	Computed Tomography
CU	Calibration Units
$D_{2\%}$	Near-maximum dose
DICOM	Digital Imaging and Communications in Medicine
DTA	Distance To Agreement
DVH	Dose Volume Histogram
EAM	Edge Area Metric
EC	Eclipse
EM	Edge Metric
EPID	Electronic Portal Imaging Device
ESAPI	Eclipse Scripting Application Programming Interface
H&N	Head and Neck
ICRU	International Commission on Radiation Units and Measurements
IMRT	Intensity Modulated Radiation Therapy
LCPE	Lateral Charged Particle Equilibrium

MCO	Multi Criteria Optimization
MCS	Modulation Complexity Score
MI	Modulation Index
MLC	Multi Leaf Collimator
MU	Monitor Units
OAR	Organs At Risk
OPP	Optimized Phantom Position
PTV	Planning target volume
QA	Quality Assurance
QC	Quality Control
RIT	Radiological Imaging Technology
ROI	Region Of Interest
RS	RayStation
SD	Standard Deviation
SSD	Source Surface Distance
SSM	Strålsäkerhetsmyndigheten
TPS	Treatment Planning System
VMAT	Volumetric Modulated Arc Therapy
2D	Two Dimensional
3D	Three Dimensional

1 INTRODUCTION

1.1.1 EXTERNAL PHOTON RADIOTHERAPY TODAY

Today, radiotherapy is widely used as a method for cancer treatments. It can be given as a single method of treatment but it is often combined with other treatment techniques, for example surgery or chemotherapy[1]. About half of all patients diagnosed with cancer will be given radiotherapy at some point during their treatment course. It can be either prior to surgery, to shrink the tumor, or after surgery to irradiate any remaining tumor cells. Radiotherapy can be used with curative intent which aims to eliminate the tumor or as palliative treatment, for example to reduce pain. Radiotherapy can be seen as a two-edged sword with the chance of treatment on one side, and the risk of radiation-induced second malignancies in the future on the other.

Modern external radiotherapy is given with high precision and high energies and an error during treatment can lead to severe consequences. If the absorbed dose to the treatment volume is too low, the treatment will not be effective, and excessive absorbed doses given to healthy critical organs, can lead to severe side effects. It is therefore important to assure the quality and safety of radiotherapy prior and during treatment. This is statutory in Sweden by the Swedish Radiation Safety Authority, as it is stated that the radiotherapy departments are obliged to verify that the planned absorbed dose is in agreement with the delivered absorbed dose for all radiotherapy plans to be approved for treatment[2].

1.1.2 SHORT SUMMARY OF RADIOTHERAPY HISTORY

The modern knowledge of radiation sciences, and the radiation therapy of today, started more than a century ago with the discovery of x-rays 1895 by the German physicist Wilhelm Röntgen and the discovery of radioactivity 1896 by the French physicist Henri Becquerel.

The application of radiation within medical health care was immediately recognized when images of a skeleton appeared as shadows on photographic plates generated by Röntgen's experiments. The first medical x-ray laboratory for medical use was established within months after the discovery and the first diagnostic examinations using x-rays performed in Sweden followed just one year after. Attempts to treat cancer using x-rays around the world were also immediate, for example treatment of breast cancer in Chicago, stomach cancer in Lyon, and a skin tumor in Vienna 1896. The first successful radiotherapy treatment using x-rays in Sweden was performed in 1899 on a woman with skin cancer on the tip of her nose.

Several decades later, in the 1950s, cobalt-60 and cesium-137 were introduced as γ-photon radiation sources to be used for external radiotherapy. This enabled treatments with higher energies which could treat tumors at a greater depth and at the same time spare the skin from high absorbed radiation dose. The pace of development has been increasing ever since and the previously used equipment has been replaced with modern linear accelerators to deliver external radiation therapy. Development of both hardware and software have led to advanced and more precise treatments.

X-rays are still essential for diagnostic images and used in radiotherapy, as computed tomography (CT) produces a three-dimensional (3D) x-ray image which is used for delineation of the treatment volume and organs at risk (OAR) of receiving critical absorbed radiation doses.

Today, an individual treatment plan is optimized to generate an absorbed dose distribution within the patient geometry based on the 3D CT images to achieve a prescribed absorbed dose to the target, i.e., treatment volume including the tumor with a margin, while at the same time delivering as low absorbed dose as possible to the healthy surrounding tissue.

1.2 MODULATED RADIOTHERAPY

1.2.1 BACKGROUND TO MODULATED TREATMENT TECHNIQUES

The procedure of planning treatments, which includes calculation of absorbed dose distributions, and the delivery technique of photon radiotherapy went through great progress during the 1980s. A breakthrough of modulated radiotherapy came with the publication by Brahme et al. in 1988[3], which describes and solves the so-called reverse problem. The optimal treatment beams, built up by several different beam aperture shapes and incident beam angles, is determined by inverse back projection of the desired dose distribution. The beam was previously modified using physical compensators and wedges. Today, the beam apertures are shaped by multileaf collimators (MLC), which is a set of multiple symmetrical leaf pairs of high-density material shielding blocks, with a width of a few millimetres up to one centimetre. The development of a more advanced treatment planning, along with the development of more advanced hardware, have led to the introduction of so-called intensity modulated radiation therapy (IMRT)[3,4] with the possibility of delivering radiotherapy with a non-uniform fluence.

The most considerable difference between 3D conventional radiation therapy (CRT) and the conformal IMRT technique is the possibility to form a high absorbed dose distribution very precisely in an arbitrary geometry, such as an irregularly shaped target, and at the same time achieve a homogenous dose distribution in the target volume. IMRT is delivered from several, in most cases 5-9, static gantry angles. For the IMRT with step-and-shoot delivery technique the beam consists of several sub-beams of static MLC beam apertures delivered one at a time, and for the IMRT with sliding window delivery technique the beam aperture is dynamically shaped by the MLC during delivery[5,6].

The modulated radiotherapy technique has been further developed into treatment delivery with simultaneously rotating gantry, continuously moving collimators (both MLC and jaws) and varying dose rate. This technique, referred to as volumetric modulated arc therapy (VMAT) was proposed in 1995[7] but has been further developed before clinical implementation[8]. The advantage of a rotating gantry is to allow for beam entries from all possible incidence angles around the patient. The VMAT technique led to a more even distribution of low doses and enabled a more time efficient treatment delivery which led to higher patient comfort and a lower risk of patient motion during treatment delivery.

1.2.2 MODULATED TREATMENT PLANNING

A modulated treatment plan is created in a treatment planning system (TPS) using an optimization algorithm. The algorithm aims to find the optimal settings for the plan parameters, i.e., the gantry position, gantry speed, collimator position, collimator angle, MLC positions and dose rate, to achieve the desired dose distribution. The optimization goals for the desired dose distribution are defined by the treatment planner as dose optimization objectives or constraints, the latter used for critical structures where the dose limit is non-negotiable, used as input to the algorithm. These objectives and constraints are expressed as dose volume histogram (DVH) metrics, i.e., a specified volume receiving a specified dose. As stated earlier, the optimization algorithm within the TPS finds the ultimate solution to generate a modulated treatment plan that fulfils the user defined dose optimization objectives according to the reverse problem. The plan parameters described previously are modulated within the optimization process, but the collimator angle is adjusted manually, and the collimator position is automatically adjusted to the MLC opening as a function called jaw tracking. The plan parameter information is defined in discrete control points evenly spaced for the treatment arc, about every other degree of gantry angle, which is sent to the treatment machine for treatment delivery.

The optimization module includes several tools for the user to steer the process of how the dose will be distributed. Most optimization modules include an option to limit the output of the treatment machine, expressed in monitor units (MU). This limitation affects the beam aperture shape, as a more open aperture is needed to be able to deliver the same dose with a lower number of MU. Another tool that can be used to affect the shape of the aperture is the aperture shape controller (ASC), provided as a tool in the optimization algorithm by Varian Medical Systems Inc. (Palo Alto, USA). The ASC algorithm penalizes irregularly shaped apertures by minimizing the distance between adjacent MLC leaf tips[9-11].

Multi-criteria optimization (MCO) is an optimization method available in TPSs from different vendors[9,12-14]. With MCO, a set of pareto optimal treatment plans is generated, i.e., dose distributions that are optimal in the sense that a better dose-volume objective for one structure can only be fulfilled by worsening it for another. MCO allows the user to navigate between the different options to choose the most appropriate treatment plan for the patient.

Given all the different planning strategies and tools available for treatment planning, it is possible to create treatment plans of different designs that generate similar dose distributions.

1.3 QUALITY OF MODULATED RADIOTHERAPY

The design of the treatment plan, i.e., the combinations and values for specific plan parameters, can be more or less advantageous with respect to the dosimetric accuracy, i.e., the agreement between planned and delivered absorbed dose distributions. Dosimetric uncertainties, i.e., the differences between planned and delivered dose distribution, can be caused by e.g., limitations in the dose calculation or variations of machine parameters during delivery. The dosimetric uncertainties can be evaluated by different methods. A common nomenclature within the topic of evaluation and reasons for dosimetric uncertainties is not yet established; different definitions of plan quality and plan complexity can be found in different studies[15-19]. Different terms and explanations of how they are applied in this thesis are defined in Table 1 and further described below.

Table 1. The nomenclature used within this thesis and included papers.

Nomenclature	
Plan quality:	A treatment plan with a high plan quality fulfils the individual clinical dosimetric treatment goals and is designed to be deliverable with minimized uncertainty, i.e., as small differences as possible between the planned and delivered absorbed dose distributions.
Plan complexity:	A measure of the impact of treatment plan design, i.e., the combination of plan parameters, in terms of differences between planned and delivered absorbed dose distribution.
Plan robustness:	A measure of the impact of variations of machine parameters at the occasion for irradiation in terms of the effect on the delivered absorbed dose distribution.
Dosimetric uncertainty:	Dose differences between planned and delivered absorbed dose distribution.

The design of a treatment plan, i.e., a certain combination of plan parameters, could lead to smaller or larger dosimetric differences, i.e., the difference between the planned and the delivered absorbed dose distribution, due to reasons related to plan complexity. Differences between planned and delivered dose distribution will occur to a higher extent for fields of a smaller size and/or irregularly shaped beam apertures, as shown in Paper II and described in the following section *Dosimetric challenges of small and irregular fields.* An example of a complex treatment plan would be a plan that is designed to deliver the treatment with fields that consist of smaller and/or more irregular beam apertures which could lead to larger dosimetric uncertainties.

The concept of robustness is commonly used to address the dosimetric differences because of patient geometry variations, which has a larger impact when it comes to proton therapy compared to photon therapy. Any changes in patient position or geometry compared to the geometry for which the treatment was planned, e.g., a spatial mismatch of the patient position or an internal anatomical change, could lead to differences between planned and delivered dose distribution. This type of treatment uncertainty is, however, not included in the studies of uncertainties in this thesis.

Within this thesis, the term *plan robustness* refers to the susceptibility of a treatment plan to variations of machine parameters, e.g., the positions of MLC leaves, collimator jaws, and the gantry, on the delivered absorbed dose. For example, a robust treatment plan would imply smaller dose differences between planned and delivered dose distribution compared to a non-robust treatment plan. A non-robust treatment plan is therefore also considered as a complex treatment plan.

1.3.1 DOSIMETRIC CHALLENGES OF SMALL AND IRREGULAR BEAM APERTURES

Modulated treatment techniques deliver a conform and homogenous dose distribution by treatment delivery of several beam apertures that are often of irregular shapes or of small sizes. Dosimetric accuracy for small beam apertures is challenging because of limitations in dose calculation accuracy and susceptibility to delivery variations. Dosimetry of small beam apertures is also challenging because of measurement uncertainties specific for situations with small beam apertures.

The reasons for dosimetric uncertainties related to dose calculation for small beam apertures have been described by Das et al.[20] and are summarized here. The model-based dose calculation algorithms, which are commonly used in TPSs today, can generally be divided in two steps. The first step models the fluence that comes out of the treatment head. The second step is to combine this fluence and the patient geometry to calculate the absorbed dose distribution. A correct model of the MLC, e.g., leakage, shape of leaf ends, and leaf position, is important for an accurate dose calculation and of special concern when calculating the dose for small fields[21,22]. Further, it is important to perform a correct modelling of the direct beam source size. The size of the direct source from a specific point of view affects the width of the penumbra, as a larger source size leads to a wider penumbra. If the field size is small, and the primary source is partially blocked, the penumbra will be wider in comparison to the smaller field size and will overlap with the penumbra from other nearby field edges. This will reduce the actual contribution of primary fluence and the dose output in the center of the field, and as a result the dose will be overestimated in the TPS.

A lack of lateral charged particle equilibrium (LCPE) will further reduce the absorbed dose. LCPE is the state for a volume where the amount of charged particles, i.e., secondary electrons, that laterally exit and enter the volume, is equal. A field can be considered small when the distance between the center of the field and the field edge is about the same size as the lateral range for the secondary electrons. The scattering and attenuation of the secondary electrons is challenging for the dose calculation algorithm to simulate correctly in a time efficient way. The lateral range of the secondary electrons increases with increasing energy of the incident photons and what is considered as a small field can therefore vary.

It is challenging to measure the dose correctly in small field conditions since the detector dimension will be large in comparison to the field size. A large detector in relation to the small field could lead to a volume averaging effect and the detector will measure a lower dose than actual in the field center[23]. Also, density differences between the detector and the media will lead to that normally used perturbation factors will not be applicable[20]. Commonly used TPSs are therefore often recommended to be commissioned for field sizes from 3 cm x 3 cm and larger; fields smaller than this are generally considered as small[24].

Another aspect of dosimetric challenges of small fields is that small fields are more susceptible to variations of machine parameters during delivery. The differences between planned and actual positions of mechanical parts during delivery, for example the position of MLC leaves, will have a larger effect for smaller fields as the dosimetric differences caused by these variations will constitute a larger part of the total dose distribution[21,22].

For static 3DCRT fields, a positioning error of MLC leaves generally affects the dose distribution in the peripheral region of the treatment volume. For modulated treatments which are planned to deliver the dose with fields consisting of several small beam apertures, such a MLC leaf positioning error could cause dosimetric differences within the treatment volume.

1.3.2 ADDITIONAL UNCERTAINTIES

Different parameters regarding the dynamic delivery of modulated treatment techniques, such as the variation in speed and acceleration of the MLC leaves, the gantry rotation and the dose rate have been investigated as sources of uncertainty for treatment delivery[25]. However, differences between planned and delivered absorbed dose distributions could also be linked to uncertainties of the treatment technique, the patient position, or a combination of them. The interaction between the treatment technique with dynamic beam delivery of VMAT and the continuous motion of the target volume, e.g., in cases of lung cancer treatments, is referred to as interplay effect[26]. The interplay effect can make it difficult to deliver the treatment as planned due to moving MLC leaves and moving target. This could lead to an over- or underdosage of the target and a higher dose to the healthy surrounding tissue[26].

Inter- or intra fraction variation in patient geometry is also recognized as a cause for dosimetric uncertainty. Dosimetric uncertainty related to uncertainties of position and/or motion of the patient is not investigated further within this thesis; it is only mentioned here to be recognized as an additional source of dosimetric uncertainty.

1.3.3 QUALITY CONTROL

Radiotherapy departments have a statutory responsibility to assure a safe treatment of the patient. This is conducted through local routines of quality assurance (QA) that comply with the regulation directives of national authorities. The QA routine involves quality controls (QC) that are specific for each step in the radiotherapy treatment chain, from the first step of treatment preparation to the very last fraction of treatment delivery.

One control within the QA is to ensure the agreement between the delivered absorbed dose distribution and the planned absorbed dose distribution. Previous regulations from the Swedish Radiation Safety Authority (SSM) stated the requirements for performing measurement-based QC for this purpose. For modulated treatment techniques, e.g., IMRT and VMAT for which in vivo dosimetry by using diode detectors is not applicable, it has earlier been mandatory to perform patient-specific pre-treatment measurements for treatment delivery verification due to the regulation by SSM. The national regulation have since been updated and the requirements are now stated "… *verify the agreement between planned and delivered dose to the patient" (SSMFS 2018:5, kap 8, §11)* without stating the choice of method[2]. This enables the use of other methods, that are not measurement-based, to perform this type of control of patient-specific treatment plans.

Several working groups, national and international, have evaluated and recommended different QC methods specific for modulated treatment techniques which are measurement-based[6,27-35] or use analytical QC methods [17,36-45]. The International Commission on Radiation Units and Measurements (ICRU) report no 83 included discussions of different methods for measurement-based QC and recommend that the dosimetric deviations detected during QC should be evaluated with respect to the effect on the dose distribution and clinical relevance[30]. The same report recommended to consider the creation of a new treatment plan, possibly with a non-modulated treatment technique, if the reason for eventual dose differences of clinical relevance cannot be explained. The performance of the evaluation method for measurement-based QC recommended by ICRU 83 is investigated in Paper I.

In a report from a national working group two possible QC tracks was suggested : a measurement-based track and a process-oriented track including the use of analytical QC methods[46]. The intention with the process-oriented track is described by the national working group as a way to save resources and be able to put more effort on the evaluation of treatment plans that are non-standard. Different analytical methods have been suggested to distinguish

which treatment plans that qualify as either standard or non-standard, but these methods need further validation before clinical implementation.

Regardless the choice of QC method, it is of great importance that each department verifies that their choice of QC method, is able to detect dose errors of clinical relevance or the type of errors that the method is intended to detect[33]. This could be done e.g., by creating treatment plans with known dosimetric errors that should be found by the applied QC method.

1.3.4 MEASUREMENT-BASED METHODS FOR QUALITY CONTROL

The method used for measurement-based patient-specific QC depends on the intention with the QC, and measurements can either be done in one point, 2D or 3D. There are several commercially available dosimetry systems that include a measurement device, e.g., different types of phantoms inserted with small detectors, and an associated evaluation software for comparison of planned and measured dose* distribution[33].

Measurement-based QC can be resource-consuming, which is one of the major drawbacks of the method. For example, film or gel dosimetry needs a lot of effort both before and after the measurement. But in return, it offers measurements of very high spatial resolution. Measurement devices with inserted detectors, e.g., ion chambers or diodes, produce measurement results with a spatial resolution of typically 0.5-1 cm. Measurements with these types of quasi-3D phantoms can give a direct result, compared to film and gel dosimetry which need different types of processing before a measurement result can be evaluated. Quasi-3D indicates that the measurement is only performed in specific spatially spaced measurement points and thereby it does not qualify as full 3D. There is software available that uses interpolation algorithms to generate a full 3D dose distribution based on the measurements in the measurement points. However, this will introduce calculation uncertainties to the already present uncertainties related to the measurement. Some treatment machines offer the possibility to perform measurements in 2D by using an integrated imaging modality, an electronic portal imaging device (EPID), that is mounted on the treatment machine. There has been some recent development of a method for in vivo measurements using EPID[47,48]. In this method, a suggested algorithm would generate back projected 3D dose distributions of the measurements, for evaluation of differences between planned and delivered dose distribution within the patient geometry. This method would include daily variations of the treatment machine if the in vivo EPID measurements were to be applied for all fractions.

* N.B. that the phrase “measured dose” is not correct. The phenomena that are measured is a signal, for example a current or an opacity difference depending on the type of detector, which is converted to a result in terms of absorbed dose. Yet, “measured dose” is sometimes used for simplicity in this thesis.

However, for the most commonly used measurement-based methods, the comparison of measured and calculated dose is generally done in the measurement geometry and not in the relevant patient geometry. Eventual errors will only be captured for that specific occasion of measurement, which can be of different magnitude for different occasions due to machine variations.

The intention behind measurement-based QC could be to detect errors during treatment plan delivery, for example positioning errors of the MLC leaves or other mechanical issues. It could also be a control of a correct export and import of treatment plan files between the planning system and treatment machine. Or, it could be seen as an end-to-end test of the individual patient treatment, prior to the actual treatment. It is difficult to pinpoint the actual reason for an eventual dosimetric difference when using measurement-based QC, since uncertainties are present in the dose calculation, the delivery and the measurement[49,50]. Further, eventual dosimetric differences between delivered and planned dose distributions will be detected very late in the treatment preparation process.

1.3.5 ANALYTICAL METHODS FOR QUALITY CONTROL

Dosimetric uncertainty due to treatment plan design can be estimated analytically using e.g., complexity metrics or treatment delivery simulations. These types of methods have been suggested as direct methods for complexity assessment[38,39]. They have the potential to be more time efficient compared to measurement-based QC. A suggested intention behind analytical QC is to distinguish complex and/or non-robust treatment plans with a type of complexity score that is related to the dosimetric uncertainty. Complexity scores do not result in a measure of actual dose difference as measurements do.

The use of analytical QC methods leads to higher requirements for controls of the treatment machine, as treatment plan delivery will not be controlled on a regular basis, as is the case for patient-specific measurements[46]. The delivery uncertainty is, however, theoretically included as most complexity metrics are developed to include parameters that contribute to delivery uncertainty. The usage of delivery simulations is also a method to assess the theoretical delivery uncertainty without introducing additional uncertainty related to the procedure of measurement.

A time efficient analytical QC method can be important for the implementation of online adaptive radiotherapy, where measurement-based verifications are not suitable[51].

1.3.6 COMPLEXITY METRICS

The first generation of complexity metrics, called fluence-based complexity metrics, was derived for the IMRT technique and based on the beam fluence distribution. The concept of a fluence distribution is not valid for VMAT plans because of the varying beam incidence angle due to the rotating gantry. The derived fluence-based complexity metrics were, therefore, not applicable for VMAT plans, but some of the suggested metrics could be adjusted to be used for both IMRT and VMAT plans. Complexity metrics were further developed and instead based on the plan parameters, i.e., the settings of gantry, collimator jaws, MLC leaves, and dose rate, and divided into two categories as suggested by Crowe et al.[52], called deliverability metrics and accuracy metrics. Accuracy metrics have been developed with focus on parameters that are related to uncertainties of dose calculation, e.g., the dosimetric challenges of small treatment fields. This type of metric is often aperture-based, i.e., based on the shape and size of the MLC openings. The other category, called deliverability metrics, consists of metrics that are based on treatment parameters related to the dynamic delivery e.g., modulation indices[45].

Complexity metrics have been developed and suggested for different intentions. Complexity metrics can either be used in the early stage of treatment planning, as an action guide for the treatment planner. Or, as a part within the optimization process to lower the treatment plan complexity[37]. It can also be used after treatment planning to determine if the treatment plan should be considered as a standard or non-standard plan, based on comparison to a group of earlier created, controlled, and approved treatment plans, as a guidance for following actions of QC.

Studies on the performance of complexity metrics have, to my knowledge, not yet been thoroughly validated. Most of the validation studies are based on the correlation of calculated complexity scores and results of patient-specific measurement-based QC. However, results of previous studies have shown that this approach is often insufficient to detect relevant dosimetric differences (Paper I). This means that the connections between dosimetric clinical relevance and complexity scores are not yet clarified. Validation needs to be done in a systematic approach with measurements dedicated for the purpose. It is also possible that validation is necessary for each combination of diagnose group, treatment technique, treatment machine and treatment planning system.

The main advantage of complexity metrics, besides its time efficiency, is that this type of evaluation will yield additional information about treatment plan complexity, and for some metrics this information will be available on a control point level.

1.3.7 DELIVERY SIMULATIONS

A difference between the planned and the actual position for mechanical parameters, e.g., the MLC leaves, gantry, couch, and collimator, can be related to uncertainties in the calibration of these parameters. Differences can also occur during delivery because the mechanics within the treatment machine are allowed to operate within specified tolerances, which can entail daily variations. The effect on the dose distribution due to these differences between planned and actual parameters value can be estimated by simulations. The advantage of simulation is that no actual measurements are performed, and uncertainties related to the measurement procedure can be avoided. For the simulations, the values of the parameters can be manipulated to simulate either a realistic variation or to simulate a worst-case scenario (i.e., a worst-case scenario within realistic terms). However, it is difficult to in advance predict which changes of parameter values would lead to a worst-case scenario with respect to dosimetric clinical relevance. The planned calculated dose distribution and the calculated dose distribution for the simulated delivery are compared to evaluate the dosimetric effect of the simulated variations. The dosimetric effects can be evaluated in the patient geometry for the CT images and the clinical relevance of eventual dose differences can be evaluated directly. However, simulated deliveries do not include assessment of dosimetric uncertainties related to limitations in dose calculation of the planned dose distribution.

2 AIM

The primary purpose of this project, presented in this thesis, is to develop and evaluate methods to meet the challenges of quality control of modulated radiotherapy plans. With the long-term intention to enable treatments of higher quality by minimizing the dosimetric uncertainty, the studies aim to increase the knowledge of what information that is achievable when using different methods to assess dosimetric uncertainty.

To achieve this, the following sub-aims have been formulated:

Paper I: Examine the performance of a commonly used measurement-based QC method in combination with international recommendations for measurement evaluation criteria.

Paper II: Develop and compare measures for aperture-shape complexity of static fields.

Paper III: Evaluate the use of a new developed aperture-based complexity-metric for clinical VMAT plans.

Paper IV: Evaluate different methods for dosimetric uncertainty assessment of VMAT plans.

3 MATERIAL AND METHODS

3.1 STUDY MATERIAL

The study material of this thesis included both static and dynamic radiation fields. Some radiation fields were created for the purpose of the study and other were selected from treatment plans that had been accepted and used for treatments at Sahlgrenska University Hospital. The study material included in this thesis is listed in Table 2.

Table 2. Detailed list of study material included in each Paper I-IV. The study material involves static fields and modulated treatment plans planned for different techniques with different treatment planning systems for different treatment machines mounted with different multileaf collimators.

Paper	Study material	Details
Paper I	3 IMRT plans, accepted and used for treatment	Eclipse and AAA, version 10 Calculation grid size 2.5 mm Clinac iX and Millennium 120 MLC
Paper II	30 static fields, created in series with varying complex shapes	Eclipse and AAA, version 11 Calculation grid size 2.5 mm Clinac iX and Millennium 120 MLC
Paper III	18 static fields, selected control point beam openings 200 VMAT plans, accepted and used for treatment	Eclipse and AAA, version 13 Calculation grid size 2.5 mm Clinac iX and Millennium 120 MLC TrueBeam STx and High Definition 120 MLC
Paper IV	30 VMAT plans, created using different optimization methods	Eclipse and AAA, version 15 RayStation 8B and CC version 5 Calculation grid size 2.5 mm TrueBeam and Millennium 120 MLC

Note: High Definition and Millennium are multileaf collimators (MLC) with 120 leaves with a leaf width of 2.5 and 5 mm for the central leaves and 5 and 10 mm for the outer leaves, respectively.
AAA: Anisotropic Analytical Algorithm, CC: Collapsed Cone

3.2 OVERVIEW OF THE METHODOLOGY

The initial studies in Paper I investigated the performance of a recommended QC method based on three head and neck (H&N) IMRT treatment plans. IMRT was the representative modulated treatment technique used at the department at the time of the conduction of the study. These plans consisted of 9 IMRT beams with sliding window technique. The calculated absorbed dose distributions of these three plans were modified by introducing known dosimetric differences to enable an evaluation the detection ability of the investigated QC method.

Further studies were conducted in Paper II to study dosimetric differences between calculated and delivered dose distributions. Six series including five static beam apertures of gradually varying aperture shape complexity were created. A more specific investigation of how the aperture shape affects the dosimetric differences was enabled with the use of static beam apertures, as the uncertainties of dynamic delivery was avoided. The series of static beam apertures were designed to either be of constant or varying circumference and area to investigate these parameters impact on dosimetric uncertainty. Three aperture-based complexity metrics were developed based on the theory of the challenges of small field dosimetry within the work of Paper II. These metrics, and additional five other metric previously suggested in the literature, were evaluated as the linear correlation of the calculated complexity scores and the dosimetric differences measured for the static beam apertures.

The static beam apertures were created for the only purpose of the study, and the performance of the metrics were therefore needed to be validated on a set of more clinically relevant material in Paper III. A set of static fields originating from control point beam openings from clinically used VMAT plans was used for this purpose. A total of 18 control points were selected from six treatment plans planned for different diagnoses. Three control points of varying complexity from each of the six treatment plans were selected for inclusion of the whole range of complexity based on EAM scores. Also, control points originating from different treatment plans but with the same EAM score were selected to study EAM performance in a situation of the same EAM score but for different beam aperture size and shape.

Next step of the studies in Paper III was to include uncertainties of parameters related to the dynamic delivery. Treatment plans that have been accepted and used for treatments of different diagnoses were selected in four groups with the intention to include treatment plans planned for treatment targets of different sizes. A total of 200 VMAT plans whereof 50 plans in each group. The plans were grouped to include plans with larger targets in the pelvic and

H&N region, smaller targets for prostate treatments and treatments in the thorax region. VMAT was the representative modulated treatment technique used at the department at the time for the study.

The assessment of dosimetric uncertainty using different methods was studied based on a total of 30 VMAT plans created using three different optimization methods. Three treatment plans of different design but similar dose distributions were created for ten patient cases. The ten patient cases included five plans created for H&N cancer treatments and five plans created for prostate cancer treatments. The ten original plans have been accepted and used for treatment at our department but have in our study been reoptimized for the purpose of the study.

3.3 MEASUREMENT-BASED ASSESSMENT OF UNCERTAINTY LEVEL

The methods used for measurement and evaluation have been carefully selected based on the requirements of spatial resolution, dosimetric accuracy, time efficiency and the possibility to evaluate the aspect of interest. Some methods have been applied for more than one study with few adjustments to serve the different purpose of each study.

Paper I: Delta^4 measurements and ICRU 83 recommended evaluations.

Paper II: EPID and EBT3 film measurements. Dose difference evaluation.

Paper III: EBT3 film and Delta^4 measurements. Dose difference and gamma evaluation.

Paper IV: Delta^4 measurements and dose difference evaluation.

3.3.1 TWO-DIMENSIONAL MEASUREMENTS USING EPID AND FILM

The dosimetric challenges for static fields due to size and shape of the beam aperture were evaluated in Paper II and Paper III based on 2D Gafchromic EBT3 film measurements. Film dosimetry was chosen since it was a recommended method appropriate for 2D measurements of penumbra regions[53,54]. But film is a time-consuming measurement procedure, referring to time for preparation, developing of the film and read out, but offers a high spatial resolution which enable a careful evaluation of spatial differences. The calibration procedure, preparation, and read out of the films is based on a double exposure method[55] and described in more detail in Paper II.

The film measurements were performed on one occasion in Paper II and three occasions in Paper III. The repeated measurements for each field in Paper III were to account for uncertainties related to differences in sensitivity between different, and within the individual, film sheets. Daily variations of the treatment machine were thereby also included. All measurements were done with the gantry at 0° and the film was positioned at 10 cm depth in a solid water phantom with a source-surface-distance (SSD) of 90 cm (Figure 1).

FIGURE 1. Measurement setup for the EBT3 film measurements using a solid water phantom with the films placed at 10 cm depth and the gantry at 0°.

The evaluation of the film measurements was performed in the software RIT113 (Radiological Imaging Technology©). The film measurements were corrected for any rotational errors based on markings of the hair cross in the margin of the films. The measured dose distributions in Paper II were aligned to the planned absorbed dose distribution, calculated in the same geometry, based on vertical and horizontal dose profiles. This alignment method might correct for true deviations due to non-proper modelling of the MLC leaves within the TPS. In Paper III, the alignment method was instead based on a registration of the measurements to the calculations according to the common point of the isocentre defined by the hair cross markings on the film.

The median dose value of a region of interest (ROI) placed centrally in the most open part of the field in both the measured and calculated dose distribution was used for normalization. The size of the ROI was in most cases 0.5 cm x 0.5 cm but adjusted for the smallest field openings to a minimum of 0.25 cm × 0.25 cm.

The series of static fields included in Paper II were also measured using an amorphous silicon electronic portal imaging device (EPID aSi1000, Varian Medical Systems) which was mounted on the treatment machine and used for 2D imaging and dosimetry of the treatment beam. The data collection enabled a digital image processing and an immediate evaluation. The measurements were evaluated in the integrated software module Portal Dosimetry (Varian) in calibration units (CU).

The EPID measurements of the static fields in Paper II were repeated for three occasions with the gantry at 0° and a source-to-measurement distance of 100 cm. Each occasion of measurement started with a calibration of the EPID and thereby no normalization of the measurements was needed. The stability of the calibration was checked by measurements during the series of measurement on each occasion. A 10 cm x 10 cm MLC opening was measured before, during, and after the series of measurement and the mean value of a suggested shift in two directions according to an auto-alignment of these three measurements was used to align the measurements of the static fields.

3.3.2 THREE-DIMENSIONAL MEASUREMENTS USING DELTA^4 DOSIMETRY SYSTEM

The quasi-3D Delta^4 dosimetry system (ScandiDos, Uppsala) was a solution that consisted of both a hardware for measurement and a software for measurement analysis[56]. The hardware was two orthogonal detector planes inserted in a cylindrical PMMA phantom. A total of 1069 diode detectors were positioned in a grid pattern on the detector planes with a spacing of 0.5 cm in the central area of 6 cm x 6 cm and a spacing of 1 cm outside this area. Measured data was given for the measurement points, where the detectors were positioned. The measurement could be evaluated in the measurement points, but the software provided the possibility for a calculated 3D dose distribution based on interpolation of the data in the measurement points.

The Delta^4 dosimetry system have been used for measurement and analysis for treatment plans in Paper I, Paper III and Paper IV. Each of the included treatment plans has been measured once on one occasion. The calculated absorbed dose distributions, used for reference during measurement evaluation, for each treatment plan have been re-calculated for the same geometry as the phantom.

The measurements were normalized prior dose evaluation to correct for daily variations in dose output of the treatment machine. The mean value of the relative dose deviations between the measured and calculated dose for all measurement points for the three H&N IMRT plans was used to normalize the measurements in Paper I. The measurements of the VMAT plans included in the studies of Paper III and Paper IV were corrected for daily variations based on measurements of a half arc (270°-90°) 10 cm x 10 cm field shaped by the collimator jaws prior to each measurement occasion.

The Delta^4 software provided Optimized Phantom Position (OPP) which was a tool that suggested a virtual positional shift in three directions for a better dosimetric agreement between measurement and calculation. An individual shift was applied for each of the measurements for the 200 VMAT plans included in Paper III, since these plans were measured different days as part of the routine patient-specific QA at the department. The mean value of the shifts suggested individually for the 30 plans included in Paper IV was used for alignment of the measurements as the measurements were carried out all at the same occasion. The OPP tool was not applied for the IMRT measurements in Paper I.

3.3.3 EVALUATION OF MEASUREMENT RESULTS

Corrections of the measured data, due to positioning errors, variations in detector response, or daily output variations of the treatment machine, were performed before evaluation of the dose difference between measured and calculated absorbed dose distributions. The applied corrections have been described in the previous sections for each measurement procedure, respectively. The software used for the evaluations have also been specified in the same section.

The work in Paper I were aimed to study if a measurement-based QC method in combination with an evaluation method recommended by the ICRU report 83 could detect introduced dose errors. The ICRU report 83 recommended a separated evaluation for the high (>20%/cm) and the low (<20%/cm) absorbed dose gradient regions. The low dose gradient region, where the dose level is stable, was recommended to be evaluated using a global (normalized to the prescribed dose) dose difference evaluation. The tolerated deviation between measurement and calculation for this region was defined as a standard deviation of ±3.5%. The tolerance level for the high dose gradient region was defined as a distance-to-agreement (DTA) criterion of a standard deviation of 3.5 mm. The ICRU report 83 translates these tolerance levels to 5% global dose difference with a pass rate higher than 85% and DTA of 5 mm with a pass rate higher than 85% for each region. ICRU 83 suggested gamma evaluation with criteria 5%/5mm and a pass rate tolerance level of 85% as an alternative, more strict, method. Gamma evaluation combines dose difference and DTA evaluation for a point-by-point comparison of two dose distributions[57].

Most departments have applied gamma evaluation within their routine QC method with more strict gamma criteria than recommended by ICRU 83. Gamma evaluation has been a widespread used method, but the choice of gamma criteria has been different between departments. However, the correlation between gamma pass rate and other measures of dosimetric differences has been shown to depend to large extent on the choice of gamma criteria[58-61]. It has been shown that gamma evaluation pass rate results not always have been optimal to use for evaluating dosimetric differences of clinical relevance between two dose distributions[49,58-65]. The concept of an evaluation method that includes a tolerance for spatial deviations, as a DTA criteria, would be to account for spatial errors due to the measurement procedure. Gamma result must be interpreted with the knowledge that true spatial deviations might be undetected.

The quasi-3D Delta^4 measurements and the 2D film measurements of Paper III was evaluated with gamma evaluation with criteria 3%/1mm. A DTA criteria this small could be used if the alignment and normalization has been done carefully. The result of the film measurements in Paper III was reported as the mean value of three measurements. These measurements were also evaluated with a global dose difference criterion of 5% for the film measurements and 3% for the Delta^4 measurements. Similarly, the Delta^4 measurements in Paper IV were also evaluated using a global dose difference criterion of 3%. The EPID and film measurements in Paper II have also been evaluated using global dose difference criteria of both 3% and 5%.

Regions of low dose, i.e., the tail of the penumbra, was not of interest in these evaluations. Therefore, a dose cut off was applied to only evaluate the dose distribution with dose levels higher than 10% of the maximum calculated dose for the measurement evaluation in Paper II and Paper III, and 20% of the maximum calculated dose or the absorbed dose in the isocenter or for Paper IV and Paper I, respectively. For comparison, the evaluation was also done with all diodes included in Paper I.

3.4 ANALYTICAL ASSESSMENT OF UNCERTAINTY LEVEL

3.4.1 COMPLEXITY METRICS

Different types of complexity metrics have been used in Paper II, III, IV to assess the dosimetric uncertainty of static fields and treatment plans.

Deliverability metrics have been described to focus on differences between the planned and delivered absorbed dose distribution due to dynamic variations of machine parameters during delivery[38]. One example of a metric within this category was the Modulation Index (MI), first suggested by Webb et al.[66] but further developed by Park et al.[45]. The MI have been developed as three different versions. The MI_t included the speed and acceleration of the MLC leaves, the acceleration of the gantry and dose rate variations. The MI_a and MI_s included only the speed and acceleration of the MLC or only the speed of the MLC, respectively. The MI_a have been used in Paper IV to assess the complexity level of 30 VMAT plans with respect to dosimetric uncertainties related to the dynamic delivery. The MI_a calculations were performed using an in-house developed script using Eclipse Scripting Application Programming Interface (ESAPI).

Accuracy metrics have been described to focus on inaccuracies in the TPS. These types of inaccuracies would mainly appear for small and irregular beam openings and these types of metrics were therefore mainly based on the characteristics of shape and size of the MLC openings. However, dosimetric differences between planned and delivered absorbed dose distributions are most often a result of a combination of both uncertainties in deliverability and inaccuracies in dose calculation. Aperture-based complexity metrics were based on characteristics of the shape and size of the MLC aperture with the intention to include both uncertainties related to the delivery, for example differences in planned and actual position of the MLC leaves, and uncertainties related to inaccuracies of the dose calculation in the TPS. These uncertainties mostly manifest as dosimetric differences in the edges of the treatment fields.

Three aperture-based complexity metrics, edge area metric (EAM), circumference/area and converted aperture metrics (CAM), were developed in the work of Paper II. An ideal complexity metric includes uncertainties related to both calculation and delivery which could lead to differences between the planned and delivered dose. Also, the metric should be designed to be easy to understand and apply.

Edge area metric

EAM was developed in the work of Paper II as an aperture-based complexity metric which was defined as a measure of the relative amount of the total aperture that was within a specified region near the MLC edge according to Eq. 1.

$$EAM = \frac{R_1}{R_1+R_2} \qquad \text{(Equation 1)}$$

The region close to the MLC edge, R_1, was defined as a specific range, that can be defined by the user, on both the inside and outside of the MLC edge. The remaining part of the aperture was defined as the open non-complex part, R_2. The regions are illustrated for an arbitrary beam opening in Figure 2. The EAM was developed with the intention to be a metric that was intuitive to interpret and understand. The EAM scores are also, according to the definition of Eq. 1, on a finite scale ranging from 0 to 1, where a higher value indicates a higher complexity level.

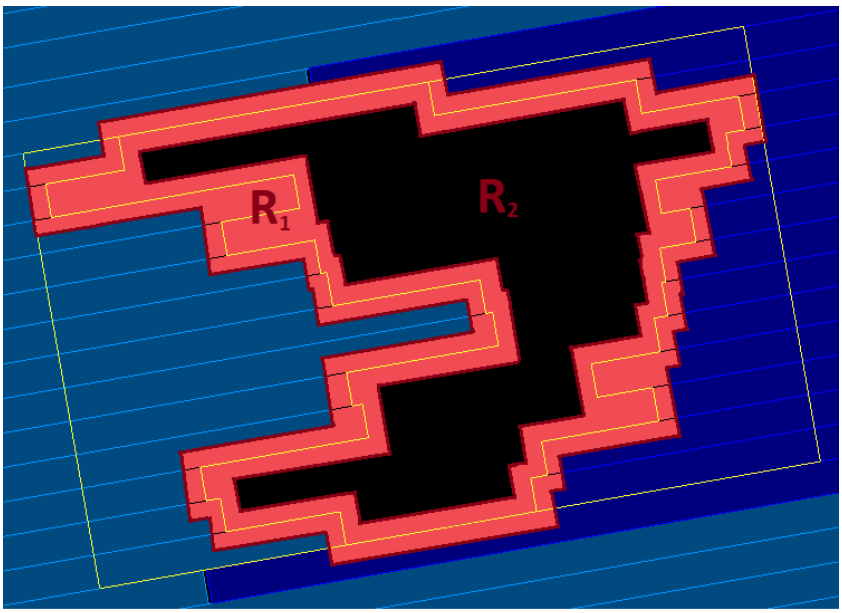


FIGURE 2. Illustrative figure of the edge area metric (EAM) and the complex (R_1) and non-complex (R_2) regions. The red area illustrates R_1, enclosing an area on both inside and outside of the MLC edge, which is marked as a yellow line, and R_2 which is the open part within the beam aperture seen as the black area in the figure.

Uncertainties related to delivery and dose calculation, which mostly manifest in the edge of the field, were therefore considered to be included in the theory behind EAM. The range of the complex region, R_1, was defined as 5 mm on both sides of the MLC edge in Paper II.

EAM was further developed by studying the optimal value for R_1. EAM scores were calculated using R_1 settings in steps from 1 mm to 5 mm for a total of 13 VMAT arcs originating from plans planned for treatments of H&N, anus, and prostate cancer and planned to be delivered by treatment machines mounted with both Millennium MLC and High Definition MLC. The largest separation between the control points within the same arc, i.e., the highest and lowest EAM score, was found when R_1 was 2.5 mm on both sides of the MLC edge. This setting, with R_1 as 2.5 mm, were used for the further EAM calculations in Paper III and Paper IV.

EAM was derived to be calculated on a control point level and a mean value of the control point scores has been used to express the EAM score on a treatment plan level in Paper III and Paper IV. A mean value does unfortunately not give any information about the complexity variations for the different gantry angles. The EAM results can, with advantage, be presented as polar plots as a function of gantry angle to enable a visual assessment of this complexity (Figure 3). Other types of information, e.g., number of MU on a control point level, can be presented and interpreted simultaneously.

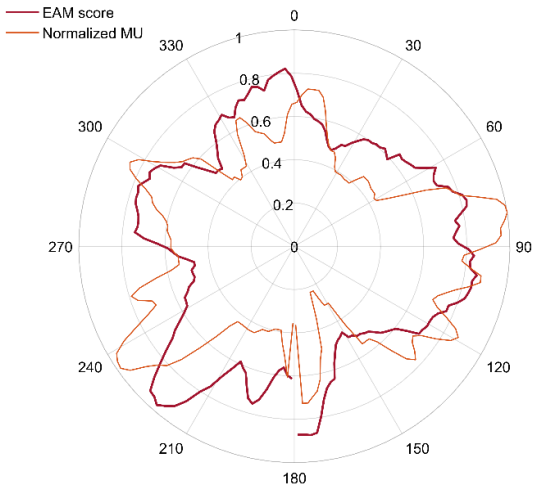


FIGURE 3. An arbitrary example of edge area metric (EAM) scores and normalized number of monitor units (MU) on a control point level plotted in a polar plot as a function of gantry angle as dark red and orange lines, respectively. In this example, the maximum number of MU that is planned to be delivered between to control points is used as normalization value.

Circumference/Area ratio

The ratio of the beam aperture circumference and the aperture area was suggested as a more intuitive alternative to EAM for compassion purpose to EAM in Paper II.

Converted aperture metric

CAM quantified the smallness and irregularity of a beam aperture based on measured distances within the aperture both parallel and opposed to the MLC travelling direction (Figure 4).

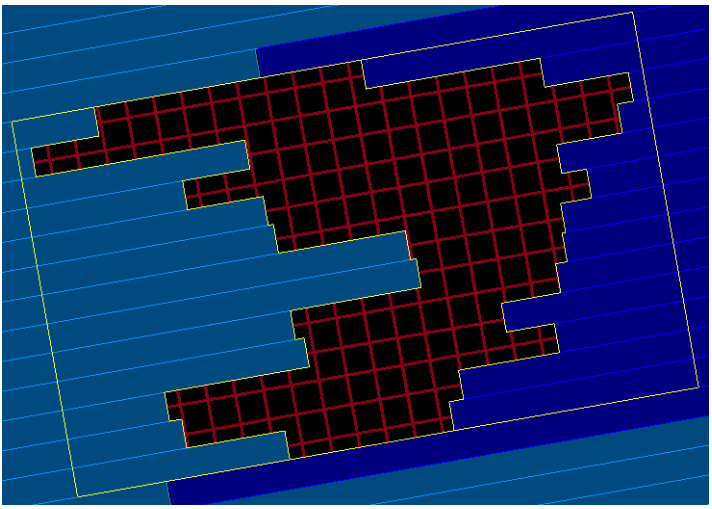


FIGURE 4. Illustrative figure of converted aperture metric (CAM) with the measured distances within the beam aperture, both parallel and opposed to the MLC travelling direction.

The distances were measured every 5 mm to give at least one measure for each MLC leaf for the MLC type used in the study of Paper II, which was the Millennium 120 MLC for which the smallest MLC in the central part was 5 mm wide. The small distances, x, are converted to a penalty score between 0 and 1 according to a function $f(x) = 1-e^{-x}$. Distances larger than 4 cm were considered to be non-complex with a penalty score equal to 1. The square root of the total aperture area was also converted according to the same function. The final CAM score was calculated according to Eq. 2 as the mean of all penalty scores of the measured distances, multiplied with the penalty score of the equivalent aperture area. A subtraction from 1 was done so that higher scores were achieved for more complex apertures, for an easier interpretation.

$$CAM = 1 - \overline{f(x_i)} \cdot f\left(a_{eq}\right) \qquad \text{(Equation 2)}$$

Evaluation of developed metrics

The developed metrics EAM, circumference/area ratio, and CAM, were evaluated by comparing their performance to other previously suggested metrics, also defined as accuracy metrics: the edge metric (EM)[37] and aperture area (AA)[67], and other metrics defined as deliverability metrics: modulation complexity score (MCS)[36], MU/Gy[43,68,69], and aperture irregularity (AI)[67]. Complexity scores of all mentioned metrics were calculated for the 30 static fields included in Paper II. Metric calculations were performed with in-house developed MatLab® software which used DICOM files as information input.

The performance of the metrics was evaluated as the correlation, expressed as the Pearson's r-value, to the 3% and 5% dose difference pass rate of EBT3 film and EPID measurements for each of the static fields in Paper II, evaluated according to the methods previously described. A r-value closer to ±1 indicates a stronger linear correlation. The performance of EAM was further evaluated for 18 static fields using film measurements and 200 VMAT plans using Delta^4 measurements in Paper III.

3.4.2 DELIVERY SIMULATIONS

The impact on dosimetric differences between planned and delivered dose distributions due to delivery variations of machine parameters, i.e., the MLC leaf positions, collimator jaw positions, collimator angle, and gantry angle, was estimated by comparing the original calculated dose distribution and the calculated dose distribution of the simulated deliveries in Paper III.

The planned values for the mentioned plan parameters were replaced in the DICOM file to values that would simulate variations that were realistic to occur during delivery. The modification of the DICOM files was performed with a self-developed Matlab® script. Realistic variations were assumed to include systematic offsets due to uncertainties of the calibration procedure, and random offsets due to daily variations of the mechanical parts. The maximum extent of these possible offsets was estimated based on the variations that can be observed in the quality controls of the treatment machine and specifications from the treatment machine vendor. The simulation values were randomly generated from a normal distribution centered around the planned value and truncated to only include values within the maximum offsets. A more detailed description can be found in Paper IV. The modified DICOM files were imported back into the TPS and the dose distributions of the simulated deliveries were calculated.

The plan robustness to treatment machine parameter variations was investigated with the method of delivery simulations for 30 VMAT plans in Paper IV. Ten calculated dose distributions, including the original plan and nine simulations, were compared for each plan. The standard deviation (SD) of the ten dose values for each voxel in the distributions was calculated as a measure of the plan robustness for each plan (with a threshold of 20% of the prescribed dose). An advantage of this method was the possibility to evaluate the extent and position of the dosimetric differences directly within the patient geometry.

3.5 SUMMARY OF METHODOLOGIES

3.5.1 IMRT PATIENT-SPECIFIC QA - PAPER I

The IMRT patient-specific QC method evaluated in this study was the combination of the Delta^4 dosimetry system and the ICRU report 83[30] recommendations for comparison of calculated and measured dose distributions.

The recommendation by ICRU 83 was to evaluate a dose distribution separated in a low dose gradient region (< 20 %/cm) to be evaluated using a dose difference criterion, and a high dose gradient region (> 20 %/cm) to be evaluated using a DTA criterion. The recommended criteria were 5% global dose difference with a pass rate higher than 85% and DTA of 5 mm with a pass rate higher than 85%, for respective region. The recommended evaluation was performed outside of the Delta^4 software, as it was not possible to evaluate a low gradient region with a separate criterion. Gamma evaluation was also performed with criteria of 5% dose difference and 5 mm DTA and a passing rate tolerance of 85%, according to the alternative method suggested by ICRU 83. The gamma evaluation was performed within the Delta^4 software but was limited to the measurement points, i.e., the positions of the diode detectors.

The QC method was evaluated by investigating its ability to detect deliberately introduced dosimetric errors. The study consisted of three H&N IMRT plans, using the sliding window technique, created in Eclipse TPS (Eclipse version 10.0, Varian Medical Systems). The three original plans were delivered at the treatment machine for which they were planned, and measured using the Delta^4 phantom. The measured dose distributions were compared to calculated dose distributions with deliberately introduced dosimetric errors, a total of 28 modified calculated dose distributions. Details about the introduced dosimetric errors can be found in Paper I.

Dosimetric differences larger than 5% in evaluated dose-volume metrics, for example the near-maximum ($D_{2\%}$) dose to the spinal cord, were considered to be of clinical relevance and were expected to be detected by an adequate QC method.

3.5.2 DEVELOPMENT OF APERTURE-BASED COMPLEXITY METRICS FOR STATIC FIELDS - PAPER II

Clinically relevant dose differences were shown to not always be detected by an evaluated measurement-based QC method (Paper I). This led to the initiative of Paper II: development and investigation of the use of aperture-based complexity metrics for estimation of dosimetric uncertainties in modulated radiotherapy plans.

Aperture-based metrics, to be valid for both IMRT and VMAT, were developed to consider the uncertainties described in a previous section about challenges of small field dosimetry. One metric developed within this study was the converted aperture metric (CAM). CAM combined the information about distances within the beam aperture, both parallel and perpendicular to the direction of the MLC leaves, and the beam aperture area. A smaller distance or area indicated that the evaluated beam aperture was small and/or irregularly shaped. The measured distances were penalized according to an inverse exponential function described in Paper II.

Another complexity metric, edge area metric (EAM), was defined as the amount of MLC edge in relation to the total beam aperture opening. The beam aperture was divided in a complex region enclosing an area of 5 mm (isocenter distance) on both sides of the MLC edge. The remaining area of the beam aperture was the non-complex region. The EAM score was defined on a finite scale between 0 and 1, according to Eq. (1) in the previous section of complexity metrics. A simpler version of EAM was the ratio of the circumference and area of the beam aperture, which was included in the study for comparison.

Additionally, five previously defined aperture-based complexity metrics were evaluated in this study: the modulation complexity score (MCS)[36], the edge metric (EM)[37], MU/Gy[43,68,69], the aperture area (AA) metric[67], and the aperture irregularity (AI) metric[67].

The performance of the complexity metrics was compared for 30 static beam apertures. The beam apertures were created in different series with apertures of different sizes and shapes to simulate IMRT/VMAT control points of varying complexities. The beam apertures were created to vary in size but with a constant circumference, and the other way around, to enable a systematic evaluation of the impact of shape and size on the complexity.

The complexity of the beam apertures was estimated based on the extent of dose differences between the calculated and measured dose distributions. All 30 beam apertures were measured using Gafchromic™ EBT3 film (Ashland) and amorphous silicon electronic portal imaging device (EPID aSi1000, Varian Medical Systems). The dose difference between planned and measured dose distribution was evaluated using dose difference criteria of 3% and 5%.

The agreement between the complexity of the beam aperture based on the 2D measurements and calculated complexity scores was compared for the different metrics.

3.5.3 DEVELOPMENT OF AN APERTURE-BASED COMPLEXITY METRIC FOR CLINICAL FIELDS - PAPER III

The edge area metric (EAM), developed in Paper II, showed good results for scoring complexity of static beam apertures. The dynamic nature of the delivery was not included in Paper II, hence the need for further studies. In Paper III, EAM was evaluated for VMAT beam apertures and plans used for treatment delivery.

Smaller adjustments of the EAM metric were performed to optimize the EAM to achieve the largest possible score separation for apertures of various complexity. The largest separation of EAM scores was achieved when the complex area around the MLC edge included 2.5 mm (isocenter distance) on both sides of the MLC edge. This was used for the EAM calculations in this and following studies.

Static beam apertures originating from control points of VMAT plans that have been used for treatment were measured using EBT3 film, similarly to the procedure of Paper II. This was done to evaluate the correlation between the complexity and the EAM score for beam apertures relevant for clinical use. Three beam apertures of various complexity from six different plans were selected for this purpose. Also, some of the beam apertures with the same EAM score, but originating from different plans, were selected to evaluate how EAM scores the complexity for beam apertures of varying shape. The VMAT plans included in this study have been used for treatment of cancer in the pelvic, thorax, H&N, and the prostate.

The complexity on a treatment plan level was derived from Delta^4 measurements that were performed as pre-treatment patient-specific QC at our department. Results from a total of 200 plans were collected – 50 plans from each of the treatment sites previously mentioned. The Delta^4 measurements were evaluated using a global dose difference criterion of 3% and gamma evaluation with 3%/1mm. The complexity on a treatment plan level based on the Delta^4 measurement results was compared with the arithmetic mean value of the EAM scores calculated for all control points of the plan (except control points with zero dose rate).

It was also investigated if higher EAM scores were related to the planning target volume (PTV). EAM should, by definition, result in higher scores for small beam apertures which it was assumed to be more of for treatment plans planned for smaller targets.

3.5.4 DOSIMETRIC UNCERTAINTY ASSESSMENT METHODS FOR VMAT PLANS - PAPER IV

Ten VMAT treatment plans, five plans planned for prostate cancer treatment and five plans planned for tonsil cancer treatment, that have been used for treatment were included in this study. These plans were replanned in Eclipse with and without using aperture shape controller (ASC) and in RayStation. The prostate cases were planned with the use of multi criteria optimization (MCO) which was available in both TPSs, but not for the plans planned with ASC since MCO and ASC could not be combined in current TPS version. The three plans for each patient case were planned with the aim to create similar calculated absorbed dose distributions, but with different treatment plan designs as a result of the different optimization methods used. The similarity of the dose distributions was evaluated by visually comparing the calculated dose distributions and analytically by comparing DVH metric values.

The dosimetric uncertainty due to the treatment plan design was evaluated for each patient case using measurement-based and analytical methods.

- Delta^4 measurements
- Delivery simulations
- Aperture shape complexity by EAM
- Modulation variations by MI_a
- Dose rate and gantry speed modulation variation analysis

Dosimetric differences between planned and delivered dose distributions were evaluated with measurements using Delta^4 dosimetry system. The measurements were evaluated using a global 3% dose difference fail rate. A method not suffering from measurement uncertainties, i.e., delivery simulations, was used to analyze treatment plan robustness to variations of machine parameters. The robustness of the plans was expressed as standard deviation of dose differences between ten calculated dose distributions for each patient case, one original dose distribution and nine simulations. The Delta^4 measurements and the delivery simulation results were evaluated for the distribution with a dose higher than a threshold dose of 20% of the calculated maximum dose. All calculations were performed with a grid size of 2.5 mm.

The aperture-based complexity metric EAM was used to score the aperture shape complexity on a control point and treatment plan level. The EAM score on treatment plan level was defined as the mean of the EAM scores on control point level as in Paper III. The modulation of MLC speed and acceleration was

evaluated using the modulation index, MI_a, described by Park et al[45]. The variation of the gantry speed and the dose rate modulation was statistically compared for the plans created with different optimization methods using a student's t-test.

Differences between the results of the applied dosimetric uncertainty methods were analyzed on a group level for the treatment plans created using the three different optimization methods.

3.6 ETHICAL CONSIDERATIONS

This thesis includes studies with focus on development and evaluation of methods to gain information about uncertainty level of modulated radiotherapy plans. The work was defined as a quality development and assurance project and no ethical consideration have been necessary. All data that originates from radiotherapy plans used for treatments have been anonymized and the results cannot be related to any patient identification information. The results of these studies have had no impact on any specific patient treatment, which therefore could not have been better or worse than the actual treatment given at the time.

4 SUMMARY OF RESULTS

The main results of each study are discussed in following sections. A more thorough and detailed presentation of the results for each study can be found in the original version of the Papers included in this thesis.

4.1 IMRT PATIENT-SPECIFIC QA - PAPER I

A patient-specific IMRT QC method based on the combination of measurements with the Delta[4] dosimetry system and evaluation of the measurement result based on the ICRU report 83 recommendations could not detect all cases of treatment plans which had a dose error considered to be of clinical relevance.

A total of 28 calculated dose distributions with introduced dosimetric deviations compared to the original dose distributions were created. All 28 plans passed the gamma evaluation with criteria recommended by the ICRU 83 of 5% dose difference and 5 mm DTA with passing rate higher than 85%. More than half of the plans, 18 of 28, had a passing rate of 100%. Only one plan had a passing rate less than 90%.

Evaluation according to the primary recommendations of ICRU 83, with a 5% dose difference criterion for the low dose gradient region and a 5 mm DTA for the high dose gradient region, resulted in pass rates lower than 85% for the low dose gradient region for six of the 28 plans.

These six plans were part of a group of 15 plans which had a dosimetric deviation exceeding a 5% difference compared to the original treatment plan for at least one specific DVH metric of interest. For example, four of the 15 plans had more than 5% higher near-maximum dose ($D_{2\%}$) to the spinal cord, 7 plans had more than 5% higher mean dose to one or both parotid glands and 9 plans had more than 5% lower mean dose to one or both target volumes. This means that nine plans with larger (more than 5%) dosimetric differences were undetected by this recommended QC method (Table 3).

Table 3. Categorization of the results for the evaluation of separated dose gradient regions recommended by ICRU 83 for 28 plans with deliberately introduced dose errors: 15 plans had more than 5%, and 13 plans had less than 5% dose difference in at least one evaluated dose volume histogram metric. Green and red color represents a correct or incorrect categorization of a QC result, respectively.

QC result, separated in low and high dose gradient regions.	**Dosimetric dev. >5%**	**Dosimetric dev. <5%**
Failed QC with pass rate <85%	6 plans	0 plans
Passed QC with pass rate >85%	9 plans	13 plans

4.2 DEVELOPMENT OF APERTURE-BASED COMPLEXITY METRICS FOR STATIC FIELDS - PAPER II

The series of static beam apertures were successfully created to be gradually more complex, as the 3% and 5% dose difference pass rates in general decreased for smaller and/or more irregularly shaped beam apertures. One series, series C, had a rather constant pass rate for all beam apertures as expected due to the similarity of the major part of the beam apertures. The overall pass rate results of 3% and 5% dose difference evaluation had similar trends but the pass rates of the 3% dose difference evaluation were at a lower level due to the stricter criterion. The EBT3 film and EPID measurement results were similar, but larger variations were observed for the pass rates within the series for the EBT3 film measurements compared to EPID. This could probably be because of larger intrinsic uncertainty for film measurement. It should be said that film measurements, in this study, had higher pass rates compared to EPID which probably can be explained by the normalization and the alignment of the film measurements compared to EPID measurements. The EPID measurements were evaluated in absolute mode in CU and the alignment was based on a measured 10 cm x 10 cm MLC opening and the same for all evaluated fields.

The dosimetric differences appeared mostly in the regions close to the MLC edge. The best correlation between EPID measurement results and complexity scores was found for the EAM with a Pearson's r-value of -0.94. Other metrics, also derived based on the theory of the challenges of small field dosimetry same as EAM, were edge metric and CAM which correlated to the measurement results with Pearson´s r-values of -0.89 and -0.88, respectively.

Some of the evaluated metrics could not differentiate between beam apertures of various complexity. For example, the beam apertures of series A which consisted of quadratic fields of various sizes had equal complexity score when calculated using Aperture Irregularity (AI) metric because the shape/irregularity of the fields was the same. In similar way, the Aperture Area (AA) metric resulted in similar scores for the beam apertures in each of series B, D and E since the areas were similar.

4.3 DEVELOPMENT OF AN APERTURE-BASED COMPLEXITY METRIC FOR CLINICAL FIELDS - PAPER III

The correlation between EAM scores and the 5% dose difference pass rate for static beam apertures, originating from VMAT plan control points, was strong with a Pearson's r-value of -0.96. A larger spread in dose difference pass rates was observed for the more complex static beam apertures according to higher EAM scores.

A relationship between total plan EAM scores and pass rates on a treatment plan level was not observed for the treatment plans in this study. Larger spread in pass rates was, however, observed for treatment plans with a higher total plan EAM score similar to the result on a control point level. Two plans could be identified with higher EAM scores in comparison to the group of evaluated plans. Further investigation showed that these two plans were planned for target volumes that were small in relation to the target volumes for the other evaluated treatment plans for respective diagnose group. This indicated that EAM could be used for detection of treatment plans that might need extra attention during QC or to be replanned to lower the dosimetric uncertainty.

This study showed that there was a difference in level of EAM scores for treatment plans planned for different diagnoses. This result was seen both on control point and treatment plan level for the 200 evaluated treatment plans. Treatments planned for smaller target volumes, which in this study were for the prostate cancer treatments, had the largest fraction of control points with higher EAM scores just as plans planned for larger target volumes, in this case vulvar cancer treatments, had a larger fraction of control points with lower EAM scores. The result indicated that EAM scores need to be interpreted in relation to a benchmark score specifically derived for the treatment group.

4.4 DOSIMETRIC UNCERTAINTY ASSESSMENT METHODS FOR VMAT PLANS - PAPER IV

The dose distributions for the three treatment plans optimized for each patient case were compared and had similar DVH metric values within 1.5-2 local percentage points and the dose was distributed similarly within the patient geometry.

The results of the dosimetric uncertainty assessment methods were summarized in Table 4.

Table 4. General results of estimated dosimetric uncertainty with the use of different complexity assessment methods on a group level for plans created with three different optimization methods. The level of uncertainty was graded as lower/higher relative between the groups of plans and based on an overall assessment of the result for each complexity assessment method. In some cases, the result cannot be separated (≈) for two groups.*

Dosimetric uncertainty assessment methods	Lower → Higher uncertainty
EAM scores	RS ≈ ASC < EC
MI_a scores	EC ≈ ASC < RS
Dose rate and gantry speed variations	EC ≈ ASC < RS
Delta^4 measurements	RS < ASC < EC
Delivery simulations	RS < ASC ≈ EC

*EC: Eclipse without aperture shape controller, ASC: Eclipse with aperture shape controller, RS: RayStation

Treatment plans generated in Eclipse without ASC were more complex than plans created in Eclipse with ASC according to patient-specific Delta^4 measurements, which was a common method to evaluate dosimetric differences between planned and delivered dose distributions. Similar results were found by the EAM analysis.

Treatment plans created in RayStation were less complex than plans created in Eclipse according to the measurement results. Plans created in RayStation were also least complex according to the delivery simulation results.

Evaluation with complexity assessments methods that focus on variations of MLC speed and acceleration, or dose rate and gantry speed variations showed a higher modulation for plans created in RayStation compared to plans created in Eclipse with and without ASC.

The results of the Delta^4 measurements and the EAM analysis showed that treatment plans planned for prostate cancer treatments were more complex than plans created for treatment in the H&N region. This result was also seen in Paper III.

The results in this study indicated that different type of complexity information can be gained by using different methods. The relation between complexity scores and dosimetric uncertainty in terms of clinical relevance is not yet determined.

5 DISCUSSION

The primary purpose of the studies presented in this thesis was to evaluate and develop methods to assess dosimetric uncertainty for modulated radiotherapy plans. The main result of Paper I was that a recommended measurement-based QC method was not able to detect deliberately introduced clinically relevant dosimetric differences between measured and calculated dose distributions. This result indicated the need for a more reliable method to identify treatment plans with relevant dosimetric uncertainties.

In Paper II, aperture-based complexity metrics were developed based on the theory that dosimetric uncertainties occurs because of the dosimetric challenges of small and irregularly shaped beam apertures. The developed metrics were based on parameters related to calculation and delivery of penumbral dose, and the performance of the developed metrics was compared to other previously suggested metrics found in the literature. A strength of Paper II was the systematic investigation of the impact of beam aperture design on dosimetric differences between calculation and measurement. This investigation showed that dosimetric differences between measurement and calculation of static beam apertures were mainly manifested in the peripheral parts of the field, i.e., the penumbra region.

One of the developed metrics in Paper II, EAM, showed promising results. EAM is defined as the relative part of a field opening that is close to the field edge compared to the part that is in the open part of the field. The strongest linear correlation to dosimetric differences assessed by 2D measurements was found for EAM. Another developed metric, CAM, also showed a good correlation to measurement results. CAM was derived based on theoretical reasonings about a non-linear relationship between complexity and aperture shape irregularity and aperture size, which unfortunately led to a rather complicated definition. However, the good correlation result for CAM gave strength to the theoretical reasoning that complexity metrics should be based on parameters related to uncertainties that manifest in the penumbra region. It is of my opinion that a method should be both intuitive to understand and easy to use and, therefore, further work has for the time being only focused on further developments of EAM.

In Paper III, EAM has been developed and evaluated for clinical cases. One result of Paper III was that the mean value of the EAM scores on control point level was found not to correlate with differences between measured and calculated dose distributions on a treatment plan level. This result was concluded based on an evaluation of a large group of VMAT plans of mixed

diagnoses. Any specific correlation between dosimetric differences and EAM scores for each diagnose type was not investigated in this study.

An evaluation of EAM on control point level showed that the beam aperture complexity could vary within the treatment field, i.e., one part of the field could consist of more complex control points leading to a higher dosimetric uncertainty while another part of the field could consist of control points that are less complex leading to lower dosimetric uncertainty. The mean value of control point complexity scores might not be able to reflect these complexity variations in a way that agrees with the overall complexity of the total treatment plan. To be useful in the clinical workflow EAM needs to be a metric applicable for evaluation on treatment plan level, and therefore further studies are needed to investigate how the EAM score information on control point level best could be presented to reflect the complexity on an overall treatment plan level. This is not straight forward since it might be necessary to include other parameters to weight the contribution of the different control points to the final complexity for the total treatment. Some metrics included the number of MU as a weight of each control point to the contribution of the overall complexity[36,70]. However, this is not an optimal method to weight the contribution of each control point beam aperture. The number of MU as a weighting factor would be used as a substitute for the actual dose contribution for each control point, but the number of MU to deliver the same absorbed dose will be higher for beam apertures that are smaller and/or of more irregular shape. This means that complex beam apertures would be weighted higher even though they might not necessarily contribute to a larger part of the treatment.

Another result in Paper III was that EAM score depends on the target volume, which indicated that EAM needs to be interpreted separately for different diagnose groups. It has been discussed in other studies that the complexity also might be varying for different TPS and treatment techniques[42,71,72]. Further development of complexity metrics must either account for these types of differences within the definition of the metric or the metric needs to be interpreted with these aspects in consideration.

The results in Paper IV, as other studies[37,73], showed that similar dose distributions can be created by plan designs of various complexity. Studies have shown that there was not always a correlation between the dosimetric quality, in terms of achieved treatment goals, and the complexity, referring to the extent of dosimetric difference between planned and delivered dose distribution due to treatment plan design[74,75]. Differences between planned and delivered dose distributions have been shown to depend on the complexity of the treatment plan[17]. The information gained by different methods for assessing

the level of complexity of a treatment plan was also evaluated in Paper IV. It was shown that different methods could generate different but valuable information when creating and selecting plans for treatment.

Another finding in Paper IV was that according to the Delta[4] measurement evaluation results, variations in modulation, i.e., acceleration and speed variations for the MLC and variations in dose rate and gantry speed, had a rather low impact on dosimetric differences. This suggests that larger variations of modulation parameters do not necessarily lead to larger dosimetric uncertainty of the treatment. The relationship between these modulation parameters and their impact on dosimetric uncertainties needs to be further evaluated.

Dosimetric uncertainty of a treatment plan was, on the other hand, shown to be related to the shape and size of the beam apertures. This can be concluded by the results of Paper II, which showed the best correlation results for the use of EAM, edge metric[37] and CAM which are aperture-based metrics. Further, Paper IV showed that the optimization method that included aperture shape controller (ASC) had the largest impact on lowering the aperture complexity and the dosimetric uncertainty, according to the EAM and Delta[4] measurement evaluations.

Quasi-3D measurements, such as those performed with the Delta[4], provides an indication of dosimetric uncertainty of the treatment plan as this type of measurement-based method includes dose calculation uncertainties and the delivery uncertainties – but limited to the delivery for the occasion of measurement. Additional uncertainties are, however, introduced in the result due to the measurement procedure and the evaluation method.

These measurement uncertainties could be overcome using simulated deliveries. The applied method for delivery simulations was derived based on the uncertainties related solely to the delivery and do not include uncertainties that are due to the limitations of the dose calculation or variations in modulation. The results of the delivery simulations were similar to the Delta[4] measurement results in Paper IV which means that the two methods confirm each other's results.

There are always room for improvements of the applied methods. For example, the evaluated analytical methods have limitations due to how they are defined and what sources of uncertainty they include. The method for delivery simulations used in Paper IV was limited to only include variations of treatment machine parameters that are involved in the mechanical delivery. For example, it could be extended to include variations in dose output. The work

of developing a method for the purpose of assessing dosimetric uncertainty is a continuous work and the insights gained by the work described in this thesis are steps towards a better understanding which can hopefully contribute to the development of current or new methods. The result of Paper IV showed that different types of methods, for example measurement-based and analytical complexity methods, could be a complement to each other as they provide different types of information.

One suggested purpose with the use of complexity metrics was to detect treatment plans that need special concern during QC. This could decrease the workload for the personnel, as the information of complexity metrics can be used to streamline the time and effort of the QA. Also, the information available from the different complexity metrics can be used to create better treatment plans, i.e., plans designed in a way that minimizes the dosimetric differences between planned and delivered dose distributions. One finding in the study of Paper IV was that the use of aperture shape controller (ASC), available in Eclipse TPS, was shown to lower the treatment plan complexity without compromising the clinical goals of the dose distribution for the plans included in that study.

Treatment plans that are to be included in a prospective clinical study could preferably be reviewed according to criteria that also include criteria for treatment plan complexity, as it has been shown to be related to the uncertainty of the delivered dose distribution (Paper II-IV). The use of criteria for dosimetric uncertainties could enable an inclusion of treatment plans that are designed in a more similar way, which would make it possible to draw more certain conclusions. Although it is not possible to state that the level of treatment plan uncertainty expressed as a complexity score would have an impact on the individual patient treatment outcome, it would ensure that the dose distributions of the treatment plans included in the study were created in a way that limits the dosimetric uncertainties.

A recurring difficulty in the studies in this thesis was to evaluate methods for assessment of specific uncertainties without introducing other uncertainties caused by the evaluation process itself. Examples of ways to get around the effect of these introduced uncertainties have been to carefully consider strategies for alignment of measurements prior to evaluation; developing a method to theoretically simulate delivery, in order to avoid measurement uncertainties; and interpreting the combined results of different methods that evaluate different aspects of uncertainties.

A strength of the work presented in this thesis was the systematic approach to first identify the problem, analyze the impact of involved parameters on dosimetric differences for both static fields and for VMAT fields, and to compare the results of different methods. There are no golden standard when it comes to validation of complexity metrics and systematic evaluations are therefore necessary.

6 CONCLUSION

The overall conclusion of this thesis was that dosimetric uncertainties occur in modulated radiotherapy and different methods can be used to estimate these uncertainties.

- It is important that each clinic defines the purpose of their QC method and validates that the method fulfils the purpose, as it has been shown that a recommended measurement-based QC was not always able to detect treatment plans that had dosimetric errors of clinical relevance.
- Complexity metrics that are aperture-based and derived to relate to the uncertainties in the penumbra region showed best agreement with dosimetric differences between measurement and calculation for static fields.
- Complexity scores need to be interpreted differently for different diagnose groups.
- The treatment plan design has impact on the dosimetric uncertainty. Dosimetric uncertainty assessment methods can be valuable information in the process of creating and selecting plan to be used for treatment.

7 FUTURE PERSPECTIVES

Continuing studies on how complexity and robustness vary for treatment plans created for different treatment techniques, treatment machines, and diagnose groups are needed. Benchmarks can then be established to enable interpretation of EAM scores for different groups. EAM has, at this point, been evaluated on treatment plan level as the mean of all control point EAM scores. Further studies are needed to investigate how control point complexity information best could be translated to an assessment of complexity on plan level.

The user experience of EAM is currently under evaluation. All personnel at our department have access to calculate EAM scores for modulated treatment plans during or after treatment planning. This will hopefully lead to further improvements of the metric regarding accessibility and simplicity to use.

The use of ASC in the clinical routine as a method to lower the uncertainty level would be of interest for further investigations.

The theory of EAM was that uncertainties of dose calculation and delivery are largest for the part of the field that was closer to the field edge. This means that larger dosimetric uncertainties are expected in voxels where the dose contribution originates, to a larger extent, from penumbra dose compared to dose contributions originating from parts of the open part of the field. An ongoing work is focused on developing a method based on the theory behind EAM to create a 3D uncertainty map with information on extent and position of probability of uncertainties in the planned dose distribution[76]. In this way, complexity information can be evaluated in more clinically relevant terms.

ACKNOWLEDGEMENT

Jag är så oändligt tacksam för all tid, engagemang, stöttning och uppmuntring från mina fantastiska handledare **Anna Bäck** och **Anna Karlsson**. Det var ni som gjorde att jag vågade ta steget in i forskningsvärlden och ta mig an ett doktorandprojekt. I stunder då det känts som att uppförsbacken varit alldeles för brant så har ni alltid pushat på. Ni är fantastiska förebilder och jag har lärt mig otroligt mycket av er. Mitt största tack är till er!

Det har varit en omtumlande tid att vara doktorand. Ibland är allt upp och ibland är allt ner. Men, jag har alltid påstått att jag har världens bästa jobb. Oftast är det sant… timmarna före en deadline har det varit en smula mindre sant.

Tack till min älskade **Daniel**. Ditt lugn, tålamod, trygghet och sätt att få mig att skratta varje dag är precis vad jag har behövt för att klara av att vara en doktorerande småbarnsmamma. Tack mina fantastiska barn **Holly** och **Seth** för att ni har fått mig på andra tankar och fått helger att handla om utflykter och äventyr i stället för jobb. Jag älskar er!

Tack till mamma **Anita** och pappa **Håkan** för att ni alltid trott på mig och fått mig att tro på mig själv. Tack för att ni har visat vägen och gett mig all fart framåt.

Tack till **Emma Hedin** och **Karin Braide** som med sitt underbara sätt alltid gett mig energi och den uppmuntring som jag behövt.

Tack till **Niclas Pettersson** för din ovärderliga programmeringshjälp, och för att du alltid ställer upp när det behövs… till och med på självaste julafton!

Tack till **Emmanouil Terzidis** och **Louise Mövik** för alla script ni skapat som gjort allt så mycket enklare för mig.

Tack till **Magnus Gustavsson** som har engagerat sig i mitt forskningsarbete genom intressanta diskussioner och viktig input.

Tack till **Ulrika Lindencrona** och **Peter Bernhardt** som har stöttat mig i rollen som doktorand och alltid funnits där för mig när jag behövt det.

Tack till **alla** kollegor både i Göteborg och Borås som har hejat på och stöttat mig i detta. Ert engagemang för att skapa de bästa förutsättningarna för patienterna har motiverat mig att arbeta för att ytterligare kunna höja kvalitén av den strålbehandling som vi kan erbjuda.

Och självklart ett stort tack till alla lunchvänner i ”mattåget” som låtit mig hålla låda varje dag.

Jag vill också rikta ett stort tack till **Strålsäkerhetsmyndigheten, Jubileumsklinikens cancerfond** och **Regionala FoU-medel Västra Götalandsregionen** som med sin finansiering gjort detta arbete möjligt.

Tack till min **opponent**, **ordförande** och **betygskommitté** för att ni tagit er tid att bedöma mitt arbete och hålla i disputationsakten.

REFERENCES

1. Cancerfonden. https://www.cancerfonden.se/om-cancer/behandlingar/stralbehandling. Accessed Sept. 2022.
2. SSM S. Strålsäkerhetens föreskrifter om medicinska exponeringar. *SSMFS.* 2018;5.
3. Brahme A. Optimization of stationary and moving beam radiation therapy techniques. *Radiother Oncol.* 1988;12(2):129-140.
4. Webb S. Optimisation of conformal radiotherapy dose distributions by simulated annealing. *Physics in medicine and biology.* 1989;34(10):1349-1370.
5. Arthur L. Boyer PD, E. Brian Butler, M.D., Thomas A. DiPetrillo, M.D., Mark J. Engler, Ph.D., Benedick Fraass, Ph.D., Walter Grant, III, Ph.D., C. Clifton Ling, Ph.D., Daniel A. Low, Ph.D., Thomas R. Mackie, Ph.D., Radhe Mohan, Ph.D., James A. Intensity-modulated radiotherapy: current status and issues of interest. *International journal of radiation oncology, biology, physics.* 2001;51(4):880-914.
6. Ezzell GA, Galvin JM, Low D, et al. Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee. *Med Phys.* 2003;30(8):2089-2115.
7. Yu CX. Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy. *Physics in medicine and biology.* 1995;40(9):1435-1449.
8. Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys.* 2008;35(1):310-317.
9. Varian Medical Systems. Eclipse Photon and Electron Algorithms 15.5 Reference Guide. 2017(P1020505-002-B).
10. Binny D, Spalding M, Crowe SB, et al. Investigating the use of aperture shape controller in VMAT treatment deliveries. *Medical dosimetry : official journal of the American Association of Medical Dosimetrists.* 2020;45(3):284–292.
11. Scaggion A, Fusella M, Agnello G, et al. Limiting treatment plan complexity by applying a novel commercial tool. *Journal of applied clinical medical physics.* 2020;21(8):27-34.
12. Craft D, McQuaid D, Wala J, Chen W, Salari E, Bortfeld T. Multicriteria VMAT optimization. *Med Phys.* 2012;39(2):686-696.
13. Miguel-Chumacero E, Currie G, Johnston A, Currie S. Effectiveness of Multi-Criteria Optimization-based Trade-Off exploration in combination with RapidPlan for head & neck radiotherapy planning. *Radiat Oncol.* 2018;13(1):229.
14. Spalding MB, Walsh AM, Clarke HB, Aland TM. Evaluation of a new hybrid VMAT-IMRT multi-criteria optimization plan generation

algorithm. *Medical dosimetry : official journal of the American Association of Medical Dosimetrists.* 2019.
15. Mohan R, Arnfield M, Tong S, Wu Q, Siebers J. The impact of fluctuations in intensity patterns on the number of monitor units and the quality and accuracy of intensity modulated radiotherapy. *Med Phys.* 2000;27(6):1226-1237.
16. Glenn MC, Hernandez V, Saez J, et al. Treatment plan complexity does not predict IROC Houston anthropomorphic head and neck phantom performance. *Physics in medicine and biology.* 2018;63(20):205015.
17. Hernandez V, Hansen CR, Widesott L, et al. What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans. *Radiother Oncol.* 2020;153:26-33.
18. Shen L, Chen S, Zhu X, et al. Multidimensional correlation among plan complexity, quality and deliverability parameters for volumetric-modulated arc therapy using canonical correlation analysis. *Journal of radiation research.* 2018;59(2):207-215.
19. Nguyen M, Chan GH. Quantified VMAT plan complexity in relation to measurement-based quality assurance results. *Journal of applied clinical medical physics.* 2020;21(11):132-140.
20. Das IJ, Ding GX, Ahnesjo A. Small fields: nonequilibrium radiation dosimetry. *Med Phys.* 2008;35(1):206-215.
21. LoSasso T, Chui CS, Ling CC. Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy. *Med Phys.* 1998;25(10):1919-1927.
22. Wang J, Jin X, Peng J, Xie J, Chen J, Hu W. Are simple IMRT beams more robust against MLC error? Exploring the impact of MLC errors on planar quality assurance and plan quality for different complexity beams. *Journal of applied clinical medical physics.* 2016;17(3):147-157.
23. Calcina CS, de Oliveira LN, de Almeida CE, de Almeida A. Dosimetric parameters for small field sizes using Fricke xylenol gel, thermoluminescent and film dosimeters, and an ionization chamber. *Physics in medicine and biology.* 2007;52(5):1431-1439.
24. Das IJ, Francescon P, Moran JM, et al. Report of AAPM Task Group 155: Megavoltage photon beam dosimetry in small fields and non-equilibrium conditions. *Medical Physics.* 2021;48(10):e886-e921.
25. Park JM, Wu HG, Kim JH, Carlson JN, Kim K. The effect of MLC speed and acceleration on the plan delivery accuracy of VMAT. *The British journal of radiology.* 2015;88(1049):20140698.
26. Court L, Wagar M, Berbeco R, et al. Evaluation of the interplay effect when using RapidArc to treat targets moving in the craniocaudal or right-left direction. *Med Phys.* 2010;37(1):4-11.

27. Miften M, Olch A, Mihailidis D, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218. *Med Phys.* 2018;45(4):e53-e83.
28. Wal E, Wiersma J, Ausma AH, et al. *Code of Practice for the Quality Assurance and Control for Intensity Modulated Radiotherapy Report 22 of the Netherlands Commission on Radiation Dosimetry.* 2013.
29. (ESTRO) ESfRao. Guidelines for the Verification of IMRT. *Brussels: ESTRO.* 2008.
30. ICRU83. *Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT).* Journal of the ICRU2010.
31. (SSRMP) SSfRaMP. Quality control for Intensity modulated radiation therapy. *http://wwwsgsmpch/.* 2007;Recommendations No. 15.
32. Hartford AC, Palisca MG, Eichler TJ, et al. American Society for Therapeutic Radiology and Oncology (ASTRO) and American College of Radiology (ACR) Practice Guidelines for Intensity-Modulated Radiation Therapy (IMRT). *International journal of radiation oncology, biology, physics.* 2009;73(1):9-14.
33. Low DA, Moran JM, Dempsey JF, Dong L, Oldham M. Dosimetry tools and techniques for IMRT. *Med Phys.* 2011;38(3):1313-1338.
34. Ling CC, Zhang P, Archambault Y, Bocanek J, Tang G, Losasso T. Commissioning and quality assurance of RapidArc radiotherapy delivery system. *International journal of radiation oncology, biology, physics.* 2008;72(2):575-581.
35. LoSasso T, Chui CS, Ling CC. Comprehensive quality assurance for the delivery of intensity modulated radiotherapy with a multileaf collimator used in the dynamic mode. *Med Phys.* 2001;28(11):2209-2219.
36. McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys.* 2010;37(2):505-515.
37. Younge KC, Matuszak MM, Moran JM, McShan DL, Fraass BA, Roberts DA. Penalization of aperture complexity in inversely planned volumetric modulated arc therapy. *Med Phys.* 2012;39(11):7160-7170.
38. Chiavassa S, Bessieres I, Edouard M, Mathot M, Moignier A. Complexity metrics for IMRT and VMAT plans: a review of current literature and applications. *The British journal of radiology.* 2019;92(1102):20190270.
39. Antoine M, Ralite F, Soustiel C, et al. Use of metrics to quantify IMRT and VMAT treatment plan complexity: A systematic review and perspectives. *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology :*

official journal of the Italian Association of Biomedical Physics (AIFB). 2019;64:98-108.
40. Crowe SB, Kairn T, Kenny J, et al. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. *Australasian physical & engineering sciences in medicine.* 2014;37(3):475-482.
41. Agnew CE, Irvine DM, McGarry CK. Correlation of phantom-based and log file patient-specific QA with complexity scores for VMAT. *Journal of applied clinical medical physics.* 2014;15(6):4994.
42. McGarry CK, Agnew CE, Hussein M, et al. The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. *The British journal of radiology.* 2016;89(1057):20150445.
43. McGarry CK, Chinneck CD, O'Toole MM, O'Sullivan JM, Prise KM, Hounsell AR. Assessing software upgrades, plan properties and patient geometry using intensity modulated radiation therapy (IMRT) complexity metrics. *Med Phys.* 2011;38(4):2027-2034.
44. Park JM, Kim JI, Park SY. Modulation indices and plan delivery accuracy of volumetric modulated arc therapy. *Journal of applied clinical medical physics.* 2019;20(6):12-22.
45. Park JM, Park S-Y, Kim H, Kim J, Carlson J, Ye S-J. Modulation indices for volumetric modulated arc therapy. *Physics in medicine and biology.* 2014;59:7315-7340.
46. Olofsson J, Gustavsson M, Isacsson U, Olevik-Dunder M, Westermark M, Benedek H, Hållström P. Strategier vid kvalitetssäkring av intensitetsmodulerad strålbehandling. *SVENSK FÖRENING FÖR RADIOFYSIK.* 2014;RAPPORT 2014:1.
47. Bossuyt E, Weytjens R, Nevens D, De Vos S, Verellen D. Evaluation of automated pre-treatment and transit in-vivo dosimetry in radiotherapy using empirically determined parameters. *Physics and Imaging in Radiation Oncology.* 2020;16:113-129.
48. SunNuclear. *SunCHECK Patient Reference Guide.* 2020.
49. Nelms BE, Zhen H, Tome WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys.* 2011;38(2):1037-1044.
50. Kruse JJ. On the insensitivity of single field planar dosimetry to IMRT inaccuracies. *Med Phys.* 2010;37(6):2516-2524.
51. Crijns W, Defraene G, Van Herck H, et al. Online adaptation and verification of VMAT. *Med Phys.* 2015;42(7):3877-3891.
52. Crowe SB, Kairn T, Middlebrook N, et al. Examination of the properties of IMRT and VMAT beams and evaluation against pre-treatment quality assurance results. *Physics in medicine and biology.* 2015;60(6):2587-2601.
53. Parwaie W, Refahi S, Ardekani MA, Farhood B. Different Dosimeters/Detectors Used in Small-Field Dosimetry: Pros and Cons. *J Med Signals Sens.* 2018;8(3):195-203.

54. Wong CJ, Ackerly T, He C, et al. Small field size dose-profile measurements using gel dosimeters, gafchromic films and micro-thermoluminescent dosimeters. *Radiation Measurements.* 2009;44(3):249-256.
55. Zhu Y, Kirov AS, Mishra V, Meigooni AS, Williamson JF. Quantitative evaluation of radiochromic film response for two-dimensional dosimetry. *Med Phys.* 1997;24(2):223-231.
56. Bedford JL, Lee YK, Wai P, South CP, Warrington AP. Evaluation of the Delta4 phantom for IMRT and VMAT verification. *Physics in medicine and biology.* 2009;54(9):N167-176.
57. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys.* 1998;25(5):656-661.
58. Steers JM, Fraass BA. IMRT QA: Selecting gamma criteria based on error detection sensitivity. *Med Phys.* 2016;43(4):1982.
59. Nelms B, Jarry G, Chan M, Hampton C, Watanabe Y, Feygelman V. Real-world examples of sensitivity failures of the 3%/3mm pass rate metric and published action levels when used in IMRT/VMAT system commissioning. *Journal of Physics Conference Series.* 2013;444:2086.
60. Fredh A, Scherman JB, Fog LS, Munck af Rosenschöld P. Patient QA systems for rotational radiation therapy: a comparative experimental study with intentional errors. *Med Phys.* 2013;40(3):031716.
61. Heilemann G, Poppe B, Laub W. On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance. *Med Phys.* 2013;40(3):031702.
62. Zhen H, Nelms BE, Tome WA. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. *Med Phys.* 2011;38(10):5477-5489.
63. Nilsson J, Karlsson A, Bäck A. IMRT patient-specific QA using the Delta4 dosimetry system and evaluation based on ICRU 83 recommendations. *7th International Conference on 3D Radiation Dosimetry.* 2013;IC3DDose 2012.
64. Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V, Gabriele P. Pretreatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram. *Med Phys.* 2012;39(12):7626-7634.
65. Nelms BE, Chan MF, Jarry G, et al. Evaluating IMRT and VMAT dose accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels. *Med Phys.* 2013;40(11):111722.
66. Webb S. Use of a quantitative index of beam modulation to characterize dose conformality: illustration by a comparison of full beamlet IMRT, few-segment IMRT (fsIMRT) and conformal

unmodulated radiotherapy. *Physics in medicine and biology.* 2003;48(14):2051-2062.
67. Du W, Cho SH, Zhang X, Hoffman KE, Kudchadker RJ. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. *Med Phys.* 2014;41(2):021716.
68. Masi L, Doro R, Favuzza V, Cipressi S, Livi L. Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy. *Med Phys.* 2013;40(7):071718.
69. Nauta M, Villarreal-Barajas JE, Tambasco M. Fractal analysis for assessing the level of modulation of IMRT fields. *Med Phys.* 2011;38(10):5385-5393.
70. Younge KC, Roberts D, Janes LA, Anderson C, Moran JM, Matuszak MM. Predicting deliverability of volumetric-modulated arc therapy (VMAT) plans using aperture complexity analysis. *Journal of applied clinical medical physics.* 2016;17(4):124-131.
71. Hernandez V, Saez J, Pasler M, Jurado-Bruggeman D, Jornet N. Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy. *Physics and Imaging in Radiation Oncology.* 2018;5:37-43.
72. Kairn T, Papworth D, Crowe SB, Anderson J, Christie DR. Dosimetric quality, accuracy, and deliverability of modulated radiotherapy treatments for spinal metastases. *Medical dosimetry : official journal of the American Association of Medical Dosimetrists.* 2016;41(3):258-266.
73. Lee MT, Purdie TG, Eccles CL, Sharpe MB, Dawson LA. Comparison of simple and complex liver intensity modulated radiotherapy. *Radiation Oncology.* 2010;5(1):115.
74. Craft DPD, Süss P, Bortfeld TPD. The Tradeoff Between Treatment Plan Quality and Required Number of Monitor Units in Intensity-modulated Radiotherapy. *International journal of radiation oncology, biology, physics.* 2007;67(5):1596-1605.
75. Jurado-Bruggeman D, Hernandez V, Saez J, et al. Multi-centre audit of VMAT planning and pre-treatment verification. *Radiother Oncol.* 2017;124(2):302-310.
76. Bäck A, Nordström F, Gustafsson M, Götstedt J, Karlsson A. Complexity metric based on fraction of penumbra dose - initial study. *Journal of Physics: Conference Series.* 2017;847:012002.