Universidade de Aveiro Departamento de Física
Ano 2020

Tiago Emanuel Pereira Ventura

Avaliação de planos em radioterapia – a caminho do planeamento automatizado

Plan evaluation in radiotherapy – on the way to automated treatment planning

Universidade de Aveiro Departamento de Física
Ano 2020

Tiago Emanuel Pereira Ventura

Avaliação de planos em radioterapia – a caminho do planeamento automatizado

Plan evaluation in radiotherapy – on the way to automated treatment planning

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Física, realizada sob a orientação científica da Professora Doutora Maria do Carmo Carrilho Calado Antunes Lopes, Professora Associada Convidada do Departamento de Física da Universidade de Aveiro e a co-orientação da Professora Doutora Brígida da Costa Ferreira, Professora Adjunta da Escola Superior de Saúde do Instituto Politécnico do Porto

FCT Fundação para a Ciência e a Tecnologia

Apoio financeiro do projeto da FCT "RT- Computational Approaches for Radiotherapy Planning of Excellence" (POCI-01-0145-FEDER-028030) numa parceria entre FEUC/INESCC e o IPOCFG, E.P.E.

Para a minha querida família, Inês e Francisco.

o júri

presidente

Prof. Doutor António Manuel Pereira
professor catedrático da Universidade de Aveiro

Prof. Doutor Ben J. M. Heijmen
professor catedrático no Erasmus MC – Erasmus University Rotterdam

Prof.ª Doutora Sandra Correia Vieira
especialista em física médica do Departamento de Radioterapia da Fundação Champalimaud

Prof. Doutor Eduardo Martins Carlinhos Netto
radioncologista do Serviço de Radioterapia do Instituto Português de Oncologia de Lisboa Francisco Gentil, E.P.E.

Prof.ª Doutora Maria Isabel Silva Ferreira Lopes
professora catedrática do Departamento de Física da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Prof.ª Doutora Maria do Carmo Carrilho Calado Antunes Lopes
professora associada convidada do Departamento de Física da Universidade de Aveiro

Como refere o grande poeta espanhol António Machado «se hace camino al andar». Sim, o caminho faz-se caminhando… passo a passo… degrau a degrau. Estes percursos e processos são tão mais enriquecedores e desafiantes quanto a diversidade de pessoas que conhecemos e que altruistamente se disponibilizam a acompanhar-nos no decurso da caminhada e descoberta, feita muitas vezes de (des)motivações, avanços e inflexões. Essas pessoas fazem a diferença na nossa vida, marcam e preenchem de forma indelével o nosso quotidiano e o nosso coração.
Neste sentido, considero-me um privilegiado. Tenho a sorte e a prerrogativa de conhecer e de trabalhar com pessoas excecionais e trabalhar numa área que me faz sentir útil, por estar ao serviço ao próximo. Tenho a sorte e o privilégio do meu trabalho conseguir preencher o meu desejo insaciável de querer saber sempre mais… porque sempre mais, raramente chega.

Agradecimentos

Ao longo de todo este processo de aprendizagem e de investigação contei com o apoio inestimável de várias pessoas que quero aqui mencionar e endereçar o meu apreço e gratidão. Em primeiro lugar um agradecimento muito especial à Doutora Maria do Carmo Lopes, minha orientadora e diretora de serviço. Obrigado pelos prestimosos conselhos, incentivos, questionamentos e reflexões, pelo inexcedível apoio e amizade ao longo de todo o projeto doutoral e também por toda a confiança que sempre senti em mim depositada. Agradeço igualmente o privilégio de poder partilhar no quotidiano da radioterapia da sua experiência e visão. O meu sincero e profundo agradecimento à Doutora Brígida da Costa Ferreira, minha co-orientadora, por todos os conselhos, apoio e amizade. Foi um prazer muito grande poder trabalhar com alguém tão talentoso e que tanto me ensinou e inspirou.
Não posso deixar de agradecer também à Doutora Joana Dias e ao Doutor Humberto Rocha, membros do grupo de investigação do INESC Coimbra, por toda disponibilidade, apoio e dedicação, cuja partilha de conhecimentos e experiências contribuíram de forma decisiva para o resultado final deste projeto. A eles o meu muito obrigado.
Um reconhecido agradecimento ao conselho de administração do Instituto Português de Oncologia de Coimbra Francisco Gentil, E.P.E. por me ter proporcionado a oportunidade e as condições para a realização deste trabalho de investigação.
À Josefina, ao Miguel, à Carla, à Carmen e à Tânia, com quem partilho o dia-a-dia, muito obrigado pela amizade, apoio e compreensão. Sem eles a conclusão deste meu grande desejo não teria sido possível. Um agradecimento muito especial também à Ana, à Fernanda, à Liliana, à Mary e ao João, pela compreensão e amizade sempre demonstrados. O meu reconhecimento à Dra. Leila Khouri e à Dra. Tânia Serra, médicas radioncologistas, com quem trabalho mais diretamente numa verdadeira parceria entre físico e médico. Obrigado pela amizade e por todos os conselhos clínicos necessários e imprescindíveis para a conceção, desenvolvimento e conclusão deste projeto.
Ao Doutor Sebastiaan Breedveld e ao Doutor Ben Heijmen pela possibilidade de utilizar o sistema de planeamento de IMRT multicritério Erasmus-iCycle e por todos os conselhos imprescindíveis ao desenvolvimento deste projeto.
Ao Doutor Eduardo Netto (Instituto Português de Oncologia de Lisboa Francisco Gentil, E.P.E.) e ao Dr. André Santos (Instituto Português de Oncologia do Porto Francisco Gentil, E.P.E), o meu agradecimento por terem aceitado participar neste projeto doutoral com a sua disponibilidade e experiência clínica.
Ao Paulo Rachinhas o meu obrigado, por todo o apoio informático dispensado ao longo deste projeto. À professora Andreia Hall um profundo agradecimento pela disponibilidade, apoio e suporte científico na análise estatística.
Aos meus amigos Pedro, Nuno e Tânia por toda a amizade e pela ajuda no design gráfico da tese.

Por fim, um agradecimento especial aos meus pais e ao meu irmão por todo o apoio, compreensão, pelo exemplo de vida e de perseverança, que sempre me guiaram e inspiraram.
À Inês e ao Francisco obrigado por todo o amor, compreensão e dedicação. Obrigado pela cor e alegria que dão ao meu dia-a-dia. Juntos somos a equipa TIF.

Muito Obrigado

palavras-chave

avaliação de planeamentos, suporte clínico à seleção de planeamentos, automação, otimização angular, radioterapia, física médica.

resumo

A Radioterapia é uma modalidade terapêutica multidisciplinar de elevada complexidade e exatidão que utiliza, de forma controlada, radiação ionizante para eliminar as células tumorais poupando o máximo possível os tecidos normais. O doente é tratado ao longo de várias sessões de acordo com um plano de tratamento personalizado otimizado, *a priori*, por um planeador num sistema de planeamento. As novas técnicas de planeamento, como o planeamento inverso com campos estáticos ou com arcos volumétricos, requerem a utilização de algoritmos de otimização de dose que, de forma iterativa, permitem ao planeador encontrar uma ou mais soluções que satisfaçam o melhor possível a prescrição de dose e cumpram os limites de dose de tolerância definidos para os órgãos de risco. A automatização do processo de planeamento utilizando algoritmos de otimização multicritério permite que todos os doentes tenham planos de tratamento com qualidade equivalente e reduzir drasticamente a dependência do processo de otimização da destreza do planeador. A qualidade das distribuições de dose pode ainda ser melhorada se durante a otimização da intensidade dos feixes de radiação forem incorporados algoritmos de otimização das direções e trajetórias dos feixes de tratamento. Este trabalho tem como principais propósitos contribuir para o processo de automatização do planeamento em Radioterapia através do desenvolvimento de novas ferramentas para avaliação de planos e de novos algoritmos de otimização de direções e trajetórias dos feixes de tratamento. Neste sentido foi desenvolvida uma ferramenta gráfica de avaliação e comparação de planos de apoio à decisão do radioncologista designada SPIDERplan. Para além da informação dosimétrica do plano, esta ferramenta incorpora na sua metodologia as preferências clínicas dos radioncologistas e permite avaliar e comparar de forma independente planos otimizados com diferentes algoritmos, diferentes técnicas ou provenientes de diferentes sistemas de planeamento. A SPIDERplan foi validada clinicamente para os carcinomas de nasofaringe por três médicos radioncologistas dos três institutos portugueses de oncologia de Lisboa, Porto e Coimbra. Os resultados obtidos provaram que a avaliação dos planos por esta realizada é comparável à avaliação clínica dos médicos pelo que a sua utilização foi alargada a outros estudos e patologias. Assim, a SPIDERplan foi utilizada na comparação de dois algoritmos de otimização angular baseados na otimização de fluências para geometrias coplanares e não-coplanares num sistema de otimização multicritério. A investigação das possíveis vantagens da otimização angular foi alargada aos tumores de sistema nervoso central, mais concretamente aos meningiomas. Neste caso, para além da avaliação da qualidade dos planos, a SPIDERplan foi também utilizada para guiar o algoritmo de otimização angular para geometrias não-coplanares. Ainda no estudo dos meningiomas foi proposto um novo algoritmo de otimização de trajetórias de feixes em arco baseado no conceito de pontos de ancoragem, cuja otimização foi também guiada pela SPIDERplan. Quer para o estudo da nasofaringe quer para os meningiomas, a SPIDERplan foi usada para avaliar a qualidade dos planos gerados de um ponto de vista global, avaliando a qualidade média de todos os casos. Em casos chave foi usada, mais especificamente, para evidenciar as vantagens que podem advir para a otimização dos planeamentos em Radioterapia com os algoritmos de otimização de direções e trajetórias dos feixes de radiação.

keywords

plan assessment, clinical suport to plan selection, automation, beam angular optimization, radiation therapy, medical physics.

abstract

Radiation therapy makes use of ionization radiation to eliminate in a controlled way tumours cells sparing as much as possible the normal tissues. The treatment is delivered to the patient during several sessions according to a personalized plan optimized a priori by a planner in a treatment plan system. The most advanced planning techniques, such as inverse planning optimization of static beams or volumetric arcs, require the use of dose optimization algorithms, that allow the planner to find one or more solutions that satisfies as much as possible the prescription dose for the target and the tolerance dose criteria for the normal tissues. Treatment plan automation based on multicriteria optimization algorithms generates treatment planning solutions with equivalent dose distribution quality and reduced the interaction from the planner. The quality of the dose distribution can even be substantially enhanced when beam angle optimization and arc trajectory optimization algorithms are incorporated in plan optimization. The main purpose of this work is to contribute to the treatment planning automation process by developing new plan assessment tools and use them for testing new directions/trajectory optimization algorithms. To support clinical decision-making, a graphical method incorporating the clinical aims of the radiation oncologist was developed to evaluate and compare treatment plans independently from the algorithm, the treatment technique or the treatment planning system used. This plan quality assessment tool, named SPIDERplan, was clinically validated for the nasopharynx pathology by three radiation oncologists from the three Portuguese Oncology Institutes in Lisbon, Porto and Coimbra. The performance of SPIDERplan proved to be comparable with the radiation oncologists' evaluations. Its configuration and resulting scoring were discussed enabling its generalized application. SPIDERplan was then extensively used. Firstly, for nasopharynx tumour cases, in the comparison of two-fluence based beam angle optimization algorithms for coplanar and non-coplanar geometries in a multicriterial optimization framework. The direction optimization topic was afterwards extended to intracranial tumours, namely to meningioma cases. SPIDERplan was used, this time, not just to assess plans quality but also to guide the non-coplanar beam angle optimization algorithm. Furthermore, a new arc trajectory optimization algorithm based in the anchor point concept was proposed, driven again by SPIDERplan. In both nasopharynx and meningioma pathologies, SPIDERplan was used to assess the plans quality, averaged over the used patient samples. In addition, its application to specific-patient situations enabled to evidence the advantages of the optimization of direction/trajectory in a very impressive way.

table of contents

list of figures

Chapter 5

Chapter 6

Appendix 1

list of tables

list of abbreviations

ATO	arc trajectory optimization
BAO	beam angle optimization
CN	conformation number
CI_{DD}	conformity index incorporating dose and distance
COIN	conformal index
COSI	critical organ scoring index
CPQI	composite quality index
CT	computed tomography
DVH	dose-volume histogram
EUD	equivalent uniform dose
FMO	fluence map optimization
gEUD	generalized equivalent uniform dose
GI	gradient index
GPD	group plan diagram
IGRT	Image-guided radiation therapy
IMRT	Intensity-modulated radiation therapy
K_{B-P}	Brennan-Prediger kappa coefficient
LTCP	logarithmic tumour control probability
MCO	multicriteria optimization
MLC	multi-leaf collimator
mMLC	micro multi-leaf collimator
MPLM	mixed linear programming model
MRI	magnetic resonance imaging
NPC	nasopharynx cancer
NTCP	normal tissue complication probability
OAR	organs-at-risk
PCI	plan conformity index
PET	positron emission tomography
PQI	plan quality index
PTV	planning target volume
RO	radiation oncologist

RT	radiation therapy
RTOG	radiation therapy oncology group
SGD	structures group diagram
SPD	structures plan diagram
TCI^{+}	uncomplicated target conformity index
TCP	tumour control probability
TPS	treatment planning system
VMAT	volumetric modulated arc therapy
3DCRT	3D conformal radiation therapy

chapter 1

Introduction

1.1. An overview of cancer care and radiation therapy

Cancer encompasses the group of diseases characterized by the uncontrolled growth and division of abnormal cells that may have the ability to infiltrate adjacent tissues and/or to spread throughout the body. According to the World Health Organization, cancer is the second leading cause of death in the world after cardiovascular diseases.[1] In 2018, 18.1 million new cases and 9.6 million cancer deaths were registered worldwide. In 2030, 24.1 million new cases and 13 million deaths are expected.[2] The disease burden is higher in the low and middle-income countries where around 70% of the deaths caused by cancer were reported.[1] In Europe, cancer is responsible for 24% of the mortality with 4.2 million new cases and 1.9 million deaths in 2018.[2] The most common cancers are lung, female breast and colorectal, and they are responsible by one-third of the cancer incidence and mortality.[1] Cancer treatment involves a multidisciplinary approach, where surgery, radiation therapy and systemic therapy (e.g. chemotherapy, hormonal therapy, immunotherapy and target therapy) come as the possible treatment modalities.

Radiation therapy is a fundamental component of comprehensive cancer care management programs. Used alone or in combination with other treatment modalities, it provides high local and local-regional tumour control and plays a crucial role in palliative treatments. Comprehensive epidemiological evidence-based studies have estimated that around 50% of the new cancer cases and 25% of the retreatments do receive, in any given moment of the course of the treatment, radiation therapy either for cure or palliation.[3-5] For cured patients, i.e. patients surviving 5 years after the diagnosis, it is estimated that approximately 40% were treated with radiation therapy.[6]

Recent reports indicate that radiation therapy has been used in a suboptimal way.[5,7] In Europe, a large and heterogeneous gap between the actual daily applied radiation therapy and its optimal utilization was verified. On average only 71% of the patients with indication for radiation therapy got access to this treatment.[5] Worldwide the situation is even more dramatic. In low-middle income countries, 50% of the patients do not have access to radiation therapy and in low-income countries this number increases to 90%.[7] The need of radiation therapy should be definitely placed in the national health policy agenda. Issues such as improved referral patterns and reimbursement policies, better accessibilities to radiation therapy facilities, new and improved patient's accommodations and increase in the investment funds to install new equipment and set up new facilities should be covered. For the low- and middle-income countries, beyond the sustainable investment in infrastructures and equipment, the cancer care and control programs should also include the recruitment and training of highly specialised multi-professional staff and the necessary efforts to retain a suitable health workforce.[7,8]

The development of new radiation therapy platforms using digital data and cloud-based data management, such as automated treatment planning and quality control, teledosimetry or peer review support, can play a crucial role in the shortening of the technological gap in the medium term.[8]

1.2. Radiation therapy framework

Radiation therapy makes use of ionizing radiation to eliminate tumour cells in a controlled way, minimizing as much as possible normal tissue damage. The ionizing radiation will impart its energy, directly or indirectly, to the volume of irradiated tissue inducing DNA damages, e.g. deletions or replacements of the basis and/or breaks in the chain that will originate cell death when repair processes are not effective.

There are two types of radiation therapy: external beam radiation therapy, where the source of radiation is located outside the patient, and brachytherapy, where the source of radiation is

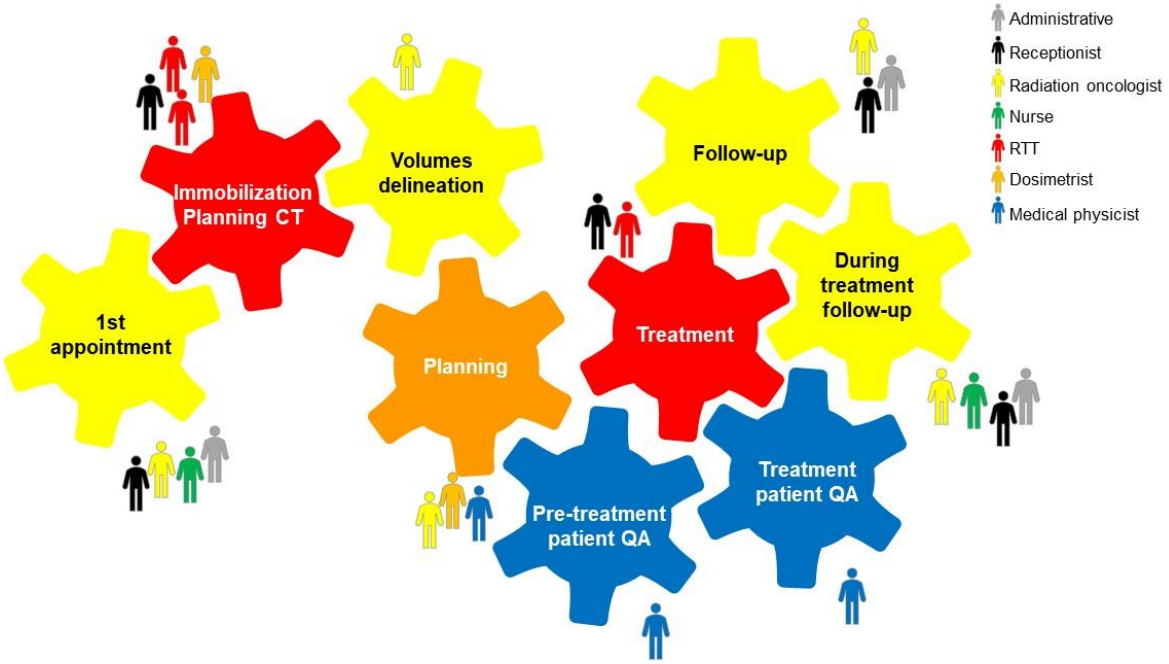

Figure 1.1 - Radiation therapy treatment phases implemented in Portuguese Oncology Institute of Coimbra.

implanted inside or placed as close as possible to the tumour. The latter is beyond the scope of this work.

In external beam radiation therapy, the radiation beam is usually generated by a linear accelerator mounted on a gantry that rotates around the patient that is immobilized on a movable couch. Several treatment techniques can be used for treatment delivery. The most common in clinical routine are: 3D conformal radiotherapy (3DCRT), intensity-modulated radiotherapy (IMRT), volumetric modulated arc therapy (VMAT) and helical IMRT. For 3DCRT treatments, uniform beams geometrically shaped by a multi-leaf collimator (MLC) irradiating from multiple directions are used. If static or dynamic beams with non-uniform intensities are used instead, the treatment delivery technique is named IMRT. For VMAT, non-uniform intensity modulated fields are delivered while the gantry rotates around the patient. Recently, technological improvements in linear accelerators enabled the simultaneous rotation of the couch and of the gantry during irradiation.[9,10] In helical IMRT, performed in a dedicated treatment unit named Tomotherapy, non-uniform modulated beams produced by a linear accelerator mounted in O-arm gantry rotates with constant velocity around the patient, while the couch continuously moves inwards.

Depending on the tumour type and on the delivery technique, a standard fraction dose, of 1.8 or 2 Gy, is administered to the patient on a daily basis during five to seven weeks. For hypofractionated treatments, higher doses per fraction (>2 Gy) are prescribed and the treatment is delivered in a single fraction, daily or in alternated days during one or two weeks.

External beam radiation therapy, hereafter designated just by radiation therapy, is a complex and patient individualized process that involves a dedicated multidisciplinary team performing a set of related steps, Figure 1.1. The radiation therapy treatment process starts, after the therapeutic decision has been taken (preferably by a multidisciplinary board), with a clinical evaluation of the patient by the radiation oncologist in the first radiation therapy medical appointment. Next, the immobilisation of the patient is defined along with any necessary accessories. A computerized tomography (CT) scan with the patient in the treatment position is performed and the structures of interest, planning target volume (PTV) and organs-at-risk (OAR), are delineated on the acquired

images. Other image sets of interest, such as magnetic resonance imaging (MRI) or positron emission tomography (PET) images, can also be acquired and fused with the reference planning CT images to enhance the accuracy of structures delineation. A treatment plan, based on the structures outlined and on the dose prescription to the PTV and on the tolerance dose criteria to the OARs, is then created using a treatment planning system (TPS). To guarantee that the generated treatment plan will be delivered to the patient as planned, pre-treatment patient quality assurance is usually done. Each treatment session/fraction intends to deliver to the patient the approved treatment plan. Before each fraction, portal or cone-beam CT images are acquired and compared with the reference CT images. This procedure, called image-guided radiation therapy (IGRT) will detect and desirably correct any deviations from the planned setup. Deviations may occur due to wrong patient positioning on the couch and/or to anatomical changes (e.g. different bladder filling). During and after the course of radiation therapy, follow-up appointments with the radiation oncologist team evaluate the patient's condition, the evolution and the outcome of the treatment.

1.3. Treatment planning

In the treatment planning phase, a plan able to accomplish the delivery of the prescribed dose to the PTV keeping the dose in the OARs at least below the tolerance doses is sought by the planner. The plan and the corresponding dose distribution optimization are performed in a commercial TPS using either a forward or an inverse planning approach. In forward planning, used in 3DCRT treatments, the number of beams, the beam angle orientations (gantry, couch and collimator angles), MLC shapes, the beam modifiers (dynamic or physical wedges) and the weight of each beam must be manually defined by the planner prior to the dose distribution calculation. Uniform or uni-directionally modulated (when modified by wedges) fields are generated. If the plan does not meet the expectations or does not accomplish the radiation oncologist prescription, a manual tune of the plan parameters is iteratively done in a trial and error process until an acceptable solution is found. In inverse planning optimization, non-uniform intensity fields are used to generate highly conformal dose distributions to the target volume. The planner starts by specifying the number of beams and corresponding orientations, in the case of static or dynamic IMRT treatments, or the number of arcs and the corresponding amplitude and trajectory for VMAT treatments. The plan objectives are described by weighted descriptors that are incorporated in a single objective function that will guide the optimization of the beam intensities by scoring the goodness of the plan.[11] An iterative manual adjustment of plan parameters (e.g. objectives weights, number of beams or arcs, beams or arc directions, number of segments) may be required to reach an acceptable plan. The final plan solution will be dependent of the planner skills and experience and of the case complexity. A common problem of this iterative process is that due to the interdependencies that exist among the adjustable parameters, the impact on the dose distribution of changing one parameter is often unpredictable. Furthermore, the objective function value has only a mathematical meaning that cannot be linked with the clinical outcome. As a result, it is not possible to guarantee that the plan selected to be approved by the radiation oncologist is the optimal solution.

More recently, multicriteria optimization methods, requiring much less manual planner interaction, have been proposed.[12-15] Instead of a single objective function, usually applied in the standard inverse planning approach, multiple objectives are simultaneously minimized (or maximized). The planner, who is named decision-maker in the multicriteria context, can select a posteriori the desired solution from a set of Pareto-optimal treatment plans or a priori establishing constraints and objectives that have to be met in order to generate a single Pareto-optimal solution.

At the end of the optimization, the planner submits its proposals to the appraisal of the radiation oncologist. In most of the TPSs, the assessment of the quality of the dose distribution is made by visually inspecting the isodose lines displayed on the CT slices seeking for hot spots, PTV coverage, dose gradients, and low dose dispersion in the OARs and in the normal tissues. Clinical plan

assessment is completed by the evaluation of the dose-volume histograms (DVH) and the corresponding dose statistics. If this individual analysis fulfils the dose prescription to the PTV and the tolerance dose criteria of the OARs the plan is approved.

When several good plans are presented by the planner to the radiation oncologist, there is an inherent subjectivity in the choice of the treatment plan. Even if one of them is optimal from the mathematical point of view, clinically it is impossible to predict which is the best plan for each patient. This topic would lead to considerations that are out of the scope of this thesis, as for instance, the individual patient radiosensivity or other patient specificities. Anyway, the radiation oncologist decision-making is driven by a group of technical, social and personal factors that will influence plan selection.[16] Due to the large number of dosimetric parameters to be evaluated, the traditional plan assessment tools described above, do not allow an efficient, objective and consistent plan evaluation. Thus, the decision-making process is often characterized by some intra- and inter-variability among radiation oncologists.

1.4. Scope and driving line of the thesis

Treatment planning is a highly demanding patient-specific procedure in which the quality of the generated plan solution is very dependent of the case complexity (e.g. shape and size of the tumour, critical structures proximity, prescription dose, etc.), the TPS features/algorithms and of the skills and experience of the planner. To reduce the variability of the output solutions, it is needed to grant higher consistency and quality to the planning process.[17,18] Benefiting from important hardware developments and enhancements on parallel computation methods, automated treatment planning approaches, with little or no intervention from the planner, have been progressively implemented within commercial systems.[17] The generalization of these options in the clinical routine will allow the generation of treatment plan solutions with equivalent quality with increasing benefits in terms of efficiency and accuracy of the treatment planning process.[17,18] Moreover, the time window necessary to generate a plan solution will be compatible with image-guided four-dimensional adaptive radiation therapy, which pave the way to an even more personalized radiation therapy daily treatment.

Most of the treatment planning automation efforts of the last years have been focused on plan optimization. Automated optimization algorithms able to provide optimal solutions, from the mathematical point of view and also deliverable, have been actively investigated.[22,23] The plan optimization problem can be decoupled in two main problems: the choice of the best irradiation directions, through static or arc beams - the beam angle optimization (BAO) problem - and the determination of best shape and intensity of the irradiation beams - the fluence map optimization (FMO) problem. In addition to optimize the plan, also the quality assessment and the selection of the best plan have to be considered.

Over the last years, the BAO and the FMO problems have had different developments and implementations in the clinical routine. The BAO problem is a highly non-convex multi-modal optimization problem in a large search space with many local minima which cannot be solved exactly. Heuristics methods, sparse optimization and local search procedures considering discrete or continuous search spaces in combination with FMO algorithms have been used to address the BAO problem.[19-21] An appropriate selection of the irradiation directions can lead to important plan quality improvements.[22,23] However, BAO is complex, time-consuming and often presents non-intuitive solutions, which has prevented its full integration in most TPSs. Furthermore, for VMAT plans, fluence-based BAO solutions for non-coplanar irradiation geometries with simultaneous gantry and couch movements are under intensive investigation and are seldom available. The FMO topic has been object of important developments. Automated planning solutions based on knowledge-based planning, protocol based automatic iterative optimization, multicriteria optimization and more recently on deep learning methods for dose prediction are already available in commercial treatment planning systems.[24,25] The first clinical application studies showed a higher harmonization degree of plan

quality intra- and inter-centres and an increased efficiency in the radiation therapy planning process. Nevertheless, further investigations are needed to understand the real impact of these solutions in the radiation therapy departments' organization and in the patient's treatment outcome.

This research work had as primary objectives to support the clinical implementation of automated plan optimization methods and the development of treatment plan quality evaluation tools to assist the radiation oncologist decision-making. Thus, during this investigation the following topics were addressed:

- development of a new tool for plan quality assessment;
- quantification of the plan quality gains obtained by applying BAO optimization in head-and-neck and central nervous system pathologies;
- development of a new optimization approach for the arc trajectory optimization problem.

To assess the impact of automated treatment plan optimization in the quality of the planned dose distributions, the treatment planning system Erasmus-iCycle, an *a priori* multicriterial optimization IMRT engine developed by Breedveld *et al.*[15] in Erasmus MC was used. Based on a wish-list, containing the user-defined constraints and prioritized objectives, this system uses a constraint-based method, 2pεc, to automatically generate a single Pareto-optimal IMRT plan solution from a pre-selected beam angle ensemble. An iterative combinatorial BAO algorithm is integrated in Erasmus-iCycle FMO framework.

The first task of this work consisted in the customization and validation of a wish-list adapted to the clinical criteria used for head-and-neck IMRT plan optimization at Portuguese Oncology Institute of Coimbra using five representative nasopharynx cancer patients. This wish-list template was applied in the comparison of two BAO algorithms. The iterative BAO algorithm embedded in Erasmus-iCycle and a local search BAO method developed by Rocha *et al.*[20] were used retrospectively in 40 nasopharynx cancer (NPC) cases, already treated, to determine the best incidence of 5, 7 and 9 beams plan sets with coplanar and non-coplanar geometries. NPC cases were selected for this study due to their high complexity in radiation therapy planning (high prescription dose, proximity of critical structures, and extension of the radiation field). All structures used by the radiation oncologist in the clinical routine for plan approval were considered. This BAO run phase generated 3640 beams incidences in 520 plans with an average of 27 structures for each patient. To handle the comparison of around 14040 dosimetric parameters, a graphical evaluation method named SPIDERplan was developed. This tool was designed to support the radiation oncologist decision-making in plan approval by incorporating his/her clinical preferences in plan assessment. A scoring approach considering the targets coverage and the OARs sparing was incorporated into a graphical display allowing the simultaneous comparison and evaluation of multiple treatment plans.

SPIDERplan is presented in chapter 2 using a clinical nasal cavity test case. The proposal of this tool included its validation through a combinatorial search approach to incorporate the radiation oncologist clinical preferences in the plan assessment process.

To firmly support its clinical use, an extensive clinical validation of SPIDERplan performance was carried out for the NPC pathology (chapter 3). Three radiation oncologists belonging to the three hospitals of the Portuguese Oncology Institute group were invited to assess and compare a sample of twenty randomly selected NPC cases. Their choices and clinical preferences were embedded into SPIDERplan configuration using a mixed linear programming model. The performance of SPIDERplan evaluation was compared with radiation oncologist clinical evaluations. Moreover, the influence of the clinical preferences in SPIDERplan assessment and the significance of the difference between plan scores were discussed.

The dosimetric performance of the two previously referred BAO algorithms was compared using the SPIDERplan analysis. From this process, two outputs were delivered: a global comparison of the average quality of the plans generated by each of the two BAO algorithms (chapter 4) and a

specific-patient analysis where the advantages of using BAO in the coplanar and non-coplanar scenarios were highlighted (chapter 5).

The roadmap of BAO drove the research towards intracranial tumours (chapter 6). In the clinical routine, plan optimization for intracranial usually requires non-coplanar beams manually optimized by the planner to achieve high levels of PTV coverage and conformity and OARs sparing.[26] To test the potential advantages of beam angle optimization in such cases, non-coplanar BAO was performed for ten meningioma cases using a parallel multistart optimization framework and again iCycle-Erasmus to generate IMRT plans.[27] Due to the increasing importance of VMAT in radiation therapy, an arc trajectory optimization (ATO) algorithm based on the parallel multistart optimization framework was proposed. ATO relies on the recent possibility of simultaneous movement of the gantry and the couch, during irradiation. Expanding its usefulness, SPIDERplan score was here also used as the objective function in the optimization phase to guide the optimization of the irradiation directions. Non-coplanar trajectory optimization for the intracranial tumours was complemented in appendix 1 with a comparison study of the arc trajectories generated by algorithms based on iterative BAO and local search methods guided by SPIDERplan and implemented in multistart optimization frameworks.

The chapters described above correspond to submitted or published papers with the following references:

- **Chapter 2** SPIDERplan: a tool to support decision-making in radiation therapy treatment plan assessment, published in *Rep Pract Oncol Radiother*, 2016;21(6):508-516.

- **Chapter 3** Clinical validation of a graphical method for radiation therapy plan quality assessment, submitted to *Radiation Oncology* in November 2019.

- **Chapter 4** Comparison of two beam angular optimization algorithms guided by automated multicriterial IMRT, published in *Phys Medica,* 2019;64:210-221.

- **Chapter 5** Advantage of beam angle optimization in head-and-neck IMRT: patient specific analysis, published in *IFMBE Proceedings*, 2020:76:1256-1263.

- **Chapter 6** Non-coplanar optimization of static beams and arc trajectories for intensity-modulated treatments of meningioma cases, submitted to *Physica Medica* in October 2019.

- **Appendix 1** Evaluation of two arc trajectory optimization algorithms for intracranial tumours VMAT planning, submitted to *ESTRO 2020*.

1.5. References

1. WHO. Global Health Estimates 2016: Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva: World Health Organization; 2018.
2. Ferlay J, Ervik M, Lam F et al. Global Cancer Observatory: Cancer Today. Lyon; 2018. International Agency for Research on Cancer. https://gco.iarc.fr/today. Accessed September 6, 2019.
3. Delaney G, Jacob S, Featherstone C, Barton MB. The Role of Radiotherapy in Cancer Treatment Estimating Optimal Utilization from a Review of Evidence-Based Clinical Guidelines. *Cancer*. 2005;104(6):1129-1137.
4. Barton MB, Jacob S, Shafiq J et al. Estimating the demand for radiotherapy from the evidence: a review of changes from 2003 to 2012. *Radiother Oncol*. 2014;112(1):140-144.

5. Borras JM, Lievens Y, Dunscombe P et al. The optimal utilization proportion of external beam radiotherapy in European countries: an ESTRO-HERO analysis. *Radiother Oncol.* 2015;116(1):38-44.
6. WHO. WHO list of priority material medical devices for cancer management. Geneva: World Health Organization; 2017.
7. Zubizarreta E, Fidarova E, Healy B, Rosenblatt E. Need for radiotherapy in low and middle income countries – the silent crisis continues. *Clin Oncol*. 2015;27(2):107-114.
8. Atun R, Jaffray D, Barton MB et al. Expanding global access to radiotherapy. *Lancet Oncol*. 2015;16(10): 1153-1186.
9. Yang Y, Zhang P, Happersett L et al. Choreographing couch and collimator in volumetric modulated arc therapy. *Int J Radiat Oncol Biol Phys*. 2011;80(49):1238-1247.
10. Papp D, Bortfeld T, Unkelbach J. A modular approach to intensity-modulated arc therapy optimization with noncoplanar trajectories. *Phys Med Biol.* 2015;60(13):5179-5198.
11. ICRU. International Commission on radiation units and measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Geneva; 2010. *J ICRU* 10(1):1–106.
12. Thieke C, Küufer, KH, Monz M et al. A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. *Radiother Oncol*. 2007;85(2):292-298.
13. Fiege J, McCurdy B, Potrebko P, Champion H, Cull A. PARETO: A novel evolutionary optimization approach to multiobjective IMRT planning. *Med Phys*. 2011;38(9):5217-5229.
14. Craft D, Hong T, Shin H, Bortfeld T. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys.* 2012;82(1):e83-e90.
15. Breedveld S, Storchi P, Voet P, Heijmen B. iCycle: integrated, multi-criterial beam angle and profile optimization for generation of coplanar and non-coplanar IMRT plans. *Med Phys*. 2012;39(2):951-963.
16. Glatzer M, Panje C, Sirén C, Cihoric N, Putora P. Decision Making Criteria in Oncology. *Oncology*. 2018;1-9.
17. Moore K, Kagadis G, McNutt T, Moiseenko V, Mutic S. Vision 20/20: Automation and advanced computing in clinical radiation oncology. *Med Phys*. 2014;41(1):010901.
18. Moore K. Automated Radiotherapy Treatment Planning. *Semin Radiat Oncol*. 2019;29(3):209-218.
19. Dias J, Rocha H, Ferreira BC, Lopes MC. Simulated annealing applied to IMRT beam angle optimization: a computational study. *Phys Med*. 2015;31(7):747-756.
20. Rocha H, Dias J, Ferreira BC, Lopes MC. Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method. *Phys Med Biol*. 2013;58(9):2939-2953.
21. Liu H, Dong P, Xing L. A new sparse optimization scheme for simultaneous beam angle and fluence map optimization in radiotherapy planning. *Phys Med Biol*. 2017;62(16):6428-6445.
22. Das SK, Marks LB. Selection of coplanar or noncoplanar beams using three-dimensional optimization based on maximum beam separation and minimized nontarget irradiation. *Int J Radiat Oncol Biol Phys*. 1997;38(3):643–655.
23. Rocha H, Dias J, Ventura T, Ferreira BC, Lopes, MC. A derivative-free multistart framework for an automated noncoplanar beam angle optimization in IMRT. *Med Phys*. 2016;43(10):5514-5526.
24. Hussein M, Heijmen B, Verellen D, Nisbet A. Automation in intensity modulated radiotherapy treatment planning – a review of recent innovations. *Br J Radiol*. 2018;91(1092):20180270.
25. Cozzi L, Heijmen B, Muren L. Advanced treatment planning strategies to enhance quality and efficiency of radiotherapy. *Phys Imag Radiat Oncol*. 2019;11:69-70.
26. Bangert M, Ziegenhein P, Oelfke U. Comparison of beam angle selection strategies for intracranial IMRT. Med Phys. 2013;40(1):011716.
27. Rocha H, Dias JM, Ventura T, Ferreira BC, Lopes MC. Beam angle optimization in IMRT: are we really optimizing what matters? *Intl Trans in Op Res.* 2019;26:908-928.

chapter 2

SPIDERplan: a tool to support decision-making in radiation therapy treatment plan assessment

Reports of Practical Oncology and Radiotherapy, 2016, volume 21, issue 6, pages 508-516

T Ventura[1,2,3], MC Lopes[1,2,3], BC Ferreira[3,4], L Khouri[5]

[1]Physics Department, University of Aveiro, Aveiro, Portugal
[2]Medical Physics Department, IPOCFG, EPE, Coimbra, Portugal
[3]Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal
[4] School of Allied Health Technologies Polytechnic Institute of Porto, Porto, Portugal
[5]Radiotherapy Department, IPOCFG, EPE, Coimbra, Portugal

Abstract

Background: Optimized plans from modern radiotherapy are not easy to evaluate and compare because of their inherent multicriterial nature. The clinical decision on the best treatment plan is mostly based on subjective options.

Aim: In this work, a graphical method for radiotherapy treatment plan assessment and comparison, named SPIDERplan, is proposed. It aims to support plan approval allowing independent and consistent comparisons of different treatment techniques, algorithms or treatment planning systems.

Materials and Methods: SPIDERplan combines a graphical analysis with a scoring index. Customized radar plots based on the categorization of structures into groups and on the determination of individual structures scores are generated. To each group and structure an angular amplitude is assigned expressing the clinical importance defined by the radiation oncologist. Completing the graphical evaluation, a global plan score, based on the structures score and their clinical weights, is determined. After a necessary clinical validation of the group weights, SPIDERplan efficacy, to compare and rank different plans, was tested through a planning exercise where plans have been generated for a nasal cavity case using different treatment planning systems.

Results: SPIDERplan method was applied to the dose metrics achieved by the nasal cavity test plans. The generated diagrams and scores successfully ranked the plans according to the prescribed dose objectives and constraints and the radiation oncologist priorities, after a necessary clinical validation process.

Conclusion: SPIDERplan enables a fast and consistent evaluation of plan quality considering all targets and organs-at-risk.

Keywords: plan approval support, plan scoring, treatment planning, and radiation therapy.

2.1. Background

Radiotherapy is a treatment cancer modality that has as main purpose to eliminate through radiation in a controlled way tumour cells, while sparing as much as possible adjacent normal tissues. Each treatment session is delivered according to an optimized plan generated by a treatment planning system (TPS). The selection of the plan whose dose distribution better fulfils the medical prescription is not a trivial task. Visual inspection of the 3D dose distribution, a detailed analysis of the dose statistics and dose volume histograms (DVH) generated for each structure are the most common tools to assess the quality of the plan. To assist the radiation oncologist in plan selection, several graphical solutions such as the superposition of DVH and dose distributions or the side-by-side plan comparison have been made available in commercial TPSs. The large and diverse amount of data to be analysed, generally with conflicting results, makes plan selection hard and time-consuming. Final clinical decision is then made by the radiation oncologist mostly based on a subjective and qualitative assessment of the planned dose distributions taking into account only the most important features of the plan.[1] Optimal balance between the probability of tumour control and normal tissues complications is thus not guaranteed moreover when multiple plans are available for final clinical decision. Quantifying plan quality taking into account both the coverage of target volumes and the sparing of all organs-at-risk (OAR) in a simple and objective way has always been an ideal aim in the treatment planning process for helping the final clinical decision. First attempts have been proposed based on statistical decision theory,[2] multiattribute utility theory[3] and decision analysis concepts[3] for application to 3D conformal radiotherapy. However, they have never been incorporated in optimization algorithms nor implemented in treatment planning systems.

On the other hand, dose-quality treatment indexes have been considered a valuable contribution for plan assessment. Since the 90s, several indexes have been reported for external radiotherapy.[4-20] Target coverage and conformity indexes are the most common options. Target coverage index, known as RTOG index, first proposed by Shaw *et al.*[4] for radiosurgery treatment plans intended to measure the ratio of the minimum isodose to the prescribed dose in the planning target volume (PTV). Although its simplicity, the coverage index yields false positives and is extremely dependent of the selected reference isodose. A different coverage target score was proposed by the SALT group[5,6] and by Lomax and Scheib.[7] Based on a volumetric concept, this approach solves the previous drawbacks, but as for the RTOG index, it still does not take into account dose in the healthy tissues. Conformity indexes focus on the relation between the shape of the reference isodose and the PTV. Although the first conformity index types were developed for the evaluation of radiosurgery plans, Knöös *et al.*[9] extended this definition to breast, lung, prostate and head-and-neck pathologies. Though simple, these indexes neither consider the dose received by the normal tissues nor distinguish spatial mismatches between the target volume and the reference dose volume. Lomax and Scheib[7] presented a conformity index definition that includes the quantification of the irradiation of healthy tissues. To avoid false positives this index should be reported together with the target coverage index also proposed by Lomax and Scheib.[7] The conformation number (CN), proposed by van´t Riet *et al.*[10] and Paddick,[11] intended to assess the conformity of a radiosurgery plan, quantifying target coverage and normal tissues exposure. However, the CN score does not take into account the different OAR tolerances, considering all non-tumour tissues as a single critical structure with the same radiosensitivity. This was partially solved by the COIN score, initially developed for brachytherapy plans.[12] The COIN added to the CN expression a penalty factor for the unwanted doses in the OAR. As the dose values used into the score calculation are not adjusted to the tolerance level of each structure, a new index called critical organ scoring index (COSI) was proposed.[13] The COSI index allows a simultaneous quantification of target coverage and dose

irradiation of OARs. To get information about plan conformity, the COSI index must be represented in a 2D diagram against the conformity index proposed by Lomax and Scheib.[7] The inclusion of both target coverage and OAR affectation into CN, COIN and COSI definitions leads to a loss of information allowing that different dose distributions may present the same index values.

Based on the radiosurgery quality indexes, new score definitions applicable to all radiotherapy modalities have been published, namely for intensity-modulated radiation therapy (IMRT) plans. Due to their comprehensive definition and applicability the uncomplicated target conformity index (TCI+),[14] the conformity index incorporating dose and distance (CI_{DD}),[15] the plan quality index (PQI),[16] and the composite quality index (CPQI)[17] are the most representative scores. The TCI+, suggested by Miften *et al.*,[14] includes adjustable penalty functions evaluating target conformity and OARs sparing. Despite the good results, the sensitivity of the parameters of the penalty function depends of the clinical experience and the accuracy of the clinical tolerance data. A similar approach was proposed by Cheung and Law[15] with the CI_{DD}. In addition to the target coverage factor, this score also incorporates a target underdosage factor that presents a penalty function dependent on the value and localization of the cold spot. Nevertheless, the CI_{DD} neither quantifies the normal tissue sparing nor the conformity of the dose distribution to the target. The PQI developed by Leung *et al.*[16] incorporates multiple plan evaluation scores. Applying the Euclidean distance definition to the conformity of PTVs with different dose levels, target coverage and normal tissue sparing are assessed. The PQI index also includes a quantification of the cold and hot spots in the target coverage parameter. The normal tissue score includes a penalty function that allows the evaluation of different tolerance point criteria. More recently the composite quality index or CPQI[17] combines the target coverage, the homogeneity index, the equivalent uniform dose (EUD),[21,22] tumour control probability (TCP) and the normal tissue complication probability (NTCP) applying to each a relative weight. Plan quality indexes are a promising solution for treatment plan evaluation. However due to their definition or complexity, these tools were never widely adopted in clinical practice.

Using a completely different approach, alternative solutions for treatment plan evaluation may be based on radiobiological functions. TCP, NTCP and the probability of uncomplicated tumour control[23] are examples of radiobiological measures that may estimate the outcome of a treatment plan. However nowadays, the benefits that could be achieved with this approach are still overshadowed by the uncertainty associated to the variables needed during the biological modelling process.

With the implementation of inverse treatment planning optimization, a variety of objective functions were implemented into TPSs to drive the optimization algorithm in each iteration. The score corresponding to the value of the objective function should ideally be correlated with the quality of the dose distribution and thus the final plan would represent the optimal radiation treatment. Nevertheless, this is generally a pure mathematical expression without clinical meaning. Furthermore, a strict comparison between plans optimized in different treatment planning systems is only valid when the same objective function is used. Comparisons between different optimization algorithms are thus not generally made simply based on the final value of the objective function.

2.2. Aim

In this study a plan quality assessment tool called SPIDERplan is proposed. This tool intends to evaluate the quality of different plans using an intuitive graphic representation and an associated score function. These are based on the target and normal tissues objectives/constraints defined by the radiation oncologist and are completely independent of the algorithm, the treatment technique or the TPS. Without any ambition to correlate this score with treatment outcome, SPIDERplan aims to be a supporting tool to help during clinical decision-making in consistently selecting the best treatment plan available.

2.3. Materials and methods

2.3.1. SPIDERplan concept description

SPIDERplan is a graphical method developed to assess and compare the quality of different radiotherapy treatment plans. Based on a scoring approach both target coverage and individual OAR sparing are considered, allowing the quantification of the global quality of any treatment planning modality. This method can be described in two phases: 1) processing of the plan data and 2) assessment of plan quality. The outputs for plan evaluation are graphs and scores that provide the user with a fast and intuitive image of plan(s) quality. Clinical validation to fine tune SPIDERplan input variables is necessary before using it in clinical practice.

In phase one, processing of the plan data, targets and OARs are divided into groups (e.g. optical track group, including organs like optic nerves and chiasm). To each structure a score based on planning objectives that may be either dosimetric, volumetric, radiobiological or mathematical is assigned. This score is a scalar that expresses the performance of that structure on accomplishing the corresponding planning goal. A pre-defined relative weight is also attributed to each group and to each structure expressing the clinical priorities during the plan evaluation process. The definition of these groups and weights must be customized by the local clinical team according to the tumour type prior to the clinical use of SPIDERplan. The numerical weighs reflect the relative importance given by the radiation oncologist to the different planning aims. Similarly, the grouping of the different structures strongly depends on the tumour type and morphology and must reflect the relative importance the physician will give to the different structures when he/she will appreciate the dose distributions. Thus, SPIDERplan is completely designed for each tumour type according to the radiation oncologist preferences.

In phase two, assessment of plan quality, all this information is graphically represented in customized radar plots. Evaluation of plan quality can be done globally visualizing all structures and groups' information or in more detail assessing each group analysis in partial radar plots. For the global analysis mode, where all structures are considered, a radar plot named Structures Plan Diagram (SPD) is generated. The circular plotting area is divided into sections and subsections with an angular amplitude proportional to the relative weight of the respective group and structure. The score of each structure is represented by a point along the angle bisector of the respective subsection whose distance from the centre of the radar plot corresponds to the score value. By connecting the score of all structures a polygon representing the quality of the dose distribution is generated. Global plan score is determined as a weighted sum of the structures individual scores over all groups as:

$$\text{Global plan score} = \sum_{i} w_{group(i)} \sum_{j} w_{struct(j)} Score_{struct(j)} \qquad (2.1)$$

where $w_{struct(j)}$ and $Score_{struct(j)}$ are the relative weight and the score of structure j, respectively, and $w_{group(i)}$ the relative weight of group i.

The SPD information may also be condensed by groups of structures. In this option the radar plot is called Group Plan Diagram (GPD), where it is designed for the groups instead of the structures. A more detailed group evaluation can also be done with the partial group plots. The group plots are named Structures Group Diagram (SGD). Only the structures of a particular group are then represented. As for the SPD and GPD, a partial group score complementing the graphical assessment is determined for each SGD.

2.3.2. Application to a clinical case

To assess SPIDERplan ability to compare and rank radiotherapy plans, an already treated nasal cavity clinical case was selected and re-planned in iPlan RT Dose version 4.5 from BrainLAB. IMRT technique was based on an m3 micro multi-leaf collimator (mMLC) from BrainLAB. For inverse optimization, this TPS generates three IMRT plans with different OAR priorities: low, medium and high. Irradiation technique used eight non-coplanar beams composed by 15 segments. Dose calculation was done applying a Pencil Beam algorithm and a dose grid of $1x1mm^2$.

SPIDERplan independence from TPS and dose calculation algorithms was assessed by comparing iPlan plans with a plan generated in Oncentra version 4.1 SP2 optimizer module from Elekta for the same clinical case. The Oncentra IMRT plan was generated using a Siemens 82-leaf Optifocus multi-leaf collimator (MLC). Collapsed Cone algorithm with a dose grid of $1x1mm^2$ was applied to optimize 9 beams delivering 100 segments.

A dose of 60Gy delivered in 30 fractions was prescribed to the PTV. The main considered OARs were lens, retinas, chiasm, optical nerves, pituitary gland, brainstem, cochlea, parotids and oral cavity. The respective tolerance doses values, shown in Table 2.1, were established in agreement to the institutional protocols defined for this pathology.

Structures were grouped into PTV group, Optics group and Other group, Table 2.1. These groups were defined according to structures localization and clinical importance for the concerned pathology. The PTV group just comprised the PTV. The Optics group was constituted by all the optical structures: lens, optical nerves, retinas and chiasm, which were next to the PTV and so acted as critical structures for planning. In the Other group all other structures that were distant from the PTV or whose tolerance dose values were not critical for this particular planning context were included: brainstem, pituitary gland, cochlea, parotids and oral cavity (Figure 2.1). This grouping reflects the relative importance that the radiation oncologist gives to the different structures when appreciating the different treatment options.

The score of each structure may be determined by any function locally used to evaluate plan quality. In this study, a score based on the ratio between clinical tolerance criteria and the planned dose was used. Thus, a value of one is expected if the dose for that structure is equal to the respective tolerance value. When a better organ sparing or target coverage is achieved, a score less than one will be obtained. In the SPIDERplan diagrams, this becomes easily perceived by the relation obtained with two circles: the inner circle with unitary radius representing the limit of acceptability

Table 2.1 - Groups and structures considered for SPIDERplan processing.

Groups	**Structures**	
Name	**Name**	**Tolerance criteria**
PTV group	PTV	$D_{98\%} \geq 95\%$ of 60Gy
Optics group	Chiasm	$D_{max} \leq 50Gy$
	Left Optical Nerve (OPTNRVL)	$D_{max} \leq 50Gy$
	Right Optical Nerve (OPTNRVR)	$D_{max} \leq 50Gy$
	Left Retina (RETINAL)	$D_{max} \leq 45Gy$
	Right Retina (RETINAR)	$D_{max} \leq 45Gy$
	Left Lens (LENSL)	$D_{max} \leq 12Gy$
	Right Lens (LENSR)	$D_{max} \leq 12Gy$
Other group	Brainstem	$D_{max} \leq 54Gy$
	Pituitary Gland (PITUITARY)	$D_{max} \leq 60Gy$
	Left Cochlea (COCHLEAL)	$D_{mean} \leq 45Gy$
	Right Cochlea (COCHLEAR)	$D_{mean} \leq 45Gy$
	Oral Cavity (ORALCAVITY)	$D_{mean} \leq 45Gy$
	Left Parotid (PAROTIDL)	$D_{mean} \leq 26Gy$
	Right Parotid (PAROTIDR)	$D_{mean} \leq 26Gy$

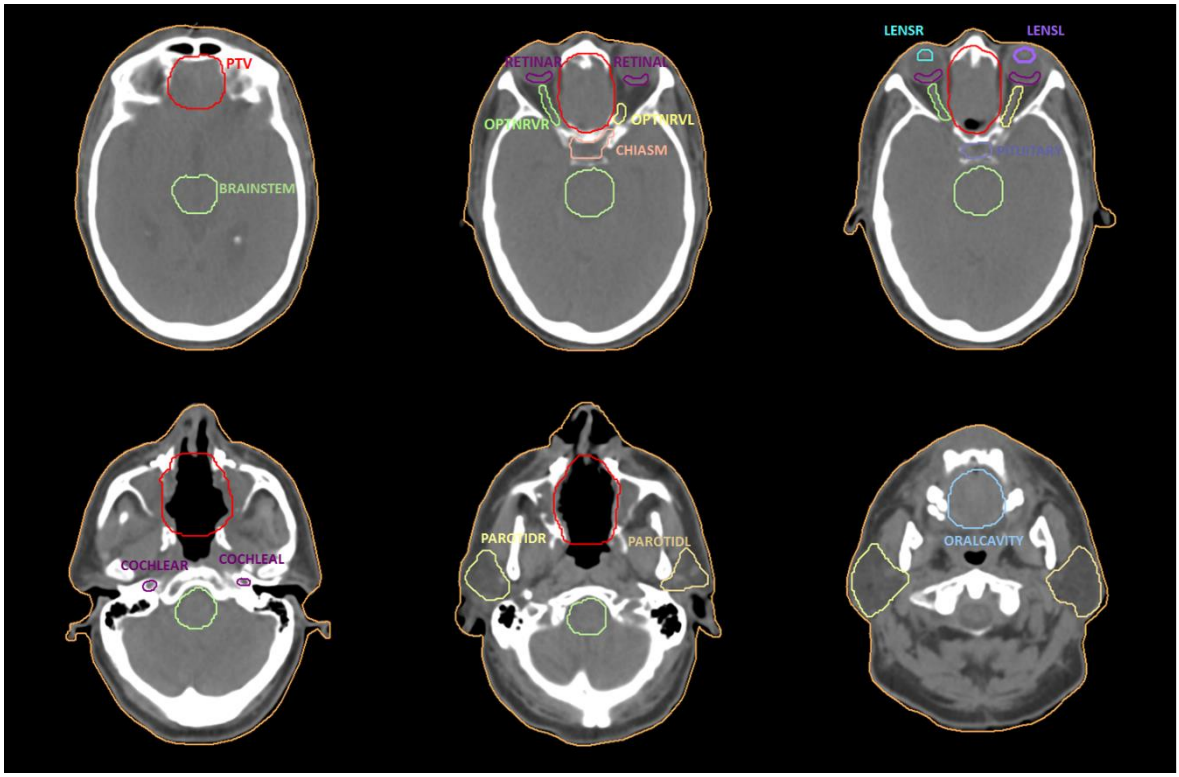

Figure 2.1 - Nasal cavity structures. Nasal cavity case CT images slices containing all the structures considered for planning.

and the outer circle with radius equal to two representing failures. Optimal scores will converge to the radar plot centre.

For the PTV, the score was calculated according to the following expression:

$$Score_{PTV} = \frac{D_{TC,PTV}}{D_{P,PTV}} \tag{2.2}$$

where $D_{TC,PTV}$ corresponds to the tolerance criteria for the PTV (in this case the dose in 98% of the PTV that should be at least 95% of the prescribed dose) and $D_{P,PTV}$ is the planned dose in the PTV. This is a target coverage criterion.

For the OARs, the score was set as:

$$Score_{OAR} = \frac{D_{P,OAR}}{D_{TC,OAR}} \tag{2.3}$$

where $D_{P,OAR}$ is the OAR planned dose and $D_{TC,OAR}$ is the tolerance dose for each OAR. In this clinical example, the score of each structure was based on the tolerances commonly used in our clinical practice corresponding to ICRU Report 83[24] planning aims. For some pathologies there are broad consensus guidelines (e.g. RTOG0615 for nasopharyngeal cancer). Yet, seldom will a radiation oncologist be able to give the planner more than accepted tolerance doses for OARs either expressed as maximum doses or mean doses in a volume, considering more or less subjective

values for the probability of complications in a tissue while selecting treatment plans. The plan is considered acceptable for treatment when these criteria are met. In this study the function score adopted compares the ratio between the prescription and the planned dose. But this function serves just as an example. SPIDERplan admits any type of score function as long as the unity value separates admissible from inadmissible planning aims. These definitions are not static and can be customized in any moment according to the clinician criteria and to the case specificity.

2.3.3. SPIDERplan clinical validation

As SPIDERplan was developed to reflect the radiation oncologist criteria when he/she is approving a plan, a clinical validation of the tool was performed prior to its application. First, all plans were assessed and ranked by the radiation oncologist using traditional tools: the DVHs, the dose statistics and the dose distribution visualization, in a completely independent way from SPIDERplan. To complete clinical plan evaluation the radiation oncologist classified the plans as 'Good', 'Admissible with minor deviations' or 'Not admissible'.

To calibrate SPIDERplan to the radiation oncologist criteria and preferences, a matrix, named here weight sensitivity matrix, was determined. This matrix was calculated by varying the weights of all groups and considering all possible group weights combinations. The weight of the group representing the OAR was varied from 0% to 100% in steps of 10%. For the PTV group the weight varied from 10% to 100% also in steps of 10%. The set of group weights that better fitted the radiation oncologist evaluation was selected in SPIDERplan processing.

2.4. Results

2.4.1. SPIDERplan outputs

The proposed plan evaluation tool functionality and results will be illustrated through the chosen nasal cavity clinical case. SPIDERplan processing phase used the groups and structures defined in Table 2.1, and the group clinical weights defined by the radiation oncologist.

The subsequent assessment phase for SPIDERplan generated different radar plots that can be analysed by the planner or the radiation oncologist. The SPD, GPD and SGD are shown in Figure 2.2 and 2.3.

From the SPD the best overall plan was the iPlan OAR High (displayed polygon closer to the plot centre). By contrary Oncentra and the iPlan OAR Low plans were the plans with lower quality. The diagram is complemented by the global score that supports that the plan iPlan OAR High, with a global score of 0.940, was the best dose distribution while the iPlan OAR Low plan, with a score of 1.040, was the worst plan. GPD is the condensed diagram showing the results for the groups. For the nasal cavity tumour case the plan iPlan OAR High was the only that achieved a global score equal or less than one for all groups. iPlan OAR Medium and iPlan OAR Low obtained a poor result for the Optics group and Oncentra for the PTV and the Optics groups. As the global score is derived using the same methodology in all diagrams, both radar plots may be used for plan selection.

A more detailed analysis of plan quality can be done evaluating the score of each structure in a given group (Figure 2.3). For structures belonging to a group with just one element, like the PTV group, this can be done directly in the Structures Plan Diagram. For the PTV, the lowest score was achieved by the iPlan OAR Low plan with a value of 0.959, followed by the iPlan OAR High, Medium and the Oncentra plans with scoring values of 0.969, 0.997 and 1.009, respectively. For larger groups, a more comprehensive analysis can be performed by visualizing the SGD (Figure 2.3). For structures belonging to the Optics group, iPlan OAR High was the only plan that accomplished all the clinical criteria (the red polygon is included in the inner circle). By contrary iPlan OAR Low plan

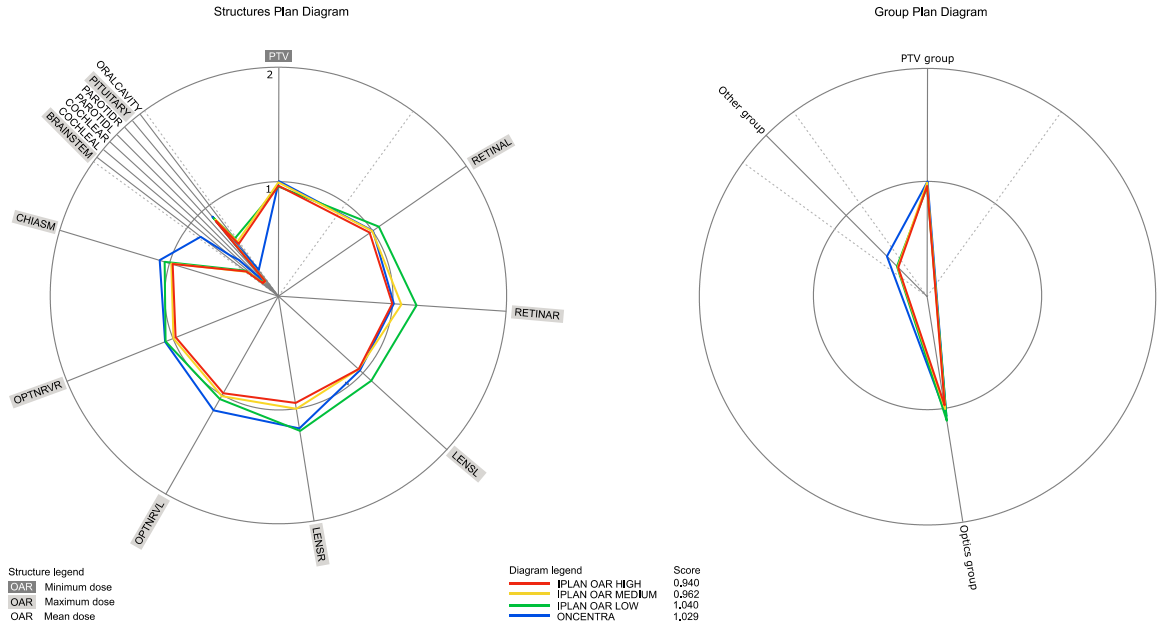

Figure 2.2 - Structures Plan Diagram and Group Plan Diagram. SPIDERplan Structures Plan Diagram and Group Plan Diagram obtained for a nasal cavity tumour case. The solid rays of the diagram correspond to a structure and the dashed rays limits each group.

was unable to achieve any of the clinical constraints for this group (the green polygon is external to the unitary circle), presenting the worst score values for the retinas and the lens. For the optic nerves and the chiasm the worst results were obtained for the Oncentra plan (blue line). In general, for the Optics Group diagram the best plan scoring was achieved by iPlan OAR High followed by iPlan OAR Medium, Oncentra and iPlan OAR Low with group score values of 0.972, 1.003, 1.069 and 1.105 respectively. For the Other group, all plan optimizations met the respective clinical criteria defined

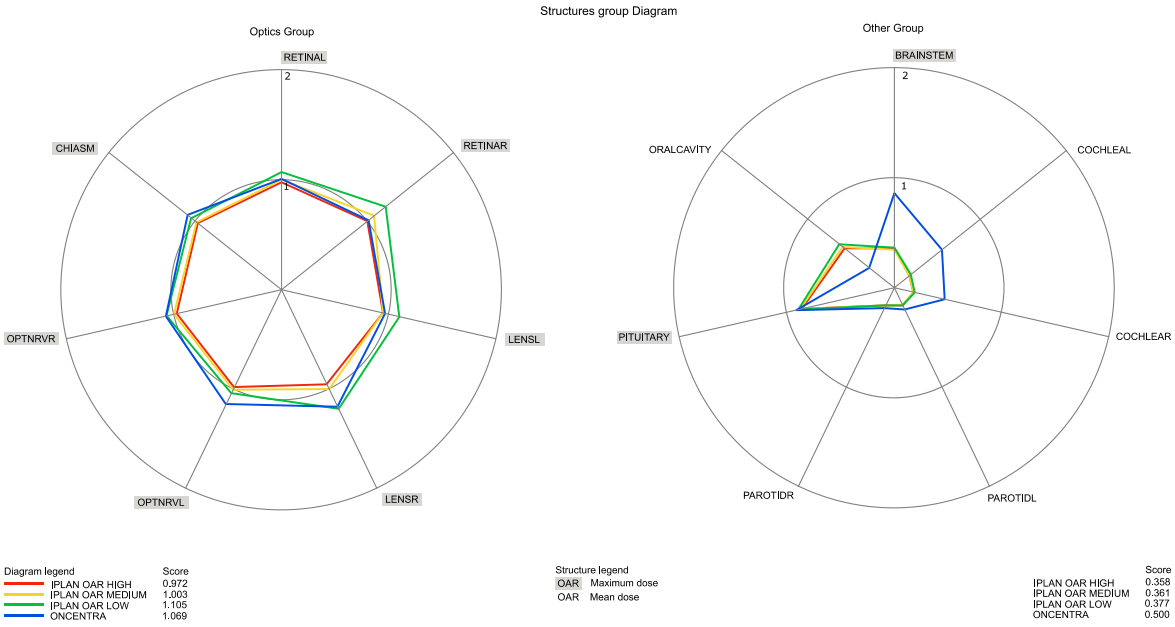

Figure 2.3 - Structures Group Diagram. SPIDERplan Structures Group Diagram generated for Optics group and for Other group, respectively.

for the OARs in this group. The absolute value of these partial group scores can just be used for relative comparisons within each group.

2.4.2. SPIDERplan clinical validation

Prior to SPIDERplan configuration, the radiation oncologist was asked to evaluate all iPlan and Oncentra plans using traditional tools (DVHs, planar dose distributions, etc). The results are shown in Table 2.2. All plans were considered acceptable for treatment. iPlan OAR High was the plan that achieved the best rank followed by the iPlan OAR Medium, Oncentra and iPlan OAR Low plans respectively. All plans with exception of the iPlan OAR High plan were classified as 'Admissible with minor deviations. iPlan OAR High plan was qualitatively evaluated as 'Good'.

To test the sensitivity of the weights needed for SPIDERplan configuration, the weight sensitivity matrix was calculated. From all combinations of group weights, the cluster with those weight combinations that maintained the plan ranking of the radiation oncologist (Table 2.2) has been selected. Figure 2.4 shows the six combinations of group weights of this cluster clearly illustrating that while the weights of the Optics group ranged from 60% to 90%, the PTV group weights ranged from 10% to 40% and the Other group weights ranged from 0% to 10%, plan ranking doesn't change. In all six group weights combinations, a common trend was verified: the Optics group presented a higher weight than the PTV group and the Other group. Furthermore, the weight assigned to each group is not critical as there is a relatively large range where the quality of the plans is sorted by SPIDERplan and the radiation oncologist in the same way.

Using this analysis a robust weight assignment could be done in SPIDERplan processing for nasal cavity. The mean value of the weight group ranges of the weight sensitivity matrix were taken as: 75% for Optics group, 20% for PTV group and for 5% Other group.

Table 2.2 - Clinical evaluation of the optimized IMRT plans in iPlan and Oncentra.

Name	Clinical plan assessment	
	Rank	Qualitative classification
iPlan OAR High	1	Good
iPlan OAR Medium	2	Admissible with minor deviations
Oncentra	3	Admissible with minor deviations
iPlan OAR Low	4	Admissible with minor deviations

2.5. Discussion

The ideal tool to support treatment plan evaluation would be able to provide an objective measure identifying the best treatment by quantitatively assess the quality of the dose distribution and consistently compare different plans.[25] During the last two decades several proposals have been presented. Nevertheless, due to their complexity or conceptual limitations, no solution has been systematically adopted by the radiation oncology community.

SPIDERplan was designed to be incorporated in a TPS as one of the evaluation tools aiming to complement existing ones and not to replace traditional plan evaluation methods like dose distribution visual inspection or DVH analysis. Multicriterial optimization and Pareto front navigation are eligible applications of the proposed tool. In those environments, by definition, a multiplicity of plans are to be considered. Along with playing with a display tool for Pareto front navigation, as described by Craft *et al.*,[26] the physician could also have an immediate picture of the whole set of

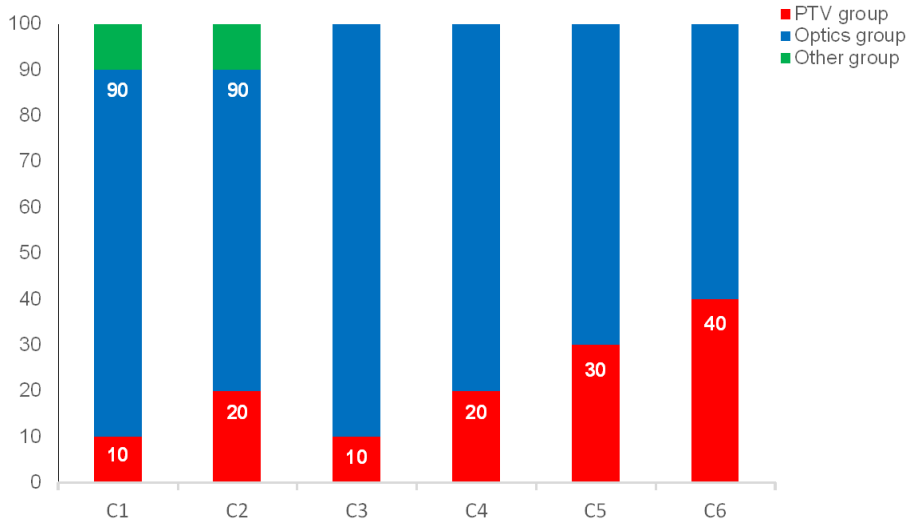

Figure 2.4 - Group weight sensitivity cluster. Cluster of combinations of group weight that produce the same plan ranking as the radiation oncologist.

plan possibilities in just one graph. With the benefit of having a complementary information on the relative value of each plan according to his/her own criteria.

The novelty of the proposed tool is not the use of radar plots. There are many examples in the literature of their usage.[27,28] The novelty of SPIDERplan is the way the radar plot is built. By incorporating the clinical criteria defined in dose prescription, SPIDERplan becomes the graphical representation of radiation oncologists viewpoints leading to plan approval. With just one diagram, combined with a score index, the medical team would be able to take decisions consistently when comparing different plans. It is a simple idea using common graphics to handle a difficult problem. SPIDERplan intuitive design allows a much faster decision in selecting the best plan taking into account all structures. When the polygon runs outside the inner circle (of radius one), dose objectives or constraints were surpassed and plan may be rejected. Otherwise the dose distribution fulfils dose prescription and tolerances and the polygon with the smallest area indicates the best plan available.

Treatment plan quality assessment performed by SPIDERplan can be applied to all pathologies, dose calculation algorithms and delivery techniques. It incorporates the potential for customization for each radiation oncologist priorities and criteria needed for plan approval. This is accomplished by assigning in the processing phase a pre-defined relative weight to groups and structures that can be adjusted according to the case complexity or clinical demands. The score function determined for each structure is also configurable. Different scoring indexes based on traditional DVH information, or any other relevant physical or biological quantities like EUD, TCP, NTCP, etc.[29] may be set up. This flexibility confers to SPIDERplan the possibility to be easily adapted to new situations that are recurrent in the clinical routine. For instance, some special clinical situation where a given OAR should be given a stronger relevance or the inclusion of a structure in the optimization process that is not usually considered. For situations where mixture of different types of planning goals are specified for each structure, the user can use composite scores[16] or more than one radius per structure in the SPIDERplan radar plot. For the hot spots situations, a score including unspecified normal tissue (body minus PTVs) dose restriction may be included. Also, some penalties to structure scores can be introduced in order to point unacceptable deviations that would conduct to rejected plans.

The processing phase can be constructed on a class-solution basis, like for instance some "wish-lists" that have been proposed for optimization steering.[30] All these tool refinements must be conducted in close collaboration with radiation oncologists. Some time and effort needs to be invested in the pre-processing phase to clinically validate SPIDERplan for each tumour site. To adjust SPIDERplan plan assessment with the radiation oncologist evaluation is an essential step to

enhance the confidence and creditability in this new tool. The proposed weight sensitivity matrix method expresses the clinical preferences for plan assessment and gives a picture of the sensitivity of the clinical weights that should be used in the configuration of SPIDERplan. This initial effort will be compensated by a much faster plan evaluation and treatment selection of similar clinical cases.

SPIDERplan potential to compare and rank treatment plans was demonstrated in the chosen nasal cavity tumour case. After the clinical validation, the graphic information included in the Structures Plan Diagram, the Group Plan Diagram and the Structures Group Diagram enabled a fast and intuitive assessment of the quality of the dose distribution in the PTV and all OARs. To complement the diagrams evaluation, the plan global score provided a quantitative measure of plan quality. This measure was embedded in the SPIDERplan construction including both the score achieved by each structure and also its relative importance according to the local or individual clinical reasoning. The new tool is independent from the TPS and the dose calculation algorithm. In a simple and systematic way it is possible to compare plans from different sources and consistently infer about their relative quality. The only requirement is that the dose prescription, either in the form of dose objectives or constraints, is the same. The number of structures in the presented clinical example was restricted to 15 (Table 2.1) but this number has no limitation. Although the Structures Plan Diagram may appear too confusing for more complex cases, the Structures Group Diagrams may then be preferable and simpler to analyse providing the same global result. Moreover, the global plan score will complement the information provided by the radar plots quantifying the quality of the best plan by incorporating the dose prescription and tolerances defined to all target volumes and OARs.

2.6. Conclusions

A graphical method for comparison and assessment of the quality of radiation therapy plans was developed. A clinical planning case of nasal cavity was used to illustrate its operationality, functionalities and potential. Due to its simplicity and flexibility SPIDERplan can be applied to all types of radiotherapy delivery techniques. The graphical information generated by the Structures Plan, the Group Plan and the Structures Group Diagrams complemented by the global plan score enables a fast and intuitive plan quality assessment incorporating all clinical priorities and criteria established for treatment plan approval.

Following the local clinical practice, the complete list of clinical planning aims needs to be translated into SPIDERplan inputs. These may be established in terms of dose prescription, PTV coverage and conformity, dose tolerance to OARs, dose-volume conditions, etc. The customization process must be clinically validated by comparing the results of the SPIDERplan assessment with the radiation oncologist evaluation. This tuning phase will confirm consistency with the processing data of SPIDERplan.

It was demonstrated that SPIDERplan can be used as a useful tool for supporting clinical plan approval with full customization potential to any local clinical practice. After the clinician realizes that his/her priorities and goals are fully reflected in the graphics and scores, SPIDERplan can easily be adopted with a very steep learning curve.

2.7. References

1. Lambin P, van Stiphout RG, Starmans MH et al. Predicting outcomes in radiation oncology – multifactorial decision support systems. *Nature Reviews Clinical Oncology.* 2013;10(1):27-40.
2. Schultheiss TE, Orton CG. Models in radiotherapy: Definition of decision criteria. *Med Phys.* 1985;12(2):183-187.
3. Jain NL, Kahn MG, Drzymala RE, Emami BE and Purdy JA. Objective evaluation of 3-d radiation treatment plans: A decision-analytic tool incorporating treatment preferences of radiation oncologists. *Int J Radiat Oncol Biol Phys.* 1993,26(2):321-333.
4. Shaw E, Kline R, Gillin M et al. Radiation therapy oncology group: Radiosurgery quality assurance guidelines. *Int J Radiat Oncol Biol Phys*. 1993;27(5):1231-1239.

5. Lefkopoulos D, Grandjean P, Platoni K. Les progrès dans l'optimisation des plans dosimétriques en radiothérapie en conditions stéréotaxiques dans le groupe Salt. *Cancer/Radiother*. 1998;2(2):127-138.
6. Dejean C, Lefkopoulos D, Foulquier JN, Schlienger M, Touboul E. Définition automatique de l'isodose de prescription pour les irradiations stéréotaxiques de malformations artérioveineuses. *Cancer/Radiother*. 2001;5(2):138-149.
7. Lomax NJ, Scheib SG. Quantifying the degree of conformity in radiosurgery treatment planning. *Int J Radiat Oncol Biol Phys*. 2003;55(5):1409-1419.
8. Nedzi LA, Kooy HM, Alexander E, Svensson GK, Loeffler JS. Dynamic field shaping for stereotactic radiosurgery: A modelling study. *Int J Radiat Oncol Biol Phys*. 1993;25(5):859-869.
9. Knöös T, Kristensen I, Nilsson P. Volumetric and dosimetric evaluation of radiation treatment plans: Radiation Conformity Index. *Int J Radiat Oncol Biol Phys*. 1998;42(5):1169-1176.
10. van't Riet A, Mak CA, Moerland MA, Elders LH, van der Zee W. A conformation number to quantify the degree of conformality in brachytherapy and external beam irradiation: Application to the prostate. *Int J Radiat Oncol Biol Phys*. 1997;37(3):731-736.
11. Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. *J Neurosurg*. 2000;93:219-222.
12. Baltas D, Kolotas C, Geramani K, Mould RF, Ioannidis G, Kerchidi M. A conformal índex (COIN) to evaluate implant quality and dose specification in brachytherapy. *Int J Radiat Oncol Biol Phys*. 1998;40(2):515-524.
13. Menhel H, Levin D, Alezra D, Symon Z, Pfeffer R. Assessing the quality of conformal treatment planning: a new tool for quantitative comparison. *Phys Med Biol*. 2006;51(20):5363-5375.
14. Miften M, Das SK, Su M, Marks LB. A dose-volume-based tool for evaluating and ranking IMRT treatment plans. *J Appl Clin Med Phys*. 2004;5(4):1-14.
15. Cheung F, Law M. A novel conformity index for intensity modulated radiation therapy plan evaluation. *Med Phys*. 2012;39(9):5740-5756.
16. Leung L, Kan M, Cheng A, Wong W, Yau C. A new dose–volume-based Plan Quality Index for IMRT plan comparison. *Radiother Oncol*. 2007;85(3):407-417.
17. Jin X, Yi J, Zhou Y, Yan H, Han C, Xie C. A new plan quality index for nasopharyngeal cancer SIB IMRT. *Phys Medica*. 2014;30(1):122-27.
18. Akpati H, Kim C, Kim B, Park T, Meek A. Unified dosimetry index (UDI): a figure of merit for ranking treatment plans. *J Appl Clin Med Phys*. 2008;9(3):99-108.
19. Wagner T, Bova F, Friedman W, Buatti J, Bouchet L, Meeks S. A simple and reliable index for scoring rival stereotactic radiosurgery plans. *Int J Radiat Oncol Biol Phys*. 2003;57(4):1141-1149.
20. Wu Q, Wessels B, Einstein D, Maciunas R, Kim E, Kinsella T. Quality of coverage: Conformity measures for stereotactic radiosurgery. *J Appl Clin Med Phys*. 2003;4(4):374-381.
21. Niemierko A. Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Med Phys*. 1997;24(1):103-110.
22. Niemierko A. A generalized concept of equivalent uniform dose (EUD). *Med Phys*. 1999;26(6):1100.
23. Ågren A, Brahme A, Turesson I. Optimization of uncomplicated control for head and neck tumors. *Int J Radiat Oncol Biol Phys*. 1990;19(4):1077-1085.
24. ICRU. International Commission on radiation units and measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Geneva; 2010. *J ICRU* 10(1):1–106.
25. Feuvret L, Noël G, Mazeron J, Bey P. Conformity Index: A Review. *Int J Radiat Oncol Biol Phys.* 2006;64(2):333-342.
26. Craft D, Hong T, Shih H, Bortfeld T. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys.* 2012;82(1):e83-e90.
27. Thieke C, Küfer K, Monz M et al. A new concept for interactive radiotherapy planning with multicriteria optimization: First clinical evaluation. *Radiother Oncol*. 2005;85(2):292-298.
28. Olteanu L A M, Berwouts D, Madani I, De Gersem W, Vercauteren T, Duprez F and De Neve W. Comparative dosimetry of three-phase adaptive and non-adaptive dose-painting IMRT for head-and-neck cancer. *Radiother Oncol*. 2014;111(3):348-353.
29. Njeh CF, Parker BC, Orton CG, Evaluation of treatment plan using target and normal tissue DVHs is no longer appropriate. *Med Phys*. 2015;42(5):2099.
30. Breedveld S, Storchi P and Heijmen B. The equivalence of multi-criteria methods for radiotherapy plan optimization. *Phys Med Biol*. 2009;54(23):7199-7209.

chapter 3

Clinical validation of a graphical method for radiation therapy plan quality assessment

submitted to Radiation Oncology in November 2019

T Ventura[1,2,3], J Dias[3,4], L Khouri[5], E Netto[6], A Soares[7], BC Ferreira[3,8,9],H Rocha[3,4], MC Lopes[1,2,3]

[1]Physics Department, University of Aveiro, Aveiro, Portugal
[2]Medical Physics Department, IPOCFG, EPE, Coimbra, Portugal
[3]Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal
[4]Economy Faculty of University of Coimbra and Centre for Business and Economics Research, Coimbra, Portugal
[5]Radiotherapy Department, IPOCFG, EPE, Coimbra, Portugal
[6]Radiotherapy Department, IPOLFG, EPE, Lisboa, Portugal
[7]Radiotherapy Department, IPOPFG, EPE, Porto, Portugal
[8]School Health Polytechnic of Porto, Porto, Portugal
[9]I3N, Physics Department, University of Aveiro, Aveiro, Portugal

Abstract

Background: This work aims at clinically validating a graphical tool developed for treatment plan assessment, named SPIDERplan, by comparing the plan choices based on its scoring with the radiation oncologists (RO) clinical preferences.

Methods: SPIDERplan validation was performed for nasopharynx pathology in two steps. In the first step, three ROs from three Portuguese radiotherapy departments were asked to blindly evaluate and rank the dose distributions of twenty pairs of treatment plans. For plan ranking, the best plan from each pair was selected. For plan evaluation, the qualitative classification of 'Good', 'Admissible with minor deviations' and 'Not Admissible' were assigned to each plan. In the second step, SPIDERplan was applied to the same twenty patient cases. The tool was configured for two sets of structures groups: the local clinical set and the groups of structures suggested in international guidelines for nasopharynx cancer. Group weights, quantifying the importance of each group and incorporated in SPIDERplan, were defined according to RO clinical preferences and determined automatically by applying a mixed linear programming model for implicit elicitation of preferences. Intra- and inter-rater ROs plan selection and evaluation were assessed using Brennan-Prediger kappa coefficient.

Results: Two-thirds of the plans were qualitatively evaluated by the ROs as 'Good'. Concerning intra- and inter-rater variabilities of plan selection, fair agreements were obtained for most of the ROs. For plan evaluation, substantial agreements were verified in most cases. The choice of the best plan made by SPIDERplan was identical for all sets of groups and, in most cases, agreed with RO plan selection. The cases where there was a difference between the RO choice and the one made using SPIDERplan corresponded to very low score differences. A score difference threshold of 0.005 was defined as the value below which the plans are considered of equivalent quality.

Conclusions: Generally, SPIDERplan response successfully reproduced the ROs plan selection. SPIDERplan assessment performance can represent clinical preferences based either on manual or automatic group weight assignment. For nasopharynx cases, SPIDERplan was robust in terms of definitions of structure groups, being able to support different configurations without losing accuracy.

Keywords: decision-making, plan quality assessment, clinical validation

3.1. Background

The delivery of radiation therapy is based on a pre-calculated personalized dose plan optimized in a treatment planning system. A plan that simultaneously irradiates the target with the prescription dose and causes little or no damage to the organs-at-risk (OAR) and to the adjacent normal tissues is sought by the planner.[1] It is usually necessary to consider trade-offs between the dose delivered to the targets and the dose received by the normal tissues. So, each plan is a compromise solution between conflicting objectives. These compromises must generally be tackled by the human planner in an iterative manual trial-and-error process. Thus, plan optimization can be seen as a decision-making problem handled by a planner that attempts to simultaneously fulfil the dose prescription objectives and the tolerance dose criteria. As a result, the plan optimization phase is extremely dependent of the planner's experience and the complexity of the case and it cannot be guaranteed that the calculated plan or plans presented to the radiation oncologist (RO) are the best possible ones.[2]

The clinical assessment of plan quality is typically done by verifying the fulfilment of the prescription dose in the target volume and the tolerance dose criteria for each OAR. The most common assessment methods used in the clinical routine are the visual inspection of the isodoses displayed on top of the computed tomography images and the evaluation of the dose-volume histograms (DVHs) and the corresponding dose statistics. To yield a comprehensive appraisal of the quality of the 3D dose distribution, it is often necessary to take into account several dozens of parameters and that is not humanly possible.[3] If two or more of the best plans are to be compared, this task becomes even more demanding. As a result, plan selection is based on the information that the RO managed to hold or considered more relevant which may lead to unsystematic and/or subjective decisions.

As in many other medical fields, the RO decision about which plan should be elected for treatment is not only influenced by disease specific criteria (e.g. cancer stage, age, comorbidities or treatment toxicity) but also by the decision-maker individual characteristics (e.g. experience, emotions or degree of expertise) and by contextual factors (e.g. patient socioeconomic status, healthcare provider organization or political environment).[16] Ideally, this complex decision-making framework should be supported by clinical reasoning methods able to efficiently combine targets, OARs and other normal tissues dosimetric data with the RO experience and clinical aims for a given pathology or the specific patient case. From the plan assessment point-of-view, treatment quality indexes describing the coverage[5] and conformity[6] of the target and/or the OARs sparing[7,8] for radiosurgery treatments have been proposed some decades ago. With the generalization of inverse planning and multicriteria optimization techniques, other comprehensive figures of merit associating different types of dosimetric score combinations to assess the plan quality were also proposed.[9-14] However, the RO clinical preferences were just included in the scoring design of plan quality indexes proposed by Schultheiss and Orton[9] and by Jain *et al.,*[10] through the application of statistical decision theory and decision analysis concepts, respectively.

Recently, a graphical method, named SPIDERplan, was developed to simultaneously assess and compare the quality of radiation therapy plans.[15] SPIDERplan considers the clinical aims associated with each of the structures of interest simultaneously weighting their relative importance.

The present work aims to assess whether it is possible to successfully relate SPIDERplan plan assessment with the RO clinical preferences SPIDERplan was applied for plan selection considering nasopharynx cancer cases and the study design included two phases. In the first phase, pairs of plans were blindly and independently evaluated by three ROs. Afterwards, the configuration of SPIDERplan, in terms of groups weights, was automatically performed using a mixed linear

programming model (MLPM) for preference elicitation. The plans that corresponded to the best SPIDERplan scores were then compared with the ROs plan choices. Intra- and inter-variability of the responses from the two phases were compared to conclude in what extent SPIDERplan was able to reproduce ROs choices. Finally, a threshold value for the score difference between competing plans, representing the value below which the plans can be considered as being of equivalent quality, was estimated.

3.2. Methods

3.2.1. Patient data

A sample of twenty nasopharynx cancer cases already optimized [16] was used for SPIDERplan clinical validation. Planning target volumes (PTVs) delineation and prescription followed the guidelines of the Radiation Therapy Oncology Group and the National Comprehensive Cancer Network. A simultaneous integrated boost prescription to be delivered in 33 fractions was assigned for all plans, where the PTVs, including gross disease, were prescribed with 70 Gy and the lymph nodes PTVs with a dose comprehended between 54.0 and 59.4 Gy according to the associated risk disease level. The tolerance criteria of the spinal cord, the brainstem, the optics structures (chiasm, optical nerves, retina and lens), the pituitary gland, the ears, the parotids, the oral cavity, the temporomandibular joints, the mandible, the oesophagus, the larynx, the brain, the thyroid and the lungs, also contoured by the RO, were defined according to the nasopharynx clinical protocol of the Radiotherapy Department of the Portuguese Oncology Institute of Coimbra (Table S3.1 of Supplementary material).

3.2.2. SPIDERplan description

SPIDERplan is a graphical method, developed by Ventura et al.,[15] that uses a scoring approach to assess and compare the quality of radiation therapy treatment plans. It aims to address the dose prescription objectives, defined for the clinical case/pathology. SPIDERplan configuration is structured in two phases: the processing of the plan data and the assessment of the plan quality.

In the processing phase, targets and OARs are divided into groups according to the clinical protocol or the RO preferences. A pre-defined relative weight is attributed to each group and each structure, representing the clinical priorities during the plan evaluation. For each plan, a score based on the pre-defined planning objectives is calculated for each structure to express the fulfilment level of the corresponding planning goal.

In the plan assessment phase, a customised radar plot displays all the score information. Plan evaluation can be done by displaying all structures and groups information in a Structures Plan Diagram and in a Group Plan Diagram, respectively. Global plan score is determined as the weighted sum of the structures individual scores as:

$$\text{Global plan score} = \sum_i w_{group(i)} \sum_j w_{struct(j)} Score_{struct(j)} \quad (3.1)$$

where $w_{struct(j)}$ and $Score_{struct(j)}$ are the relative weight and the score of structure j, respectively, and $w_{group(i)}$ the relative weight of group i. A partial group score based on the dose sparing of the structures that belong to that group is also calculated and represented in the Structures Group Diagram.

For the PTVs, the score was calculated according to a coverage criterion given by:

$$Score_{PTV} = \frac{D_{TC,PTV}}{D_{P,PTV}} \tag{3.2}$$

where $D_{TC,PTV}$ corresponds to the tolerance criteria for the PTV (in this case the dose in 98% of the PTV that should be at least 95% of the prescribed dose) and $D_{P,PTV}$ is the planned dose in the PTV. For the OARs, the score was set as:

$$Score_{OAR} = \frac{D_{P,OAR}}{D_{TC,OAR}} \tag{3.3}$$

where $D_{P,OAR}$ is the OAR planned dose and $D_{TC,OAR}$ is the tolerance dose for each OAR.

A score of 1 is therefore expected when the dose of a given structure (target or OAR) is equal to the respective tolerance value. If either target coverage or OAR sparing are better than the goal set by the RO, the score will be less than one.

3.2.3. SPIDERplan clinical validation

SPIDERplan clinical validation was performed in two-steps. In the first step, three ROs (RO1, RO2 and RO3), from three different national radiotherapy institutions, ranked and assessed the quality of the dose distributions of the selected cases. For each patient case, two plans (A and B), using coplanar optimized beam directions, were simultaneously presented to each RO. Based on the analysis of the dose distribution, the DVHs and the dose statistics, the ROs were asked to select the best plan of each of the 20 pairs of plans. If the plans were considered equivalent, both plans could be selected or rejected. For the evaluation of plan quality, each RO was asked to classify the plans as 'Good', 'Admissible with minor deviations' or 'Not Admissible'. Four control cases were randomly selected and randomly introduced in the list of patients to evaluate the intra-rater variability of each RO. These control cases used the same plans of patient cases #1, #4, #6 and #9 and were displayed to the RO in a swapped position (Plan A replaced plan B and vice-versa).

In the second step, SPIDERplan evaluation was applied to the same 20-paired cases. Two sets of structured groups were used to customize SPIDERplan response: a set of groups used by the local clinical protocol and a set of groups suggested by RTOG 0615,[17] named as CLIN and RTOG, respectively (Table 3.1). For the first set, SPIDERplan was successively applied using the CLIN group weights defined by the local RO (RO1) and the groups' weights automatically generated by the MLPM method ($CLIN_{aut}$), described in section 3.2.5. For the RTOG based groups, SPIDERplan evaluation just used the group weights defined by the MLPM method ($RTOG_{aut}$).

3.2.4. Statistical analysis

The intra-rater and inter-rater variabilities of ROs for plan selection and evaluation were statistically assessed by the Brennan-Prediger kappa (K_{B-P}) coefficient for nominal and ordinal variables, respectively.[18] The relative strength of the agreement is dependent on the K_{B-P} coefficient value and was classified using the scale proposed by Landis and Koch,[19] where for $K_{B-P} < 0.00$ the agreement is 'poor', for $0.00 \leq K_{B-P} \leq 0.20$ is 'slight', for $0.20 < K_{B-P} \leq 0.40$ is 'fair', for $0.40 < K_{B-P} \leq 0.60$ is 'moderate'; for $0.60 < K_{B-P} \leq 0.80$ is 'substantial and for $0.80 < K_{B-P} \leq 1.00$ the agreement is 'almost perfect'.

Table 3.1 - SPIDERplan group of structures defined locally according to RO aims (CLIN) and to RTOG guidelines (RTOG).[17]

CLIN		RTOG	
Groups	**Structures**	**Groups**	**Structures**
PTV	PTVs	**PTV**	PTVs
Critical	Brainstem	**Critical**	Retinas
	Spinal cord		Optical Nerves
			Chiasm
Optics	Lens		Brainstem
	Retinas		Spinal cord
	Optical Nerves		TMJ
	Chiasm		Mandible
DigestOral	Parotids	**Salivary**	Parotids
	Oral cavity		
	Larynx	**Other**	Brain
	Oesophagus		Lens
			Pituitary gland
Bone	Ears		Ears
	TMJ		Oral cavity
	Mandible		Larynx
			Oesophagus
Other	Brain		Thyroid
	Pituitary gland		Lungs
	Thyroid		
	Lungs		

TMJ – Temporalmandibular joint

3.2.5. Automatic weight determination by MLPM

When a decision-maker expresses his/her preferences by one out of two alternatives, the decision-maker is giving information regarding his/her preferences. It is possible to analyse these preferences, under a set of defined criteria, and to understand what is the importance that each one of the criterion has in the choice made. The importance of each criterion can be quantified by calculating a weight.

In this work, we have followed the methodology proposed by Srinivasan and Shocker.[20] Consider the multiattribute space defined by the different criteria that are taken into account by the decision-makers when making a choice. The decision-makers are the ROs. The multiattribute space dimension is equal to the number of different structure groups defined. Each attribute (criterion) is the corresponding structure group score. Each treatment plan is evaluated regarding the score of each one of the defined groups.

It is assumed that the ROs have a point in this multiattribute space that represents an ideal point: if a plan achieves, for each and all of the structures' groups, the score defined by this ideal point, then they will be satisfied with the plan. Furthermore, it is assumed that ROs will prefer plans that are as close as possible to this ideal point. The problem of finding a vector of weights (one weight for each group) that is able to represent the ROs preferences can be represented by a mixed linear programming model, where the decision variables will be the weights. The objective will be to guarantee that the preferred plans are closer to the ideal point than the non-preferred plans.

The following notation is used:

Parameters

- $J = \{1, \dots, n\}$ represents the set of plans that are going to be evaluated by the decision-maker
- $P = \{1, \dots, t\}$ represents the *t* dimensions in which each of the plan is evaluated (each plan is evaluated considering each one of the groups so that *t* is equal to the number of groups considered)
- $Y_j = \{y_{jp}, j \in J, p \in P\}$ represents the score of the j^{th} plan for structure group *p*
- $\Omega = \{(j,k), j,k \in J\}$ represents the set of all ordered pairs (*j*,*k*) resulting from the comparison of plan *j* and plan *k* if *j* is preferred to *k*.

Decision variables

- $X = \{x_p\}, p \in P$ represents the ideal point to be determined
- $W = \{w_p\}, p \in P$ represents the weight of each one of the *t* dimensions (the weight that each group should have in the calculation of the global score).

It is possible to calculate the distance between each plan $j \in J$ and the ideal point *X*. In this work we have chosen the Euclidean distance, meaning that:

$$d_j = \sqrt{\sum_{p \in P} \left(y_{jp} - x_p\right)^2}, j \in J \tag{3.4}$$

If plan *j* was evaluated as being better than plan *k* then this should mean that $d_j < d_k, \forall (j,k) \in \Omega$. The problem can then be described as: given $Y_j, \forall j \in J$ and Ω, find *X* and *W* such that conditions (4) are violated as minimally as possible. It is thus necessary to define what is meant by "violating as minimally as possible". In this work this has been defined as finding *X* and *W* such that the number of violations of equation 3.4 is minimized.

Srinivasan and Shocker[20] showed that this problem can be represented by the following mixed linear programming model:

$$\text{Minimize} \sum_{(j,k) \in \Omega} \delta_{jk} \tag{3.5}$$

subject to:

$$\sum_{p \in P} (y_{kp}^2 - y_{jp}^2) w_p - 2 \sum_{p \in P} \left(y_{kp} - y_{jp}\right) v_p + \delta_{jk} M \geq 0, \forall (j,k) \in \Omega$$

$$\sum_{p \in P} \sum_{(j,k) \in \Omega} (y_{kp}^2 - y_{jp}^2)\, w_p - 2 \sum_{p \in P} \sum_{(j,k) \in \Omega} (y_{kp} - y_{jp})\, v_p = 1$$

$$w_p \geq 0, \forall p \in P$$

$$\delta_{jk} \in \{0,1\}, \forall (j,k) \in \Omega$$

where *M* represents an arbitrarily large positive number.

Table 3.2 - Intra-rater variability and inter-rater variability.

	Intra-rater variability			Inter-rater variability		
	Radiation oncologist	**Plan selection**	**Plan evaluation**	**Radiation oncologist**	**Plan selection**	**Plan evaluation**
K_{B-P} coefficient	RO1	0.25 **Fair**	0.78 **Substantial**	All ROs	0.38 **Fair**	0.63 **Substantial**
	RO2	0.25 **Fair**	0.89 **Almost perfect**			
	RO3	-0.13 **Poor**	0.78 **Substantial**			

In the present work, $\Omega = \{(j,k), j,k \in J\}$ is built from the combined result of the evaluation made by the three different ROs. The objective was not to find *X* and *W* that would be RO dependent, but instead to find *X* and *W* capable of representing global preferences. This was achieved by applying a majority rule: (j,k) belongs to Ω if *j* was preferred to *k* by the majority of ROs.

3.3. Results

3.3.1. Plan selection and plan evaluation performed by the radiation oncologists

The results of plan selection and plan evaluation performed by the ROs are displayed in the first four columns of Figure 3.1. For each comparison, the plan selected by the RO is represented by a filled square, the plans evaluated as 'Good' by a green square, the plans evaluated as 'Admissible with minor deviations' by a yellow square and the plans considered 'Not Admissible' by a red square. The control cases are represented in Figure 3.1 below the correspondent patient case but were randomly presented to the ROs. More than two-thirds of the plans were evaluated as 'Good' by all ROs evaluations. Globally, all plans presented to the ROs, have high-quality dose distributions, but still 19% of the plans were evaluated as 'Admissible with minor deviations' and 8% as 'Not Admissible'. For patients #4, #9, #18, #20 and the control case #c9 both plans A and B were selected by RO3, meaning that both plans were considered of equivalent quality. For patient #18, plans A and B were evaluated by RO2 and RO3 as 'Not Admissible' and for patient #20, plan A was differently evaluated by all ROs.

The intra-rater and inter-rater variabilities analyses were assessed through the calculation of K_{B-P} coefficients displayed in Table 3.2. The intra-rater variability was computed for each RO by comparing plan selection and plan evaluation of patients #1, #4, #6, #9 with the corresponding control cases. RO1 kept his plan selection for patients #1 and #9 and RO2 for patients #1, #6 and #9, which conducted to a fair agreement (K_{B-P} = 0.25). RO3 was the clinician with higher variability in plan selection, with a K_{B-P} = -0.13, as only for patient #9 he selected the same plan.

For plan evaluation, the intra-rater variability was unquestionably higher than for plan selection. When asked to grade plan quality, all ROs presented at least a substantial agreement between the first and the second evaluation (K_{B-P} = 0.78). RO1 and RO3 evaluated the quality of plans A and B as equally 'Good' for patients #1, #4 and #9 and for patients #1, #4 and #6, respectively. For RO2, the agreement was almost perfect (K_{B-P} = 0.89), as only for plan A of patient #6 the plan evaluation was not coincident.

For the quantification of inter-rater variability, the control cases were not considered. As in the previous analysis, the agreement between the ROs in plan selection (K_{B-P} = 0.38 - fair) was not as good as for plan evaluation (K_{B-P} = 0.63 - substantial). Only in 10/20 patient cases all ROs agreed

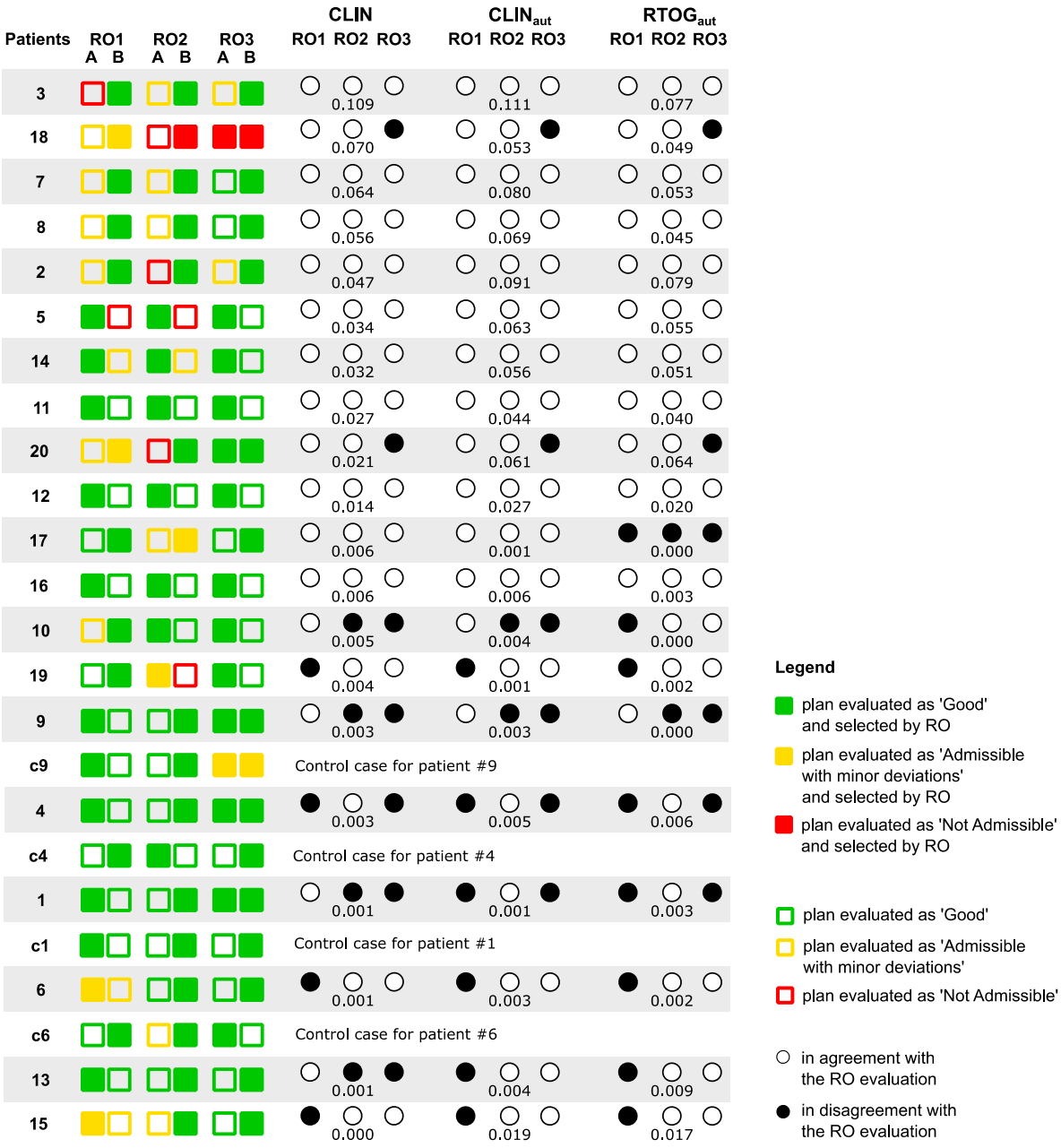

Figure 3.1 - Results of the plan selection and evaluation of the selected nasopharynx cases by RO1, RO2 and RO3 (squares) and of the agreement and disagreement between SPIDERplan evaluation and the corresponding RO plan selection for CLIN, CLIN$_{aut}$ and RTOG$_{aut}$ group of structures (white and black circles, respectively). The difference between SPIDERplan global plan scores of plan A and B for the correspondent group of structures is shown immediately below. The patient cases were sorted in descending order of the CLIN set score difference. To facilitate the graphical comparison between the control cases and correspondent patient cases, the representation of plan A and B in the control cases was swapped.

in the selection of the best plan. While the agreement between RO2 and RO3 was high, the agreement between RO1 and RO2 and between RO1 and RO3 was about 50% (12/20 and 10/20, respectively). For plan evaluation, most of the plans 39/40) had two or more coincident RO ordinal assessments and more than one half (21/40) the same classification by all ROs. Again, it was between RO2 and RO3 that there was the higher number of plan evaluation agreements (29/40). For RO1, the number of plans with the same evaluation as RO2 and RO3 was almost equal (25/40 and 26/40, respectively).

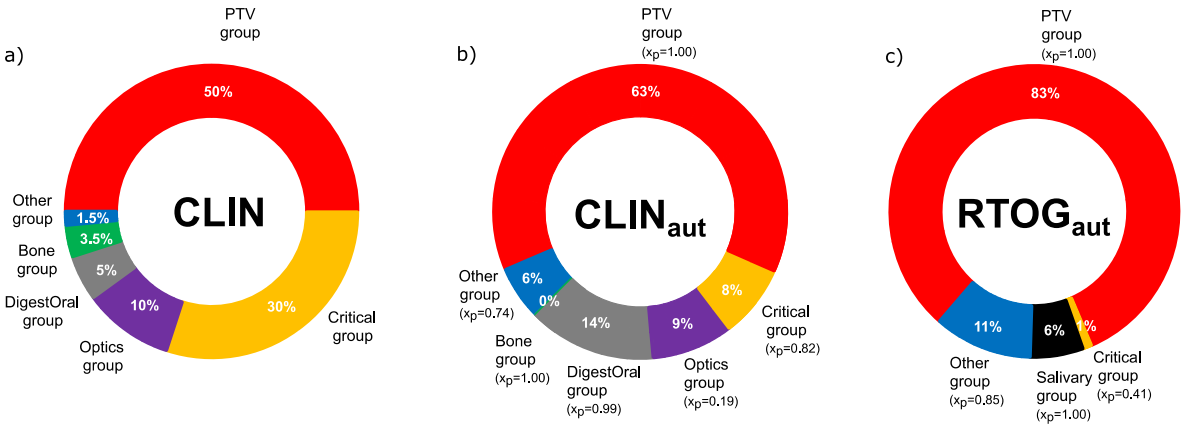

Figure 3.2 - Group weights for the CLIN, the CLINaut and the RTOGaut sets. The x_p value represents the satisfaction point of the decision-maker.

3.3.2. MLPM group weight determination

The group weights of CLIN, $CLIN_{aut}$ and $RTOG_{aut}$ group sets are shown in Figure 2. The CLIN group weights (Figure 3.2a) were defined according to the local clinical nasopharynx protocol and the $CLIN_{aut}$ (Figure 3.2b) and the $RTOG_{aut}$ (Figure 3.2c) were determined using the MPLM method. Values for vector X are also shown. Low values for x_p (namely less than one) mean that the ROs are, in reality, being more demanding with the corresponding group (finding plans satisfactory only if a low score is attained) than when greater x_p values are obtained. For the CLIN set, the PTV and the Critical groups received the higher weights (50% and 30%), while the Bone and the Other group the lowest weights. For the $CLIN_{aut}$ set, the groups' weights changed considerably. The PTV group presented the higher weight, but the Critical group weight was lower than those of the DigestOral and Optics groups. However, the x_p associated with the Critical group is less than one, showing that this is, in fact, an important structure group and the ROs will not, probably, be satisfied with plans that simply comply with the planning goal, expecting to see better organ sparing. For the Bone group the calculated weight value is 0% and the x_p is equal one, meaning that the dose received by the structures of this group will only have to comply with the prescribed value for the RO to be satisfied. For the $RTOG_{aut}$ set, the PTV group achieved the highest weight, while the lowest was computed for the Critical group. This is the group with the lowest x_p, certifying that the ROs value a low score from the structures of the group even if a relative low weight has been assigned to it.

3.3.3. SPIDERplan evaluation

SPIDERplan global plans scores were computed for all patient cases using the groups and the weights from CLIN, $CLIN_{aut}$ and $RTOG_{aut}$ sets. Its response accuracy is graphically displayed in Figure 3.1, where the agreement between the selection based on SPIDERplan global plan score and the clinical plan choice is represented by white circles and the disagreement by black circles. The patient cases are sorted by descent order of the score difference between plan A and B for the CLIN configuration. A complete agreement between SPIDERplan selection for all sets and all ROs was obtained in 9/20 of the patient cases and at least two agreements per set in 14/20 of the cases. It can be seen that the higher the global score difference between the two paired plans, the better the SPIDERplan results agree with the ROs choice and also the closer the agreement among the three ROs (e.g. patients #2, #3, #7 or #8). Globally, the agreement in plan selection between SPIDERplan and all ROs was high (>45/60), resulting in an inter-rater variability of substantial to moderate agreement. The plan selection agreement between SPIDERplan and RO1 was higher for the CLIN set than for the remaining sets whose group weights were automatically determined by the MLPM method. By contrary, the percentage of plan selection agreement between SPIDERplan and RO2

and RO3 was higher for the $CLIN_{aut}$ and $RTOG_{aut}$ sets than for CLIN where an inter-variability of almost perfect agreement was obtained. Nevertheless, the global percentage of agreement and the intra-rater variability (all ROs) was almost equal for all sets (45/60, 46/60 and 44/60). A total disagreement between SPIDERplan response and all ROs was just obtained for patient #17, when $RTOG_{aut}$ set was used (three black circles). In this case the difference in plan quality between the two plans is so small (0.0008) that in fact it is irrelevant which is the plan selected for treatment. Thus, a threshold value for the score difference between two plans was defined. This threshold, estimated as 0.005, represents the value below which two plans are judged as dosimetrically equivalent. Considering now this threshold value, the agreement between SPIDERplan and RO plan selection increases from 45/60 to 55/60 cases. The plan choices made by SPIDERplan that fail this threshold value where #15 ($CLIN_{aut}$ and $RTOG_{aut}$) #18 and #20. For patients #18 and #20, RO3 could not make a choice between plans A and B, so no agreement could ever be found, anyway. For patient #15, RO1 was not in agreement with the other ROs and, as stated in section 3.2.5, his/her choice was thus not considered for the automatic determination of the group weights by MLPM method (the majority decision was considered).

3.4. Discussion

SPIDERplan is a graphical plan assessment tool developed for supporting the clinical choice of the best plan for treatment delivery. The evaluation of the quality of the dose distribution is done by combining the graphical analysis provided by customised radar plots with a scoring index. Weighted groups of structures reflecting the RO clinical preferences for a given pathology or case must be defined and validated prior to starting using SPIDERplan in the clinical routine. In this study, SPIDERplan was clinically validated for the nasopharynx pathology by comparing the plan evaluation made by three ROs with the SPIDERplan score results.

Twenty nasopharynx cases with high-quality dose distributions were blindly evaluated and ranked by three ROs from different institutions. Four control cases were randomly selected from the list of these patient cases with the most similar plans and randomly presented to the ROs without their knowledge. The choice of the best treatment may be influenced by different factors (individual characteristics of the decision-maker and the patient, contextual factors, specific technical criteria), that can introduce some inter- and even intra-variabilities in the decision of the ROs. In this work, some of these factors were surpassed given the retrospective and anonymous character of the selected patient cases sample. The ROs assessed the plans following their own institutional protocol guidelines, using traditional treatment plan evaluation tools (dose distribution visualization, DVH and dose statistics analysis) and embedding into the final decision their personality and clinical experience. On average, just three hours were spent by each RO to complete the assessment of all the cases. On one hand, the continuous time slot dedicated to this task may have negatively influenced the consistency of his/her evaluation, as the repetition of cases assessment may have caused some inattention/fatigue to, at least, the last evaluated cases. Probably that was the reason for RO3 not having been able to select the best plan in patient cases #18 and #20 (the last ones) even with high score differences. On the other hand, it assured, in principle, the use of more consistent criteria during the process.

For plan selection, the intra-rater variability analyses presented lower $K_{B\text{-}P}$ coefficients than the inter-rater variability analysis, meaning that the agreement between different ROs was better than between themselves. This low agreement may be a result of the high-quality of the dose distributions of the control cases and also the similarity of the plans in the control pairs. This choice of control plans avoided the perception, by the ROs, that control cases have been introduced because it was harder to acknowledge that they were comparing for the second time a pair of plans already considered. Of course, the reduced number of cases used for this intra-variability analysis statistically influenced the intra-rater agreement result. From the intra-rater and inter-rater variabilities

analyses, it is also evident that the agreement between SPIDERplan and the ROs for plan evaluation was much higher than for plan selection. This finding has a direct correspondence with RO appraisal in the clinical routine as it is usually much easier for clinicians to agree upon the quality of the plans (saying if they are 'Good', 'Admissible with minor deviation' or 'Not Admissible') than it is to select the plan for treatment.

SPIDERplan response accuracy was tested using two sets of groups of structures and two methods to establish group weights. The composition and the weights of the CLIN set were defined according to the local nasopharynx protocol and the RO1 clinical preferences. SPIDERplan response reproduced ROs selection in 75% of the cases. It is interesting to note that the disagreement between SPIDERplan and ROs in plan selection was in-line with the inter-rater variability analysis between the three ROs corroborating the accuracy of SPIDERplan assessment. The small number of disagreements occurred for very low score differences between plans A and B. A threshold, in terms of score difference between plans, below which the choices were considered in agreement with ROs was thus defined. Values below this threshold reflected the difficulty showed by the ROs to choose the best plan when they were very similar. This threshold can be seen as a measure of the uncertainty associated with SPIDERplan plan assessment and also as a justification for the intra-rater and inter-rater variability of the ROs plan selection and evaluation.

The definition of groups of structures according to their clinical importance and the corresponding assignment of importance weights is a non-trivial task for ROs. In the daily routine ROs acceptability criteria and preferences are qualitatively incorporated in the process of selection and approval of the best plan and not based on a quantitative value reflecting the importance of each structure. Therefore, a MLPM method was applied to automatically determine the weights of each group ($CLIN_{aut}$). An alternative group of structures was also defined following RTOG 0615 guidelines, and the respective weights calculated using the same automated method ($RTOG_{aut}$). Compared to $CLIN_{aut}$ and $RTOG_{aut}$, SPIDERplan performance was similar to that of the CLIN set except for RO1 where the agreement for the two new sets of structures groups decreased. This is to be expected as the CLIN set was defined by the clinical protocol followed by RO1.

The group weights determined by the MLPM method considered the clinical choices of all ROs in plan selection. Indeed, the $CLIN_{aut}$ set presented a somewhat different configuration from the CLIN set (Figure 3.2). The unexpected low weight of the Critical group may have grounds on the automated method itself. The low score values (high sparing) and the associated low variability presented by this group give room to the MLPM algorithm to confer more importance to groups with scores with higher values and higher variability, such as the Optics, the DigestOral or the Salivary groups. Nevertheless, the lower x_p values showed that, although the importance of this group in the plan evaluation was not so high as initially thought, the ROs required that this group of structures presented higher levels of sparing to be satisfied.

SPIDERplan was configured for the nasopharynx pathology using a local group set and weights, a local group definition with weights automatically calculated, and thirdly using a group definition based on international guidelines with automated group weights. The performance of SPIDERplan against the ROs choices, for all sets of group weights, was similar to the inter-rater variability obtained between the ROs clinical evaluations. The flexibility in plan evaluation and comparison provided by different group weights enables the possibility to adapt with confidence any of these SPIDERplan configuration options in the clinical practice. For other pathologies, any of these SPIDERplan configuration methods could be followed. It is possible to define the structure groups and weights resorting to the local RO team clinical protocols and preferences. Alternatively, it is possible to automatically elicit these weights through the analysis of the comparison of different plans using a pool of patient cases considering either groups locally defined or in accordance to international guidelines.

3.5. Conclusions

In this work, the evaluation of SPIDERplan was successfully linked to the plan evaluation of three ROs from three Portuguese radiation therapy departments for the nasopharynx pathology using three different configuration methods. SPIDERplan plan evaluation agreed with most of the ROs assessments and presented an equivalent variability to that of the ROs choices. To handle decision uncertainty when the quality of the plans is very similar, a threshold value was determined for the score differences between the plans, below which the plans are considered of equivalent quality.

For the nasopharynx pathology, any of the configurations tested, i.e., based on local preferences or automatically determined from a pool of testing cases, can be used in SPIDERplan without loss of accuracy. For other pathologies, any of these configuration methods can/could be set before starting using SPIDERplan in clinical practice.

3.6. Acknowledgements

The authors would like to express their gratitude to Tânia Serra for the valuable support in the development of this work and to Andreia Hall for the technical support in the statistical analysis.

This work was supported by project grant POCI-01-0145-FEDER-028030 and by the Fundação para a Ciência e a Tecnologia (FCT) under project grant UID/Multi/00308/2019.

No potential conflict of interest nor any financial disclosures must be declared.

3.7. References

1. ICRU. International Commission on radiation units and measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Geneva; 2010. J ICRU 10(1):1–106.
2. Thieke C, Küfer KH, Monz M, Scherrer A, Alonso F, Oelfke U et al. A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. Radiother Oncol. 2007;85(2):292–298.
3. Miller GA. The magical number seven plus or minus two: some limits to our capacity of processing information. *Psychol Rev.* 1956;63(2):81-97.
4. Glatzer M, Panje C, Sirén C, Cihoric N, Putora P. Decision Making Criteria in Oncology. *Oncology.* 2018:1-9.
5. Lomax NJ, Scheib SG. Quantifying the degree of conformity in radiosurgery treatment planning. *Int J Radiat Oncol Biol Phys*. 2003;55(5):1409-1419.
6. Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. *J Neurosurg.* 2000;93:219-222.
7. Baltas D, Kolotas C, Geramani K, Mould RF, Ioannidis G, Kerchidi M. A conformal index (COIN) to evaluate implant quality and dose specification in brachytherapy. *Int J Radiat Oncol Biol Phys.* 1998;40(2):515-524.
8. Menhel H, Levin D, Alezra D, Symon Z, Pfeffer R. Assessing the quality of conformal treatment planning: a new tool for quantitative comparison. *Phys Med Biol*. 2006;51(20):5363-5375.
9. Schultheiss TE, Orton CG. Models in radiotherapy: definition of decision criteria. *Med Phys*. 1985;12(2):183-187.
10. Jain NL, Kahn MG, Drzymala RE, Emami BE, Purdy JA. Objective evaluation of 3-D radiation treatment plans: a decision-analytic tool incorporating treatment preferences of radiation oncologists. *Int J Radiat Oncol Biol Phys*. 1993;26(2):321-333.
11. Miften MM, Das SK, Su M, Marks LB. A dose–volume-based tool for evaluating and ranking IMRT treatment plans. *J Appl Clin Med Phys*. 2004;5(4):1-14.
12. Leung LT, Kan MK, Cheng AK, Wong WH, Yau CC. A new dose–volume-based Plan Quality Index for IMRT plan comparison. *Radiother Oncol*. 2007;85(3):407-417.
13. Nelms BE, Robinson G, Markham J et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Pract Radiat Oncol*. 2012;2(4):296-305.
14. Alfonso J, Herrero M, Núñez L. A dose-volume histogram based decision-support system for dosimetric comparison of radiotherapy treatment plans. *Radiat Oncol*. 2015;10:263.

15. Ventura T, Lopes MC, Ferreira BC, Khouri L. SPIDERplan: a tool to support decision making in radiation therapy treatment plan assessment. *Rep Pract Oncol Radiother*. 2016;21(6):508–516.
16. Ventura T, Rocha H, Ferreira BC, Dias J, Lopes MC. Comparison of two beam angular optimization algorithms guided by automated multicriterial IMRT. *Phys Medica*. 2019, 64:210-221.
17. Lee N, Garden A, Kim J et al. A phase II study of concurrent chemotherapy using three-dimensional conformal radiotherapy (3D-CRT) or intensity modulated radiation therapy (IMRT) + bevacizumab (BV) for locally or regionally advanced nasopharyngeal cancer. NRG Oncology - RTOG 0615. 2014.
18. Santiago MJ. Métodos de estimação de fiabilidade e concordância entre avaliadores. Master Thesis. 2016.
19. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
20. Srinivasan V, Shocker AD. Linear programming techniques for multidimensional analysis of preferences. *Psychometrika*. 1973;38(3):337-369.

3.8. Supplementary material

Table S3.1 - Tolerance dose criteria for PTVs and OAR.

Structures **Name**	**Tolerance criteria**
PTVs	$D_{98\%} \geq D_{p,95\%}$
Spinal cord (SPNLCORD)	$D_{max} \leq 45Gy$
Brainstem	$D_{max} \leq 54Gy$
Chiasm	$D_{max} \leq 55Gy$
Left optical Nerve (OPTNRVL)	$D_{max} \leq 55Gy$
Right optical Nerve (OPTNRVR)	$D_{max} \leq 55Gy$
Left retina (RETINAL)	$D_{max} \leq 45Gy$
Right retina (RETINAR)	$D_{max} \leq 45Gy$
Left lens (LENSL)	$D_{max} \leq 6Gy$
Right lens (LENSR)	$D_{max} \leq 6Gy$
Left parotid (PAROTIDL)	$D_{mean} \leq 26Gy$
Right parotid (PAROTIDR)	$D_{mean} \leq 26Gy$
Oral cavity (ORALCAV)	$D_{mean} \leq 35Gy$
Oesophagus	$D_{mean} \leq 40Gy$
Larynx	$D_{mean} \leq 45Gy$
Left temporal mandibular junction (TMJL)	$D_{max} \leq 66Gy$
Right temporal mandibular junction (TMJR)	$D_{max} \leq 66Gy$
Mandible	$D_{max} \leq 66Gy$
Left ear canal (EARL)	$D_{mean} \leq 45Gy$
Right ear canal (EARR)	$D_{mean} \leq 45Gy$
Brain	$D_{max} \leq 54Gy$
Pituitary Gland (PITUITARY)	$D_{max} \leq 60Gy$
Thyroid	$D_{mean} \leq 27.5Gy$
Left lung (LUNGL)	$D_{mean} \leq 5Gy$
Right lung (LUNGR)	$D_{mean} \leq 5Gy$

D_p – Prescribed dose, D_{max} – maximum dose, D_{mean} – mean dose

chapter 4

Comparison of two beam angular optimization algorithms guided by automated multicriterial IMRT

Physica Medica, 2019, volume 64, pages 210-221

T Ventura[1,2,3], H Rocha[3,4], BC Ferreira[3,5,6], J Dias[3,4], MC Lopes[1,2,3]

[1]Physics Department, University of Aveiro, Aveiro, Portugal
[2]Medical Physics Department, IPOCFG, EPE, Coimbra, Portugal
[3]Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal
[4]Economy Faculty of University of Coimbra and Centre for Business and Economics Research, Coimbra, Portugal
[5]School Health Polytechnic of Porto, Porto, Portugal
[6]I3N, Physics Department, University of Aveiro, Aveiro, Portugal

Highlights

- Comparison of two beam angle optimization strategies for head-and-neck IMRT for coplanar and non-coplanar geometries.
- Average results highlight global tendencies for beam incidences and case-by-case analysis features specific advantages.
- Beam angular distributions and plan quality of IMRT plans were assessed through graphical tools.

Abstract

Purpose: To compare two beam angle optimization (BAO) algorithms for coplanar and non-coplanar geometries in a multicriterial optimization framework.

Methods: 40 nasopharynx patients were selected for this retrospective planning study. IMRT optimized plans were produced by Erasmus-iCycle multicriterial optimization platform. Two different algorithms, based on a discrete and on a continuous exploration of the space search, algorithm *i* and *B* respectively, were used to address BAO. Plan quality evaluation and comparison were performed with SPIDERplan. Statistically significant differences between the plans were also assessed.

Results: For plans using only coplanar incidences, the optimized beam distribution with algorithm *i* is more asymmetric than with algorithm *B*. For non-coplanar beam optimization, larger deviations from coplanarity were obtained with algorithm *i* than with algorithm *B*. Globally, both algorithms presented near equivalent plan quality scores, with algorithm *B* presenting a marginally better performance than algorithm *i*.

Conclusion: Almost all plans presented high quality, profiting from multicriterial and beam angular optimization. Although there were not significant differences when average results over the entire sample were considered, a case-by-case analysis revealed important differences for some patients.

Keywords: beam angular optimization, multicriteria optimization and plan quality assessment

4.1. Introduction

Intensity-modulated radiation therapy (either static/ dynamic IMRT or volumetric modulated arc therapy, VMAT) is becoming the standard technique in radiation therapy. Non-uniform intensity fields from multiple directions are used to generate high conformal dose distributions to the tumour. For the standard approach of IMRT/VMAT treatment planning, the plan objectives are usually described by physical or biological descriptors that are typically incorporated in an objective function that will guide the fluence map optimization (FMO) procedure by scoring the goodness of the plan.[1] Searching methods such as linear least squares,[2] gradient descent[3] or simulated annealing[4] are used to compute the intensity pattern that provides the best possible trade-off between conflicting planning goals. A trial-and-error iterative manual tuning of plan parameters (like weights, objectives or beam angles) may be necessary to achieve an acceptable plan. One important difficulty in this iterative process is the fact that it is not possible to know the impact that changing one given parameter will have in the treatment plan, or what are the interdependencies that exist between the different parameters. This iterative process is thus mainly guided by the empirical knowledge of the planner. Furthermore, it is also not possible to link the parameters' values with the desired clinical planning goals. As a result, it is not possible to guarantee that the trial-and-error optimization process will lead to an optimal plan. This process is also more or less time-consuming depending on the case complexity and mostly on the planner skills.[5]

Multi-criteria optimization (MCO) methods come up as a natural option to support the IMRT treatment planning decision making process. Despite presenting less manual interaction, these methods still require some triggering from the planner to obtain good dose distributions. Multiple objectives, resulting from the goals assigned to targets and normal tissues, are simultaneously maximized (or minimized), instead of a single objective function usually applied for the standard approach of inverse planning. As most of the times it is not possible to find a single feasible solution that is simultaneously the best one for every objective,[6,7] a set of optimal plan solutions containing the best possible trade-offs between objectives are presented to the decision maker.

Beam angle selection plays also an important role in IMRT optimization. An appropriate beam angle assembly choice, based on a mathematical criterion rather than on the planner experience or on equidistant coplanar arrangement solutions, may lead to important enhancements in the final plan dose distribution solution.[8] Plan quality improvements can be even more significant if non-coplanar directions are included in the optimization process.[9] The use of non-coplanar beam angles in VMAT was also proposed to combine the benefits of arc therapy, such as short treatment times, with the benefits of noncoplanar IMRT plans, such as improved organ sparing. Selected non-coplanar beam angle directions can be used as anchor points of the arc therapy trajectory[10] which further validate the interest of studying the selection of optimal non-coplanar beam angle incidences. Mathematically, the beam angular optimization (BAO) problem can be described as a highly non-convex multi-modal optimization problem with many local minima,[11-13] which ideally requires methods with few computing iterations and able to avoid getting trapped in local minimum. BAO solutions are often non-intuitive, so it is important to use optimization approaches that are reliable considering their capacity of delivering optimal solutions.[14]

The BAO problem can be addressed in two different ways. One possibility is to decouple the beam angle selection from the FMO and solve the two problems sequentially. In this case, BAO is driven by geometrical measures (e.g., beam's-eye view metrics) or by methods that require prior knowledge of the problem.[8,15] Although computationally efficient, these methods do not fully account for the interplay of beam angles and beamlet weights and plan solution optimality cannot be fully guaranteed. Another possibility is to simultaneously address BAO and FMO problems. FMO optimal

solutions are used to assess the beams set plan quality during the BAO.[13] Two completely different mathematical formulations of the BAO problem can be found in the literature: a combinatorial formulation, where the interval of possible gantry angles, [0º, 360º[, is discretized into evenly spaced angles (e.g. {0º,10º,…,350º} for an angle increment of 10º) and a continuous BAO formulation where all possible gantry angles in the interval [0º, 360º[are considered. For the first approach, a combinatorial search for the best ensemble of beams over a discretized space search defined with all possible beam incidences can be done using heuristic methods.[12,13,16-20] However, as this formulation is considered a nondeterministic polynomial time hard problem,[21] alternative combinatorial approaches have also been developed. The iterative BAO methods wherein the beams are iteratively subtracted[22] or added[23,24] to a beam ensemble, decreasing significantly the possible number of combinations, are one the most well-known examples. BAO methods based on the continuous exploration of the solutions search space have been explored as an alternative to the combinatorial BAO, namely using pattern search methods,[9,25] or considering a parallel multistart derivative-free optimization framework.[9,26]

In the present work, the BAO problem is addressed using two algorithms, one belonging to the discrete combinatorial type[23] and the other to the continuous space search approach optimization class.[25] Both algorithms use the FMO objective function to guide the BAO process and the two problems are simultaneously addressed. The BAO algorithms were compared over a set of 40 nasopharyngeal cancer (NPC) clinical cases. The correspondent IMRT plans were optimized by an automated MCO calculation engine developed by Breedveld *et al.*[27] Coplanar and non-coplanar geometry scenarios and different number of beam incidences in treatment delivery were considered in this retrospective planning study. The plans were assessed and compared using SPIDERplan,[28] that evaluates the quality of the dose distribution through an intuitive graphic representation and an associated score function that are based on dose prescription aims.

4.2. Materials and methods

4.2.1. Patient data

Forty NPC cases, stages T1 - T4; N1 - N3a/N3b, treated with IMRT were selected for this study. Planning target volumes (PTV) delineation and dose prescriptions were based on the Radiation Therapy Oncology Group and the National Comprehensive Cancer Network guidelines. All cases had simultaneous integrated boost prescription delivered in 33 fractions, 70.0 Gy to the tumour PTV and a dose ranging between 54.0 Gy and 59.4 Gy according to the associated risk disease level to the lymph nodes PTVs. Some patients had one or more adenopathies that were also prescribed with 70.0 Gy. Spinal cord, brainstem, retinas, lens, optical nerves, chiasm, pituitary gland, ears, parotids, oral cavity, temporomandibular joints, mandible, oesophagus, larynx, brain, thyroid and lungs were also contoured by the radiation oncologist, as shown in Figure S4.1 and S4.2 in the Supplementary Material. The organs-at-risk (OAR) tolerance doses were established in agreement with the institutional protocol for the nasopharyngeal pathology (the reader is referred to the last column of Table S4.1 in the Supplementary Material).

4.2.2. Plan generation and optimization

FMO for all plans was handled by Erasmus-iCycle IMRT multicriterial optimization engine.[27] Guided by a wish-list, containing clinical constraints and prioritized objectives, a constraint-based method, 2pεc method, is used to automatically generate a single Pareto optimal IMRT solution for a given set of beams.[27] Beamlets size are set to 10x10 mm^2 with 30 mm of scatter radius for IMRT optimization and 15 mm for the BAO algorithm implemented within Erasmus iCycle. Dose calculation is performed using a pencil-beam dose algorithm with equivalent path length inhomogeneity corrections and no fluence segmentation is performed during or after multicriterial optimization. The

wish-list template built for NPC cases, was previously customized according to the established institutional clinical tolerance criteria using five test cases (Table S4.2 in the Supplementary Material). It contained clinical constraints and prioritized objectives that were divided in two optimization levels, according to the clinical tolerance doses, the proximity between PTVs and OARs and its impact on the dose distribution. This configuration, with a progressive dose optimization structure, is appropriate for complex sites, like the NPC cases. It intends to avoid possible limitations that may arise when a dose value achieved in an OAR with a high priority restricts the optimization of another one with a lower priority. The objectives associated with the PTVs were assigned with the Logarithmic Tumour Control Probability (LTCP) function, which is regulated by a cell sensitivity parameter (α). For this study, an α value of 0.75 was applied to guarantee that at least 98% of the PTV volume receives 95% of the prescription dose (D_p). To allow the minimization of lower prioritized objectives, a LTCP sufficient value of 0.5 was defined. The remaining objectives were defined according to the OAR type. For organs with a serial architecture a maximum dose objective was defined. For parallel architectures a mean dose objective was applied. Also, the dose of non-vital OARs, such as lens, optics, retinas, brain or pituitary gland, was minimized using the generalized Equivalent Uniform Dose (gEUD) function with a value of the tissue-specific parameter that describes the volume effect (*a*) equal to 12.

4.2.3. Beam angular optimization

For BAO of coplanar and non-coplanar beam geometries two different methods were tested. Both methods used the optimal FMO value to guide the BAO process and the two problems were jointly solved. The number of beams was defined a priori.

In the first method, developed by Breedveld *et al.*[23,29] and implemented within Erasmus-iCycle, BAO is integrated in the plan optimization framework, considering a discretization of the search space, i.e. the gantry angle [0º,360º[and the couch angle [-90º,90º] intervals are discretized into equally spaced angles with an angle increment of 5º. Plan generation is done by iteratively adding into the plan beams with an optimal orientation. For a given beam arrangement, all possible candidates will be combined with the beams already selected for the plan and the candidate beam that achieves the lowest score of the fluence optimization problem is added. New beams will be sequentially added until the maximum number of beams initially defined is reached. The way the iterative BAO algorithm works is described in Algorithm *i*:

Algorithm *i*

Initialization

- Define n, the number of beams;
- Define $\Theta=\left\{\left(\theta_i,\phi_j\right),i=1,\ldots,N,\ j=1,\ldots,M\right\}$ as the discrete set of possible beam directions;
- Set $x^0:=\{\ \}$ as the set of best beam directions;
- Set $k:=1$;

Iteration

1. Add each direction of Θ (that does not belong to x^{k-1}), one at a time, to x^{k-1} and compute the optimal FMO value of the corresponding beam direction ensemble;
2. $x^k:=\{x^{k-1},(\theta^k,\phi^k)\}$, where (θ^k,ϕ^k) is the beam direction of Θ that added to the directions of x^{k-1} leads to the best optimal FMO value;
3. $k:=k+1$;
4. If $k\leq n$ return to step 1 for a new iteration.

The second approach, developed by Rocha *et al.,*[25] explores the continuous BAO search space using a pattern search method, meaning that there is no need to do any kind of discretization. These class of methods are directional direct search methods and thus do not require the use of derivatives to minimize the objective function. To assure a more effective search for the best objective function local minimum, a set of points spanning as much as possible the entire search space is defined in a preliminary step of the pattern search optimization. Thus, the objective function values of a set of plans with equally spaced orientations that span the entire beam angle search space are determined. The pattern search optimization is organized around two steps: the search step and the poll step. It starts with the search step where any (global) strategy can be used to improve the best objective function value. In this implementation, minimum Frobenius norm quadratic models were used to perform a search over the whole search space.[30] These quadratic models are based on the beam angle sets already considered. If the corresponding objective value is lower than the best objective function minimum value, the search step was successful, and it is repeated. Otherwise the optimization method proceeds to the poll step, where the current best solution is locally improved using the concept of positive basis. If this step fails to obtain a decrease in the objective function value, the step-size parameter is reduced. If the step-size becomes smaller than the defined limit the process stops, otherwise a new loop of the algorithm is performed starting a new search step. When the maximum number of iterations is reached, the pattern search optimization will stop. BAO using pattern search is described in Algorithm *B*:

Algorithm *B*

Initialization

- Define n, the number of beams;
- Choose $x_i \epsilon [0, 360]^n$, i=1,…,N, the starting beam direction ensembles;
- Compute $f(x_i)$, i=1,…,N, the optimal FMO value for each of the initial points;
- Set $x_i^* := x_i$, i=1,…,N, as the best points and $f_i^* := f(x_i^*)$, i=1,…,N as the corresponding best optimal FMO values;
- Choose $\alpha_i > 0$, i=1,…,N, the initial step size and α_{min} the minimum step size;

Iteration (for each of the active searches)

1. *Search step.* Use a minimum Frobenius norm quadratic model considering the beam angle sets already tested to search the entire BAO space. If the objective function value is improved, repeat the search step, otherwise proceed to the poll step.
2. *Poll* step. Compute f(x), $\forall\, x \in N(x_i^*) = \{x_i^* \pm \alpha_i e_j, j=1,\ldots,n\}$, where e_j is the j column of identity matrix $I=[e_1 \ldots e_n]$;
3. *If* poll is successful, i.e. $\min_{N(x_i^*)} f(x) < f(x_i^*)$ then

   $x_i^* := \underset{N(x_i^*)}{\operatorname{argmin}} f(x)$;

   $f_i^* := f(x_i^*)$;

   Else

   $\alpha_i := \frac{\alpha_i}{2}$

4. If $\alpha_i \geq \alpha_{min}$ return to step 1 for a new iteration, otherwise search started by initial point x_i becomes inactive;

Output

– $f^* := \text{argmin } f_i^*$ is the best FMO found and x^* is the corresponding beam direction ensemble.

4.2.4. Study design

IMRT plans were automatically generated in iCycle for all NPC cases. Based on the defined wish-list, plan optimization was performed using 7 coplanar equidistant beams (*d7*) corresponding to the standard clinical option. In a second phase, plan optimization was done by applying beam angular optimization. Breedveld *et al.*[23] (algorithm *i*) and Rocha *et al.*[25] (algorithm B) beam angular optimization algorithms were used to generate IMRT plans with 5, 7 and 9 beams (*i5*, *i7*, *i9* and *B5*, *B7*, *B9*, respectively). For both algorithms, coplanar and non-coplanar beam geometries were considered. For the non-coplanar case it is important to guarantee that there are no collisions between the gantry and the treatment couch (Figure S4.3 in the Supplementary Material). For algorithm *B*, this is usually achieved by strongly penalizing these solutions (assigning a very large value for the FMO). This means that these solutions can be found in the algorithms' search procedures but will be discarded since better solutions will be found. For algorithm *i*, a space search is defined before the beginning of the optimization that should avoid collisions. However, due to the discretization of the space search, it is not possible to completely prevent beam incidences located in regions defined as avoidable.

4.2.5. Plan assessment and comparison

Plan assessment and comparison were performed using an independent graphical method developed by Ventura *et al.*[28] SPIDERplan, is based on a scoring approach that considers both target coverage and individual OAR sparing. In SPIDERplan framework, targets and OARs are divided into groups depending on their clinical priorities. A score is determined for each structure based on pre-defined planning objectives and relative weights. A global plan score is determined as a weighted sum of the structures' individual scores over all groups. This score is just a quality plan indicator without any direct correlation with the treatment outcome. All dosimetric plan information is graphically represented in customized radar plots. Evaluation of plan quality can be done globally by displaying all structures (Structures Plan Diagram - SPD) and groups of structures (Group Plan Diagram - GPD). Global plan score is determined as a weighted sum of the structures individual scores over all groups as:

$$\text{Global plan score} = \sum_i w_{group(i)} \sum_j w_{struct(j)} \text{Score}_{struct(j)} \tag{4.1}$$

where $w_{struct(j)}$ and $\text{Score}_{struct(j)}$ are the relative weight and the score of structure j, respectively, and $w_{group(i)}$ the relative weight of group i. For the PTVs, the score was calculated according to the following expression:

$$\text{Score}_{PTV} = \frac{D_{TC,PTV}}{D_{P,PTV}} \tag{4.2}$$

where $D_{TC,PTV}$ corresponds to the tolerance criteria for the PTV (in this case the dose in 98% of the PTV that should be at least 95% of the prescribed dose, Table S4.1) and $D_{P,PTV}$ is the planned dose in the PTV. This is a target coverage criterion. For the OARs, the score was set as:

$$Score_{OAR} = \frac{D_{P,OAR}}{D_{TC,OAR}} \tag{4.3}$$

where $D_{P,OAR}$ is the OAR planned dose and $D_{TC,OAR}$ is the tolerance dose for each OAR.

A more detailed group evaluation can also be done with the partial group plots (Structures Group Diagrams – SGD), where only the structures of that group are represented. As for the SPD and GPD, a partial group score complementing the graphical assessment is determined for each SGD.

For this study, all delineated structures were grouped according to their location and clinical importance into: PTV group (PTVs), Critical group (spinal cord and brainstem), Optics group (chiasm, optical nerves, retinas and lens), DigestOral group (parotids, oral cavity, oesophagus and larynx), Bone group (temporal mandibular joint, mandible and ear canals) and Other group (brain, pituitary gland, thyroid and lungs). To each group a relative weight of 50%, 30%, 10%, 5%, 3.5% and 1.5%, respectively, was pre-assigned by the radiation oncologist. Within each group, the same weight was attributed each structure of that group. For detailed information, the reader is referred to Table S4.1 in the Supplementary material.

The score of each structure is determined considering the ratio between the clinical tolerance criteria and the planned dose. Thus, a value of one is expected if the dose for that structure is equal to the respective tolerance value. When a better organ sparing or target coverage is obtained, a score less than one will be obtained. Optimal scores will converge to the centre of the radar plot.

4.2.6. Statistical Analysis

Statistical comparisons of the mean scores associated with each BAO algorithm and geometry sets were performed using IBM SPSS software, version 25. As the same set of patients is used to perform IMRT optimization applying the two BAO algorithms and different geometric settings, it was assumed that the samples were dependent. As the number of patients selected for this retrospective study is greater than 30, it was also considered that the samples follow a normal distribution. Statistically significant differences between the families of test were assessed with a randomized block design ANOVA test and, if applicable, a post-hoc multiple comparison test using the Tukey method. Single pair comparisons were statistically evaluated with the t-test. A level of significance of 5% was considered for all statistical tests.

4.2.7. Methodology used for the presentation of results

The results from the two BAO algorithms were structured in two subsections. In the first (section 4.3.1), a beam angle distribution analysis was performed using circular diagrams for the coplanar geometries and 2D-maps for the non-coplanar situation. The mean angle incidences (calculated by sorting the beam angles calculated for each patient) and the associated standard deviation angles of each algorithm and the angles from the equidistant beam angle solution (*d7*) were also represented. For the coplanar geometries, the circular diagrams were composed by three concentric rings with an angle section resolution of 10º that were used to represent the relative frequencies of the beam angle distribution obtained for the two algorithms. The inner ring of the circular diagrams showed the beam angle distribution of the BAO with 5 beams, the middle ring corresponds to the optimization with 7 beams and in the external ring the 9 beams results were shown. Each ring was divided into eight regions, described in Table 4.1, commonly used in the clinical

Table 4.1 - Gantry and couch regions defined for the beam angular distribution analysis. The gantry regions were applied for the coplanar and non-coplanar analysis (circular diagrams and 2D-maps, respectively) and the couch regions to the non-coplanar analysis.

Gantry		Couch	
Region label	**Angle region**	**Region label**	**Angle region**
Anterior	[340º, 20º[	Left coronal (LCOR)	[-90º, -70º[
Left oblique anterior (LOA)	[20º, 70º[	Oblique left non-coplanar (OLNC)	[-70º, -20º[
Left lateral (LL)	[70º, 110º[	Central non-coplanar (CNC)	[-20º, 20º[
Left oblique posterior (LOP)	[110º, 160º[	Oblique right non-coplanar (ORNC)	[20, 70º[
Posterior	[160º, 200º[	Right coronal (RCOR)	[70, 90º]
Right oblique posterior (ROP)	[200º, 250º[		
Right lateral (RL)	[250º, 290º[		
Right oblique anterior (ROA)	[290º, 340º[		

routine to label beam/patient orientations. For the non-coplanar plans, 2D-maps were used to perform the beam angle distribution analysis. A grid resolution of 10º was considered. The gantry angles axis (vertical) was divided in the same groups defined as for the coplanar case. The couch angles axis (horizontal) was grouped into five regions also included in Table 4.1. The dosimetric performance of BAO optimizations is presented in section 4.3.2 using SPIDERplan analysis. A mean global analysis of the plans dose distribution quality was performed using the global plan score and the partial group scores described above. Furthermore, an individual analysis for some patients with relevant global results was performed by presenting the GPD and SGDs of interest.

4.3. Results

4.3.1. Beam angle distribution

The frequency analysis of the beam angle distribution of algorithms *i* and *B* for 5, 7 and 9 beams is shown in Figure 4.1 (coplanar setting) and Figure 4.2 (non-coplanar setting).

In Figure 4.1, the beam angles for the equidistant beam angle solution (red dash lines) and the mean angle incidences (black solid pointers) with the associated standard deviation (grey solid arcs are also represented in each circular beam diagram. The two coplanar BAO approaches presented distinct beam angular distribution patterns. In algorithm *i* (*i5*, *i7* and *i9*), based on an iterative BAO framework, the beam angular distribution across the regions was asymmetric with preferred regions very well defined as the LOP, the anterior or the ROP regions. The relative frequency values were comprehended between 4% and 22%. The higher relative frequency values, corresponding to the preferred irradiation directions, were the LOP region and the anterior region in *i5* set and also the ROP region in *i7* and *i9* sets. For algorithm *B* (*B5*, *B7* and *B9*), a more evenly angular distribution was obtained, with relative frequency values ranging between 8% and 18%. For all sets, the anterior, the posterior and the LOP regions presented the higher relative frequency values. For both algorithms, anterior-oblique and lateral orientations were not often selected as good irradiation directions regardless the number of beams used. The difference between the optimal mean angle and the correspondent equidistant beam was 0.8º±34º for algorithm *i* and 1.5º±19.9º for algorithm *B*.

In Figure 4.2, for the non-coplanar BAO modality, the gantry angle distribution and the couch angle distributions are presented by a relative frequency 2D-map. The gantry angles axis and the couch angles axis were divided into regions of interest, referred in Table 4.1. For simplification each region, composed of a set of gantry and couch angles, will first be named with the gantry region

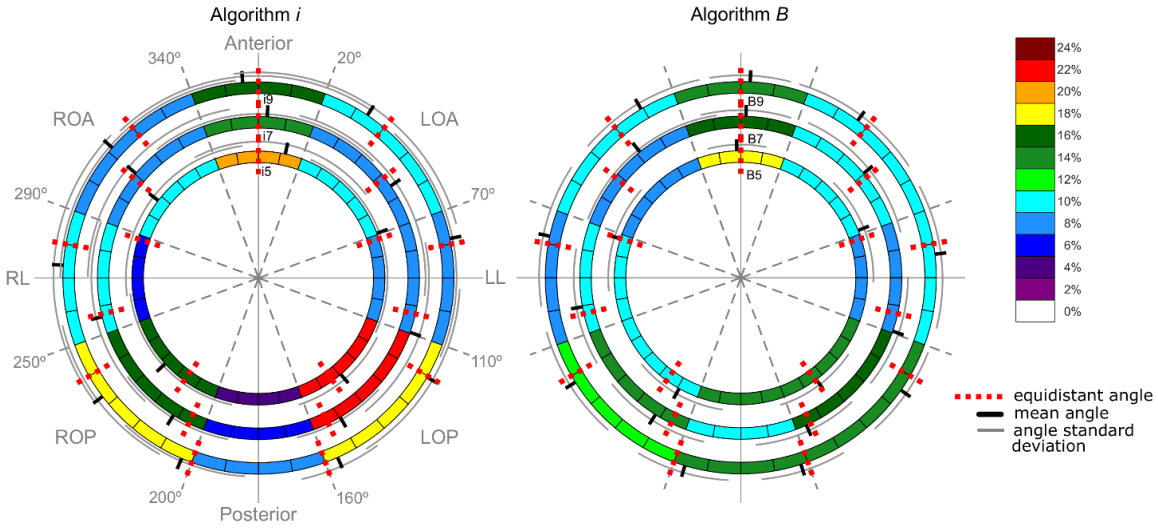

Figure 4.1 - Angular representation of the relative frequencies of the coplanar BAO of algorithms *i* and *B* for 5, 7 and 9 beams. The colour represents the relative frequencies obtained for each angle section: a hot colour is associated to a high relative frequency and a cold colour to a low relative frequency value. The mean angle incidences and the associated standard deviation angle values of each algorithm were represented with black solid pointers and grey solid arcs respectively. The red dash lines represent the beam angles of the equidistant beam angle solution (*d7*).

followed by the couch region (for instance, posterior_CNC). In Figure 4.2, the beams position for the equidistant coplanar beam plans are shown by the red dots, the individual beam incidences obtained by angular incidences by small black dots, the correspondent mean angle incidences by the large black dots and the associated standard deviation by the grey ellipses. Avoidance incidences, corresponding to potential collisions between gantry and couch, were represented by yellow grid squares. The two non-coplanar BAO algorithms presented again distinct beam distribution patterns. In algorithm *i*, most of the beams were uniformly distributed over the space, with relative frequencies ranging between 0% (white squares in Figure 4.2) and 7% (cyan). Interestingly, the preferred irradiation directions selected by algorithm *i* where almost neglected by algorithm *B* where relative frequency values of less than 1% were obtained. In algorithm *B*, with relative frequency values ranging between 0% and 13%, it is possible to define a pattern for the beam's angular distribution. In fact, the non-coplanarity is almost confined to couch angulation between -20º and 20º, corresponding to the CNC region. The remaining regions presented relative frequency values inferior to 2%.

The average beam incidences, and especially the standard deviation values, for both algorithms are quite different. Graphically, this can be perceived in Figure 4.2 by the clear separation between the ellipses for algorithm *B* while for algorithm *i* the standard deviations ellipses overlap each other. Furthermore, the distance between the mean incident angles (large black dots) and the correspondent equidistant solution (red dots) are closer in algorithm *B* than in algorithm *i*.

4.3.2. SPIDERplan scores analysis

4.3.2.1. Global plan analysis

The values of the global plan score, implemented in SPIDERplan, for the *d7* plans (equidistant beams) and coplanar and non-coplanar BAO of algorithms *B* and *i* for 5, 7 and 9 beams are shown in Figure 4.3a. The mean SPIDERplan global plan scores ranged between 0.901 and 0.947. The lowest mean plan scores, i.e. the plans with better overall score performance, were obtained by *B9* non-coplanar (*B9nc*) and *B9* coplanar (*B9c*) sets. *i9c* plans attained the third best score while *i9nc*

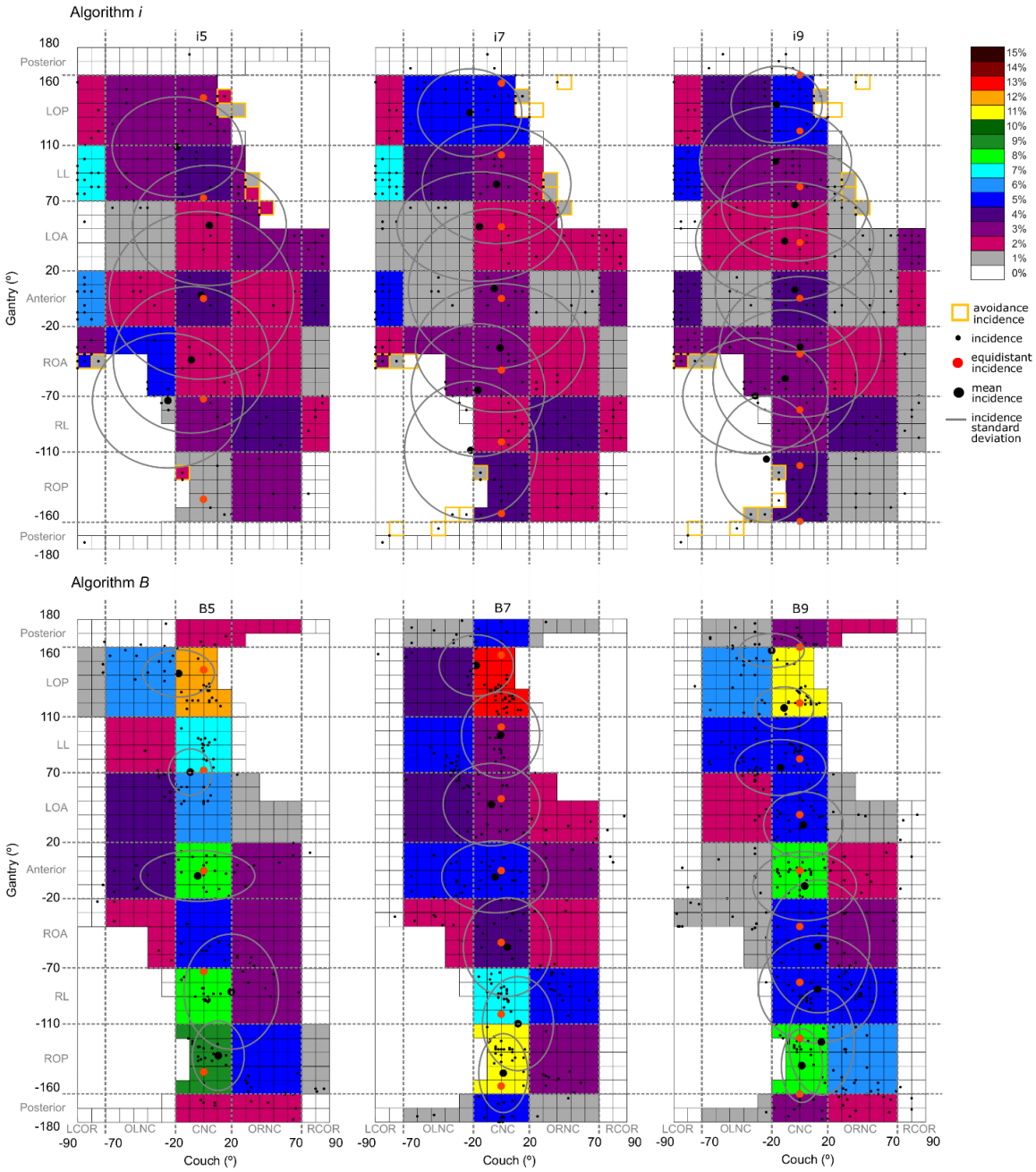

Figure 4.2 - 2D map representation of the relative frequencies of the non-coplanar BAO of algorithm *i* and *B* for 5, 7 and 9 beams. The gantry angles values are represented on the vertical axis and the couch angles on the horizontal axis. The colour represents the relative frequencies obtained for each angle section: a hot colour is associated to a high relative frequency. The mean angle incidences and the associated standard deviation angle values of each algorithm were represented with big black solid pointers and grey solid ellipses respectively. The small black solid pointers represent the angles incidence obtained with the BAO for all patients. The red solid points represent the beam angles values of the equidistant beam angle solution (*d7*).

set only achieved the eighth best score immediately below all plans using 7 beams. *B5c* and *B5nc* plans, respectively, obtained a better performance than *i7nc*. The highest mean global plan scores, and therefore the worst overall plan performances, were obtained with the *i5nc* and *i5c* sets.

The statistical analysis that was carried out allowed the identification of pairs of algorithms and beam angle configurations such that the generated treatment plans cannot be considered as being

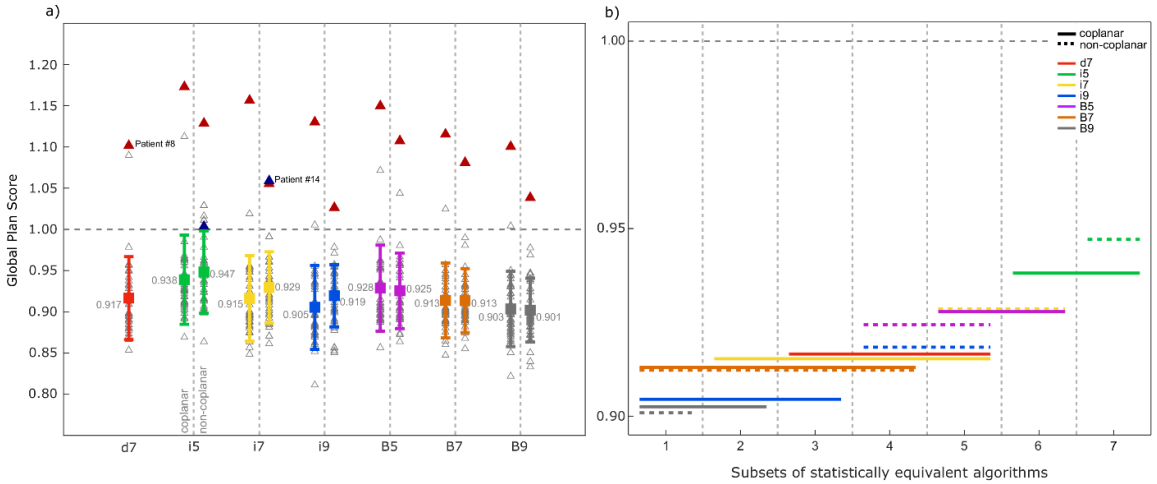

Figure 4.3 - a) SPIDERplan Global Plan scores, corresponding to all 40 clinical cases (triangles), for *d7* and for coplanar and non-coplanar BAO of algorithms i (coplanar plans: *i5c, i7c, i9c*, non-coplanar plans: *i5nc, i7nc* and *i9nc*) and *B* (*B5c, B7c, B9c, B5nc, B7nc* and *B9nc*). b) Homogenous subsets resulting from post-hoc multiple comparisons test using the Tukey method with a level of significance of 5% of the SPIDERplan global plan scores of each algorithm for coplanar and non-coplanar BAO.

different from a statistical point of view. The results of the statistical analysis and the resulting p-value of each comparison led to seven subsets, grouping the algorithms that did not present statistically significant differences. It was thus possible to build sets, as presented by the horizontal axis of Figure 4.3b, such that each set includes similar treatment planning results. As an example, subset 1 shown in Figure 4.3b, with the lowest global plan scores, includes *B9nc*, *B9c*, *i9c*, *B7nc* and *B7c* meaning that the quality of these plans is statistically equivalent. Statistically significant differences were found between plans *B9nc* (positioned in subset 1) and plans i9nc (belonging to subsets 4 and 5). These results also show that BAO may bring no benefit to plan quality when compared to the equidistant beam angle solution - all those solutions that overlap the red solid line belonging to subsets 3, 4 and 5, like *i9c*, *i7nc*, *i9nc*, *B7nc* or *B7c*, do not significantly differ from it. However, better plan scores were obtained when the number of beams increased to 9 beams, as in B9c and B9nc (subset 1), compared to 7 beams (*i7nc* or *d7*). Also, while a statistically significant difference in plan quality was found between non-coplanar and coplanar plans using 9 beams whose positions were determined by algorithm *i*, for algorithm *B*, non-coplanarity brought no improvement in terms of plan quality. It is interesting to observe that, for algorithm *i* and 9 beams, the 9 beam coplanar plans were better than the non-coplanar ones.

The mean scores of coplanar and non-coplanar sets of algorithms i and B are compared in Figure 4.4a and d, respectively. For algorithm i, the coplanar set had a lower mean score than the non-coplanar set (p=0.002) whereas for algorithm B, non-coplanar plans were statistically equivalent to coplanar ones (p=0.960). Statistically significant differences were also found between non-coplanar plans optimized by algorithm i and B (p=0.000), in favour of non-coplanar B plans, (Figure 4.4e). The overall superior performance of algorithm B over algorithm i was statistically significant (p=0.000), as demonstrated in Figure 4.4f.

4.3.2.2. Group plan analysis

The quality of the plans based on BAO algorithms *B* and *i* was assessed also using the information generated by SPIDERplan diagrams (Figure S4.4 - S4.8 in the Supplementary Material). Generally, the group score agreed with the analysis performed for the global plan score section. Almost all structure groups included in the optimization got mean scores below 1, meaning that the clinical criteria were on average accomplished. The exception was the DigestOral group, where for

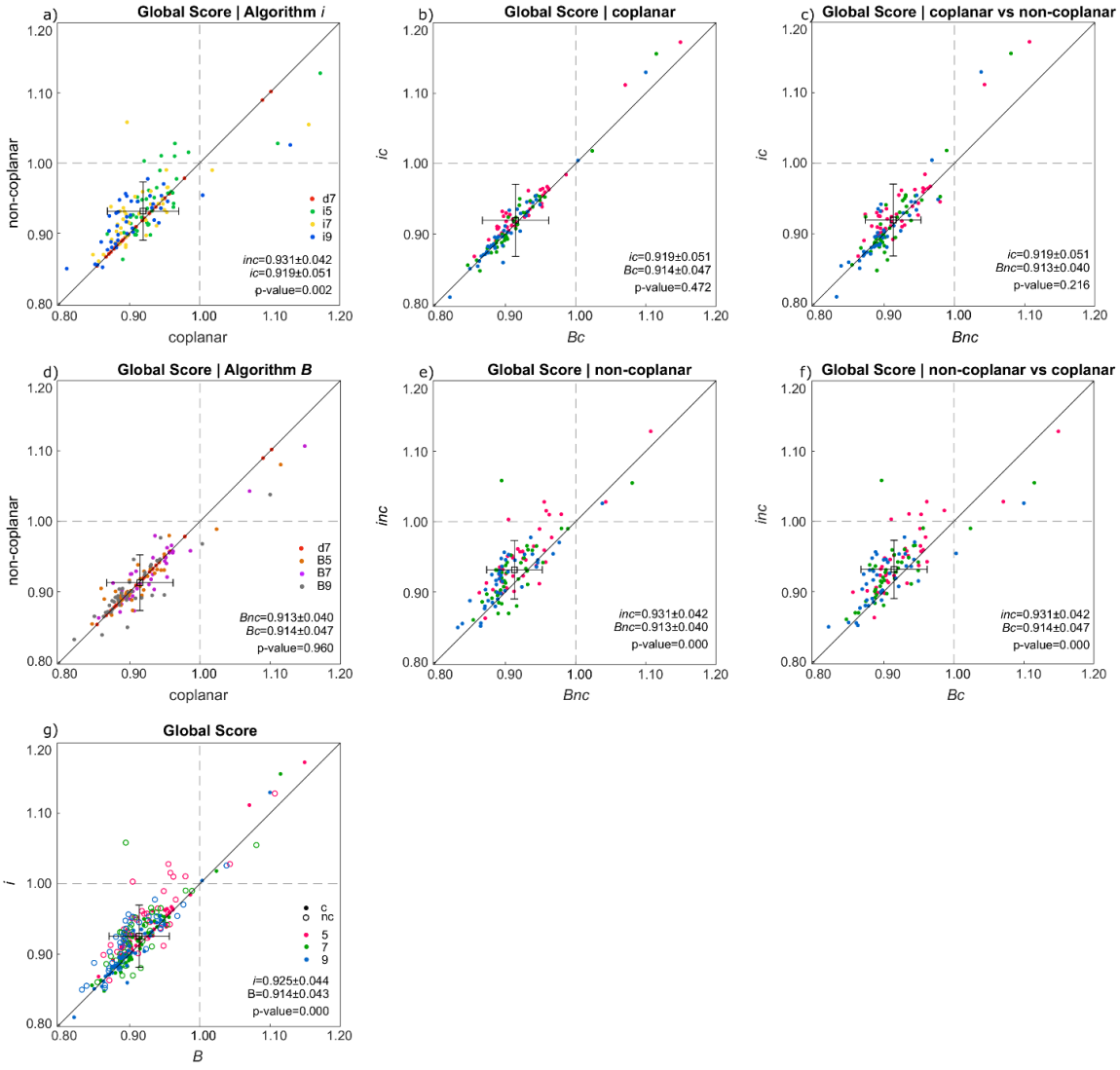

Figure 4.4 - Comparison between different plans optimized with algorithms *i* or *B* using coplanar or non-coplanar beams.

the parotids and for the oral cavity planned doses surpassed tolerance doses. Differences in the mean group scores between non-coplanar plans of algorithm *i* and the remaining plans and between 9 beams and 5 beams plans of algorithms *i* and *B* were also obtained for the Optics and the DigestOral groups, respectively.

The plans with higher number of beams and a non-coplanar geometry tend to lead to dose distributions with better quality, i.e. higher PTV coverage and higher OAR sparing. Some exceptions were found for the Optics and Bone groups. In the Optics group, the best scores were found for the coplanar beam geometries of algorithm *B* and the worst for the non-coplanar sets of algorithm *i*. For the Bone group, either the non-coplanar or the coplanar plans of algorithm *B* achieved the best performances, while the coplanar and the non-coplanar sets of algorithm *i* got the worst scores. Globally, algorithm *i* presented better scores for the two most important groups (PTV and Critical group), while algorithm *B* got the best scores for the remaining groups. However, the differences in plan quality for each structure group between the two algorithms were statistically significant just for the Optics and Bone groups, which included the OARs with the lowest clinical weight.

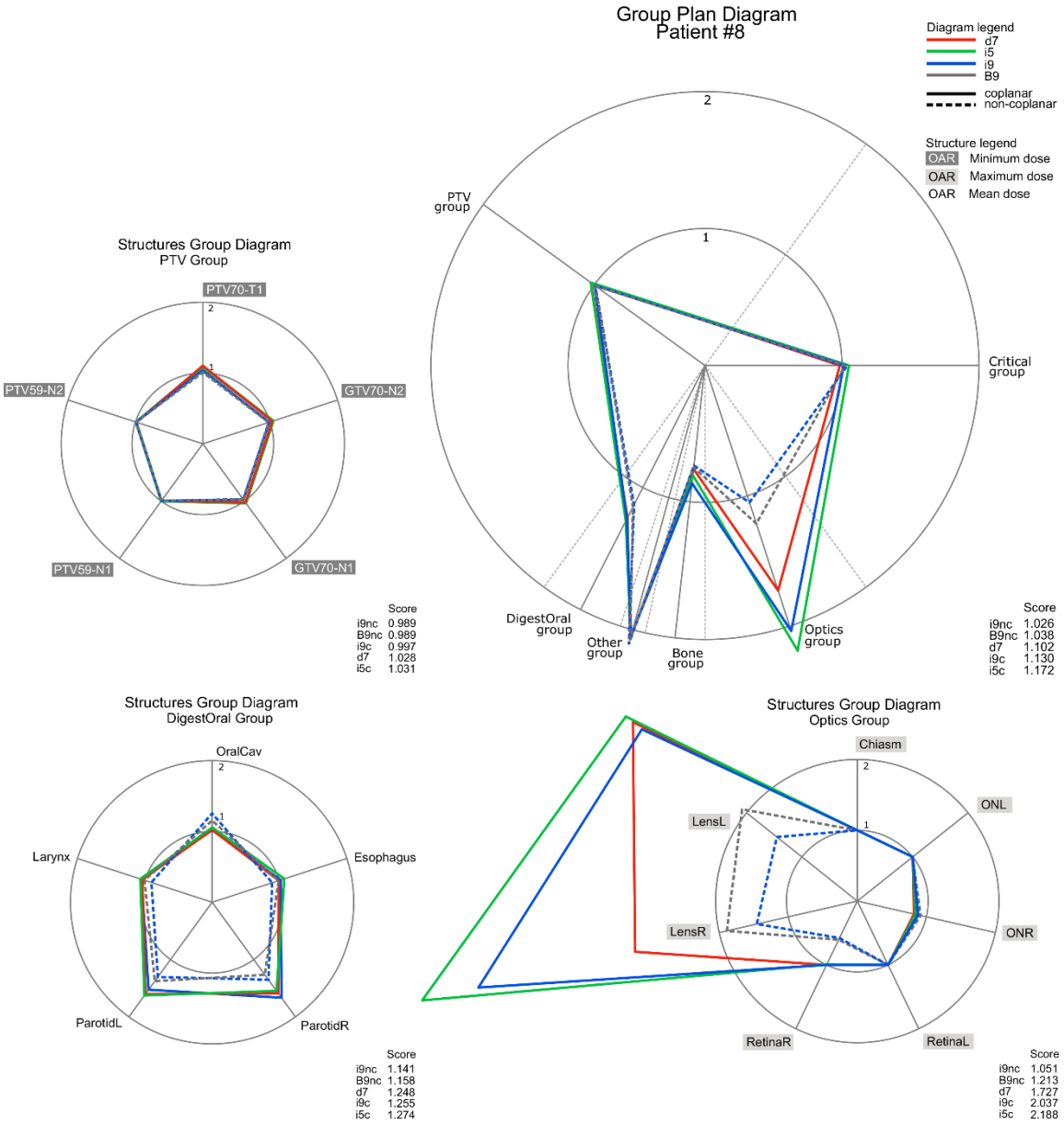

Figure 4.5 - SPIDERplan of patient number 8 and structures group diagram for PTV group, Optics group and DigestOral group.

4.3.2.3. Individual patient analysis

The decision of which beam set-up should be used in a given patient must be well pondered and clinically assessed case by case. In Figure 4.3a, two patients (patient #8 and #14) were identified with notorious high scores (worst plan quality). For patient #8 (red triangles in Figure 4.3a), all plans obtained a global score superior to 1 and presented mean percent differences between the coplanar and the non-coplanar sets for algorithms *i* and *B* of -8% and -5%, respectively. For patient #14 (blue triangles in Figure 4.3a), two plans exceeding the score threshold defined for SPIDERplan, presented an apparent contradictory score difference, wherein plan *i5nc* was better (lower score) than plan *i7nc*. The assessment of plan quality for patients #8 and #14 is presented in Figure 4.5 and 4.6, using the GPD and the SGDs of SPIDERplan. Plans using equidistant beam angles (d7) and the plans with

Figure 4.6 - SPIDERplan of patient number 14 and structures group diagram for Optics group, DigestOral group and Other group.

the best and worst global plan scores were selected for this individual analysis. One or more additional sets were also considered to emphasize some results of interest observed in each patient.

For patient #8 (corresponding to Figure S4.1 in the Supplementary Material) the best global plan score was achieved by plan *i9nc* (global plan score of 1.026) and the worst global plan score, of 1.172, was obtained with plan *i5c*. An increase of 15% percent in plan quality of *i9nc* plan, when compared with *i5c*, was achieved when SPIDERplan global score is adopted as plan quality scoring metric. A percent difference of +11% was obtained between the global plan scores of *i9nc* and *i9c*. These differences highlight the potential benefits that can arise from angular optimization including non-coplanar beam angle incidences. The largest difference between the tolerance and the planned dose was obtained for the Optics group and the DigestOral group (Figure 4.5). For the PTV group and in the Critical group some score values slightly higher than 1 were also obtained for some plans due to the proximity between the primary tumour mass, prescribed to 70 Gy, and the retinas, the optical nerves, the chiasm, the brainstem, the ears and the oral cavity. The increase in the number

of beams with non-coplanar geometries led to important improvements in the OAR sparing, especially in the lens and the parotids but also in PTV coverage. Nevertheless, these improvements were not extensible to all structures where even worst results were obtained for the oral cavity when 9 non-coplanar beams were used.

For patient #14 (corresponding to Figure S4.2 in the Supplementary Material), *B9c* presented the best global plan score and *i7nc* the worst performance. A mean percent difference in the global score of -10% was achieved when coplanar and non-coplanar sets of algorithm i were compared (Figure 4.6). For algorithm *B*, this mean percent difference was close to 0%, meaning that for this patient the non-coplanarity did not bring any advantage for algorithm *B*. Significant differences between the considered plans can be identified for the lens (Optics group), the left ear (Bone group) and for the left parotid (Digest Oral group). All structures but the right lens presented better scores for plans with higher number of beams and/or non-coplanar geometry. For the right lens, however, *i5nc* presented a better score than *i7nc*. This configures a situation where a larger number of beams did not bring improvements to the overall plan quality. Analysing the specific anatomy of patient #14, it is possible to observe that the primary mass PTV was well below the optical structures (chiasm, optical nerves, retinas and lens). This influenced the non-coplanar BAO process and probably the SPIDERplan analysis results, since some of the considered clinical criteria could probably have been relaxed.

4.4. Discussion

In this work, the plans produced by two BAO algorithms, *i* and *B,* were evaluated and compared for NPC tumour cases. Forty clinical cases were retrospectively used to automatically determine the best incidences of 5, 7 and 9 beams plan sets with coplanar and non-coplanar geometries. The BAO and the FMO problems were addressed together by using a multicriterial IMRT optimization framework to guide the process. Algorithm *i* is based on a combinatorial iterative discrete search approach and is embedded in the multicriterial optimization framework. Algorithm *B* is based on a continuous space search using a pattern search method. It is also possible to consider the optimization of the number of beams. This can be done in a trivial way, by running different optimization procedures, each one for a different number of angles. The choice of the number of angles could also be incorporated in the optimization algorithm but given the complexity of the BAO the inclusion of one more degree of freedom could actually lead to worse results (since the size of the possible solutions space would be enlarged). In the final optimization phase 240 plans with 27 associated structures were generated for each algorithm. Starting from the equidistant solution, BAO plans were considered, covering an expressive universe of 3640 beam incidences, 520 plans and 14040 dosimetric structures statistics available to be analysed. The analysis of this large amount of data was done from two perspectives: the characterization of the beam angle distribution over the space search and the assessment of the quality of the dose distribution of the generated plans. To our knowledge, this is the first work that compares these two types of class methods for head and neck cancer taking into consideration all the clinical structures using subjacent clinical criteria. Furthermore, the graphical options *ad-hoc* constructed for this purpose, the circular diagrams for the coplanar case and the 2D-map for the non-coplanar one, enable an efficient global analysis that otherwise would be difficult to be performed.

The relative frequency patterns of the beam angle distribution for coplanar and non-coplanar beams geometries seemed to be conditioned by the optimization strategy followed by each algorithm. In algorithm *i,* beams with optimal orientation were iteratively added into the plan after being combined with the beams already selected in a discretized space search. For plans using coplanar beams, this cumulative beam adding methodology generated a non-uniform angle distribution pattern where it is possible to clearly identify favourite irradiation directions and regions of low preference. For non-coplanar beam plans this asymmetric beam distribution pattern with well-defined preferred

incidences blurred into an almost uniform beam distribution pattern. This pattern change is a natural consequence of the selection of beam incidences over almost all the available space search. In algorithm *B*, the search for the best ensemble is initially done by considering a fixed number of incidences defined from the best equidistant coplanar angle set solution. This preliminary optimization is followed by the application of the pattern search method considering a continuous space search. Although the equidistant beam ensemble seems to be the most reasonable BAO starting point for this approach, the beam angle distribution maps presented patterns that may be strongly influenced by the initial solution. For coplanar geometries an almost uniform pattern, with low relative frequency values was patent in the circular diagram of frequencies. For the non-coplanar situation, the results follow the starting point option, being the non-coplanarity confined to modest deviations from zero couch position (±20º). Comparing the mean incidences and the associated standard deviations obtained by the two optimization algorithms, once again the optimization strategy of each of the algorithms is patent, leading algorithm *i* to more distributed incidence solutions and algorithm *B* proposing solutions closer to the initial equidistant case.

The quality assessment and comparison of the plans generated with BAO was performed using three types of approaches: a global plan analysis, a group plan analysis and an individual analysis of selected patients. This methodology was accomplished by the determination of SPIDERplan scores and an appropriate statistical analysis that conferred to the process the possibility to evaluate the dosimetric quality of the BAO with different levels of specificity. Increasing the number of beams brought improvements to the plan dose distribution. Nevertheless, for most cases only the comparison between 9 beam plans and 5 beam plans was significant statistical.

Algorithm *B* showed a more consistent behaviour and presented, by a moderate difference in score, a better performance than algorithm *i.* For the studied NPC tumour cases, on average, non-coplanarity brought no improvements to plan quality. For algorithm *B*, a better score was obtained when non-coplanar beams were compared with the corresponding coplanar solution but this difference was not statistically significant. These results confirmed some empirical impressions shared by many planners. In face of highly complex planning cases, beyond the manual tuning of the objectives and the associated weights, planners usually try to play with the initial beam angle incidences or to increase the number of beams in order to improve the plans. The general assumption that plan quality improves when the number of beams increases is also supported by the results usually achieved with VMAT ("infinite" number of beam directions). Nevertheless, for the studied pathology, BAO seemed to bring only marginal improvements to the plan quality. A first explanation may be related with the anatomy of the NPC cases, where the PTVs with large extensions (up to 25 cm of height), the high number of critical structures along the field of irradiation and minimum exposure requirement for the remaining normal tissues may limit the optimization of the beam incidences. Another justification can be found in the use of the same wish-list for all patients and for both BAO approaches. The improvement that can be obtained by BAO is intrinsically linked to the FMO approach. Since the resources and time needed to find an optimal beam set are costly, if manual tuning is needed then the clinical utility of BAO must be seriously appraised. If BAO can be done in an automated way, then it will represent an added-value, since it can bring interesting improvements for some patients. In Erasmus-iCycle, the treatment planning procedure is almost automatic. The only think that is asked to the planner before the planning is to build and validate a wish-list that will guide the multicriterial optimization process. For the NPC pathology, five test cases were used in the validation process. This initial configuration does not take long, and it has a reduced impact on the overall time spent with the optimization. More expressive score differences between the treatment planning sets could be achieved if the SPIDERplan score could be embedded in the BAO process as suggested by Rocha *et al.* [26] As SPIDERplan methodology incorporates the radiation oncologist preferences, it could confer to the BAO process some proximity to the clinical aims and thus improve overall plan quality.

The overall weaker performance of algorithm *i*, when compared with algorithm *B,* is related with the results of the non-coplanar optimization in the Optics and in the Bone groups, since for the remaining groups these sets presented the best SPIDERplan scores. For the Bone group, although the non-coplanar optimization of algorithm *i* presented a better performance than the coplanar set of algorithm *i*, it was inferior to the coplanar optimization of algorithm *B*. For the Optics group, the results of the non-coplanar sets of algorithm *i* were by far the worst when compared with the remaining sets. Due to the anatomic localisation of the structures of these two groups and also to the optimization methodology subjacent to algorithm *i*, it was not expected that the non-coplanar optimization presented such results that were on average below the score tolerance but were worse than the remaining sets.

The weaker performance of non-coplanar solutions for algorithm *i* compared to coplanar plans is unexpected as the coplanar problem is a sub-solution of the non-coplanar problem. This might be a result of the complexity of BAO, a highly nonconvex problem with many local minima, particularly for complex tumour sites as NPC with a large number of OARs. Obtaining the optimal solution is quite difficult, particularly for the non-coplanar case that explores a vaster search space. As it is not possible to guarantee that the optimal solution is found, and the algorithms do not perform an exhaustive search (which would be prohibitive both in terms of time and computational resources), it is possible that the best coplanar solution is not found when looking for a non-coplanar solution. A simple strategy to improve the performance of algorithm *B* would be to include the optimal solution of the coplanar BAO in the set of initial starting solutions. However, that strategy cannot be used for algorithm *i*, since it fixes one direction at each iteration. Whenever one direction is fixed, the search space is restricted in one dimension, meaning that there are solutions in the search space that cannot be visited in the subsequent iterations. The algorithm may be prevented from exploring better regions and the probability of getting trapped in local minima increases.

An important application of non-coplanar BAO is its importance in the calculation of non-coplanar intensity-modulated arc trajectories in VMAT. In fact, some of the arc trajectory algorithms are two-step approaches where, in the first step, non-coplanar BAO is performed using previously tested BAO algorithms and, in a second step, an arc trajectory optimization is performed using the beam directions found in the first step as anchor points.[10] The fact that algorithm *B* obtain solutions for a more limited range of couch angles, with a superior quality to the solutions of algorithm *i*, might represent a competitive advantage for its use in the calculation of non-coplanar trajectories in VMAT planning.

4.5. Conclusions

In this work the beam angle optimization IMRT was addressed using forty head-and-neck cancer clinical cases. Two algorithms, based on a combinatorial iterative (algorithm *i*) and on a continuous exploration of the space search (algorithm *B*) approaches, were assessed and compared for coplanar and non-coplanar beam geometries. A graphical method for plan quality assessment and comparison, named SPIDERplan, was used. The two algorithms were assessed through the analysis of the beam angle distribution and of the plan quality. The great amount of generated data was managed through graphical plots that enabled efficient global analysis and comparisons. Algorithm *i* for coplanar optimization presented a less uniform angle distribution pattern whereas for non-coplanar optimization the beam distribution pattern was almost uniform. For algorithm *B*, both beam angles geometries options were strongly influenced by the starting equidistant solution. Concerning assessment and comparison of plan quality for BAO algorithms, slightly better score performance was achieved by algorithm *B,* when compared to algorithm *i*. For algorithm *B*, coplanar and non-coplanar beam angle geometries were statistically equivalent, while for algorithm *i*, non-coplanar solutions were statistically worse than the correspondent coplanar due to the optimization

strategy followed by this algorithm. Nevertheless, for specific patients strong benefits were obtained, and angle optimization proved to be valuable.

The results of the present study can potentially be applied in VMAT planning through the calculation of non-coplanar modulated arc trajectories.

4.6. Acknowledgments

The authors would like to express their gratitude to Sebastiaan Breedveld and Ben Heijmen for making available the multicriterial optimization of IMRT plans system Erasmus-iCycle and their valuable support and coaching along the work. The authors would like also to thank to Andreia Hall for the technical support in the statistical analysis.

This work was supported by project grant POCI-01-0145-FEDER-028030 and by the Fundação para a Ciência e a Tecnologia (FCT) under project grant UID/Multi/00308/2019.

No potential conflict of interest nor any financial disclosures must be declared.

4.7. References

1. ICRU. International Commission on radiation units and measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Geneva; 2010. *J ICRU* 10(1):1–106.
2. Zhang Y, Merritt M. Dose–volume-based IMRT fluence optimization: a fast least-squares approach with differentiability. *Linear Algebr Appl.* 2008;428(5,6):1365–1387.
3. Spirou SV, Chui CS. A gradient inverse planning algorithm with dose–volume constraints. *Med Phys*. 1998;25(3):321–333.
4. Webb S. Optimization of conformal radiotherapy dose distributions by simulated annealing. *Phys Med Biol*. 1989;34(10):1349-1370.
5. Thieke C, Küufer KH, Monz M, Scherrer A, Alonso F, Oelfke U et al. A new concept for interactive radiotherapy planning with multicriteria optimization: First clinical evaluation. *Radiother and Oncol*. 2007;85(2):292-298.
6. Miettinen K. *Nonlinear multiobjective optimization*. New York: Springer US; 1998.
7. Branke J, Deb K, Miettinen K, Slowiński R. *Multiobjective Optimization. Interactive and Evolutionary Approaches*. New York: Springer-Verlag Berlin Heidelberg; 2008.
8. Das SK, Marks LB. Selection of coplanar or noncoplanar beams using three-dimensional optimization based on maximum beam separation and minimized nontarget irradiation. *Int J Radiat Oncol Biol Phys*. 1997;38(3):643-655.
9. Rocha H, Dias J, Ventura T, Ferreira B, Lopes MC. A derivative-free multistart framework for an automated noncoplanar beam angle optimization in IMRT. *Med Phys*. 2016;43(10):5514–5526.
10. Papp D, Bortfeld T, Unkelbach J. A modular approach to intensity-modulated arc therapy optimization with noncoplanar trajectories. *Phys Med Biol*. 2015;60(13):5179–5198.
11. Bortfeld T, Schlegel W. Optimization of beam orientations in radiation therapy: some theoretical considerations. *Phys Med Biol*. 1993;38(2):291-304.
12. Ehrgott M, Holder A, Reese J. Beam selection in radiotherapy design. *Linear Algebra Appl*. 2008;428(5):1272-1312.
13. Craft D. Local beam angle optimization with linear programming and gradient search. *Phys Med Biol.* 2007;52(7):127-135.
14. Stein J, Mohan R, Wang XH et al. Number and orientation of beams in intensity-modulated radiation treatments. *Med Phys*. 1997;24(2):149-160.
15. Pugachev A, Lei X. Pseudo beam's-eye-view as applied to beam orientation selection in intensity-modulated radiation therapy. *Int J Radiat Oncol Biol Phys.* 2001;51(5):1361-1370.
16. Dias J, Rocha H, Ferreira BC, Lopes MC. Simulated annealing applied to IMRT beam angle optimization: A computational study. *Phys Medica.* 2015;31(7):747-756.
17. Dias J, Rocha H, Ferreira BC, Lopes MC. A genetic algorithm with neural network fitness function evaluation for IMRT beam angle optimization. *Cent Eur J Oper Res*. 2014;22(3):431–455.

18. Aleman DM, Kumar A, Ahuja RK, Romeijn HE, Dempsey JF. Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning. *J Global Optim*. 2008;42(4):587–607.
19. Lim GJ, Cao W. A two-phase method for selecting IMRT treatment beam angles: Branch-and-Prune and local neighborhood search. *Eur J Oper Res*. 2012;217(3):609-618.
20. Bertsimas D, Cacchiani V, Craft D, Nohadani O. A hybrid approach to beam angle optimization in intensity-modulated radiation therapy. *Comput Oper Res*. 2013;40(9):2187-2197.
21. Sultan ASA. *Optimization of beam orientations in intensity modulated radiation therapy planning PhD Thesis*. Department of Mathematics, Technical University of Kaiserslautern, Germany; 2006.
22. Lim GJ, Choi J, Mohan R. Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning. *OR Spectrum*. 2008;30(2):289-309.
23. Breedveld S, Storchi P, Voet P, Heijmen B. iCycle: integrated, multi-criterial beam angle and profile optimization for generation of coplanar and non-coplanar IMRT plans. *Med Phys*. 2012;39(2):951-963.
24. Bangert M, Unkelbach J. Accelerated iterative beam angle selection in IMRT. *Med Phys*. 2016;43(3):1073-1082.
25. Rocha H, Dias J, Ferreira BC, Lopes MC. Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method. *Phys Med Biol*. 2013;58(9):2939-2953.
26. Rocha H, Dias JM, Ventura T, Ferreira B, Lopes M. Beam angle optimization in IMRT: are we really optimizing what matters? *Intl. Trans in Op Res*. 2019;26(3):908-928.
27. Breedveld S, Storchi P, Keijzer M, Heemink A, Heijmen B. A novel approach to multi-criteria inverse planning for IMRT. *Phys Med Biol*. 2007;52(20):6339-6353.
28. Ventura T, Lopes MC, Ferreira BC, Khouri L. SPIDERplan: A tool to support decision-making in radiation therapy treatment plan assessment. *Rep Pract Oncol Radiother*. 2016;21(6):508-516.
29. Breedveld S, Storchi P, Heijmen B. The equivalence of multi-criteria methods for radiotherapy plan optimization. *Phys Med Biol*. 2009;54(23):7199-7209.
30. Custódio AL, Rocha H, Vicente LN. Incorporating minimum Frobenius norm models in direct search. *Comput Optim Appl*. 2010;46:265-278.

4.8. Supplementary material

Table S4.1 - Groups, structures and weights considered for SPIDERplan processing (D_p – Prescribed dose, D_{max} – maximum dose, D_{mean} – mean dose).

Groups		**Structures**		
Name	*Group Weight*	*Name*	*Structure Weight*	*Tolerance criteria*
PTV	50%	PTVs	100%	$D_{98\%} \geq D_{p,95\%}$
Critical	30%	Spinal cord (SPNLCORD)	50%	$D_{max} \leq 45Gy$
		Brainstem		$D_{max} \leq 54Gy$
Optics	10%	Chiasm	14.3%	$D_{max} \leq 55Gy$
		Left optical nerve (OPTNRVL)		$D_{max} \leq 55Gy$
		Right optical nerve (OPTNRVR)		$D_{max} \leq 55Gy$
		Left retina (RETINAL)		$D_{max} \leq 45Gy$
		Right retina (RETINAR)		$D_{max} \leq 45Gy$
		Left lens (LENSL)		$D_{max} \leq 6Gy$
		Right lens (LENSR)		$D_{max} \leq 6Gy$
DigestOral	5%	Left parotid (PAROTIDL)	20%	$D_{mean} \leq 26Gy$
		Right parotid (PAROTIDR)		$D_{mean} \leq 26Gy$
		Oral cavity (ORALCAV)		$D_{mean} \leq 35Gy$
		Oesophagus		$D_{mean} \leq 40Gy$
		Larynx		$D_{mean} \leq 45Gy$
Bone	3.5%	Left temporal mandibular junction (TMJL)	20%	$D_{max} \leq 66Gy$
		Right temporal mandibular junction (TMJR)		$D_{max} \leq 66Gy$
		Mandible		$D_{max} \leq 66Gy$
		Left ear canal (EARL)		$D_{mean} \leq 45Gy$
		Right ear canal (EARR)		$D_{mean} \leq 45Gy$
Other	1.5%	Brain	20%	$D_{max} \leq 54Gy$
		Pituitary Gland (PITUITARY)		$D_{max} \leq 60Gy$
		Thyroid		$D_{mean} \leq 27.5Gy$
		Left lung (LUNGL)		$D_{mean} \leq 5Gy$
		Right lung (LUNGR)		$D_{mean} \leq 5Gy$

Table S4.2 - General wish-list defined for NPC cases.

Level	Priority	Structure	Type	Goal	Sufficient	Parameters
Constraints						
		PTV$x_{1,2}$-N1+N2*	maximum	$D_{p,107\%}$		
		PTV70-T1	maximum	$D_{p,107\%}$		
		PTV$x_{1,2}$-N1+N2 shell†	maximum	$D_{p,107\%}$		
		SPINAL CORD	maximum	45 Gy		
		BRAINSTEM	maximum	54 Gy		
		OPTICS+	maximum	55 Gy		
		RETINAS	maximum	45 Gy		
		RING PTV$x_{1,2}$-N1+N2‡	maximum	$D_{p,85\%}$		
		RING PTV70-T1ᵠ	maximum	$D_{p,85\%}$		
		EXTERNAL RING¥	maximum	45 Gy		
		Unspecified Tissue	maximum	69.96 Gy		
Objectives						
1	1	PTV$x_{1,2}$-N1+N2	LTCP	1	0.5	D_p=59.40 Gy, α=0.75
	2	PTV70-T1	LTCP	1	0.5	D_p=69.96 Gy, α=0.75
	3	PTV$x_{1,2}$-N1+N2 shell	LTCP	1	0.5	D_p=59.40 Gy, α=0.75
	4	EXTERNAL RING	maximum	$D_{max,95\%}$		
	5	SPINAL CORD	maximum	$D_{max,95\%}$		
	6	BRAINSTEM	maximum	$D_{max,95\%}$		
	7	OPTICS	maximum	$D_{max,95\%}$		
	8	RETINAS	maximum	$D_{max,95\%}$		
	9	LENS	gEUD	12 Gy		a=12
	10	EARS	mean	50 Gy		
	11	PAROTIDS	mean	50 Gy		
	12	ORAL CAVITY	mean	45 Gy		
	13	TMJ	maximum	66 Gy		
	14	MANDIBLE	maximum	66 Gy		
	15	OESOPHAGUS	mean	45 Gy		
	16	LARYNX	mean	45 Gy		
2	17	OPTICS	gEUD	48 Gy		a=12
	18	RETINAS	gEUD	22 Gy		a=12
	19	LENS	gEUD	6 Gy		a=12
	20	EARS	mean	45 Gy		
	21	PAROTIDS	mean	26 Gy		
	22	ORAL CAVITY	mean	35 Gy		
	23	OESOPHAGUS	mean	35 Gy		
	24	LARYNX	mean	35 Gy		
	25	BRAIN	gEUD	54 Gy		a=12
	26	PITUITARY GLAND	gEUD	60 Gy		a=12
	27	THYROID	mean	27.5 Gy		
	28	LUNGS	mean	5 Gy		

*PTVx1,2-N1+N2 - Union of PTVx1-N1 and PTVx2-N2, +OPTICS - Union of optical nerves and chiasm, †PTVx1,2-N1+N2 shell - Subtraction of a 10 mm margin of PTV70-T1 to PTV$x_{1,2}$-N1+N2, ‡RING PTV$x_{1,2}$-N1+N2 – Ring of 10 mm of thickness at 10 mm distance from PTV$x_{1,2}$-N1+N2, ᵠRING PTV70-T1 – Ring of 10 mm of thickness at 10 mm distance from PTV70-T1, ¥EXTERNAL RING - ring of 10 mm thickness next to patient outer contour

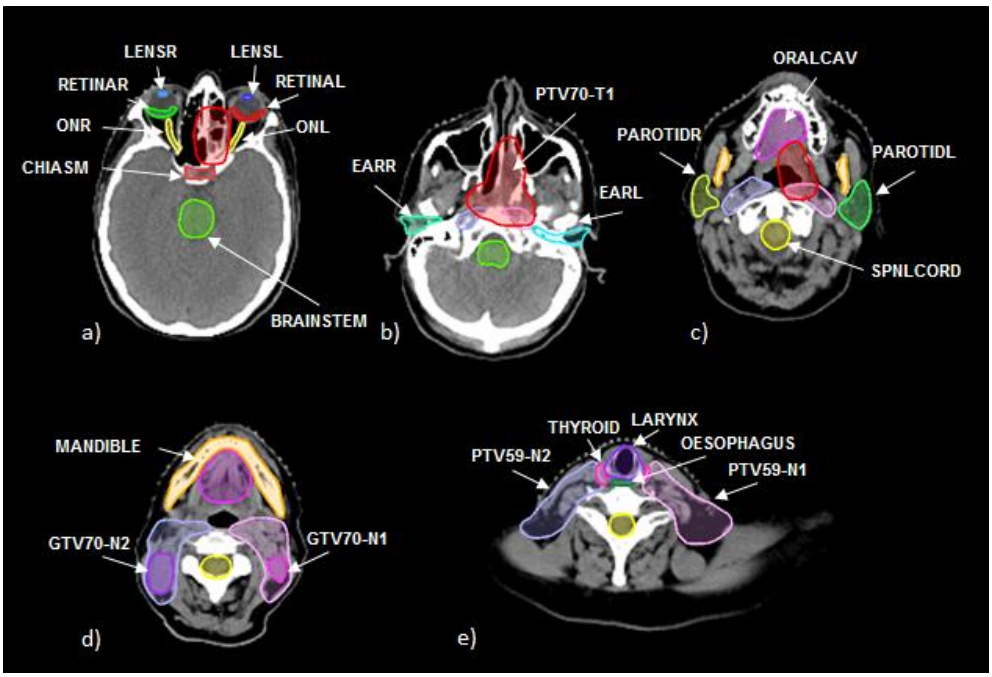

Figure S4.1 - PTVs and OARs delineated by the radiation oncologist on some of the planning CT images for NPC patient #8.

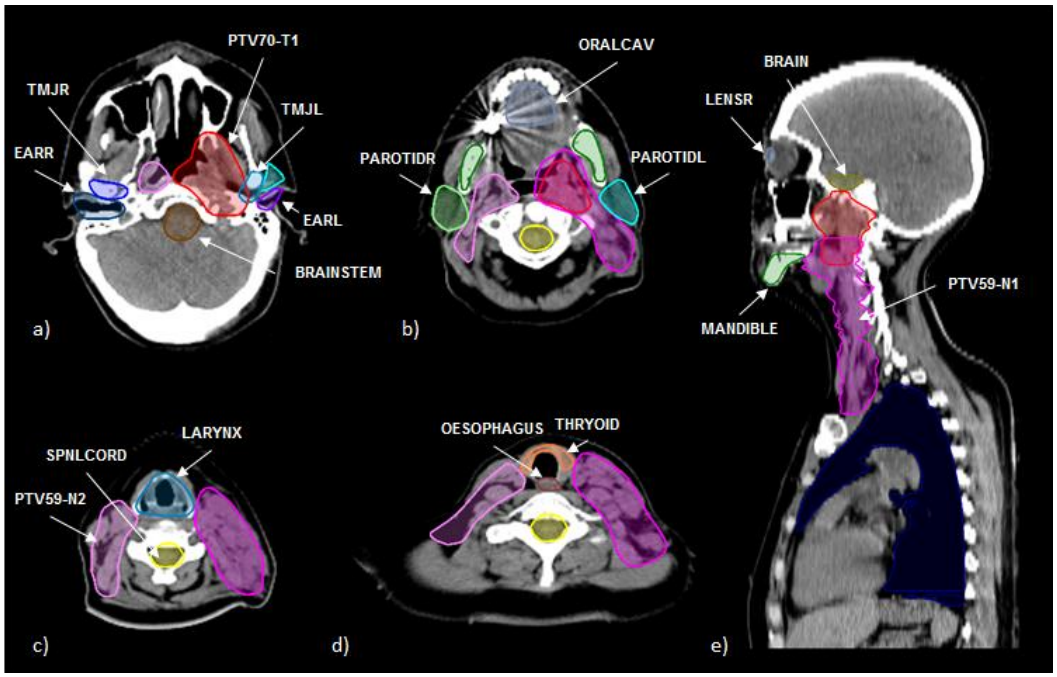

Figure S4.2 - PTVs and OARs delineated by the radiation oncologist on some of the planning CT images for NPC patient #14.

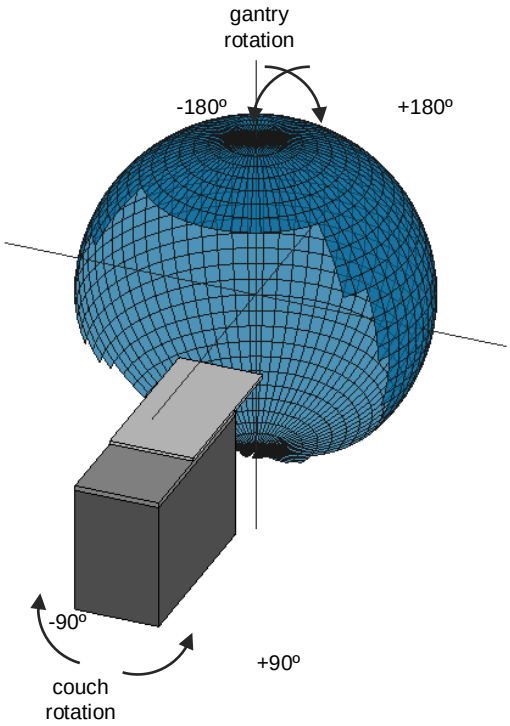

Figure S4.3 - Representation of the beam directions considered for the non-coplanar beam angular optimization.

PTV group

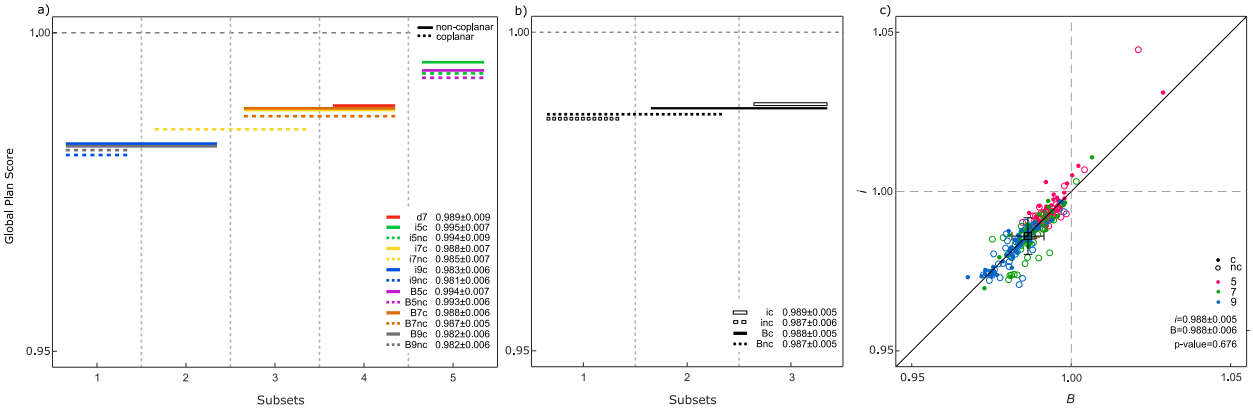

Figure S4.4 - Homogenous subsets resulting from post-hoc multiple comparisons test using the Tukey method with a level of significance of 5% of the mean PTV group scores for: a) coplanar and non-coplanar BAO of algorithms *B* (*B5c*, *B7c, B9c, B5nc*, *B7nc* and *B9nc*) and *i* (*i5c, i7c, i9c, i5nc, i7nc* and *i9nc);* b) the set of coplanar and non-coplanar algorithms *B* (*Bc* and *Bnc*) and *i* (*ic* and *inc*). c) Mean PTV group score of all sets of algorithms *i* and *B*.

Critical group

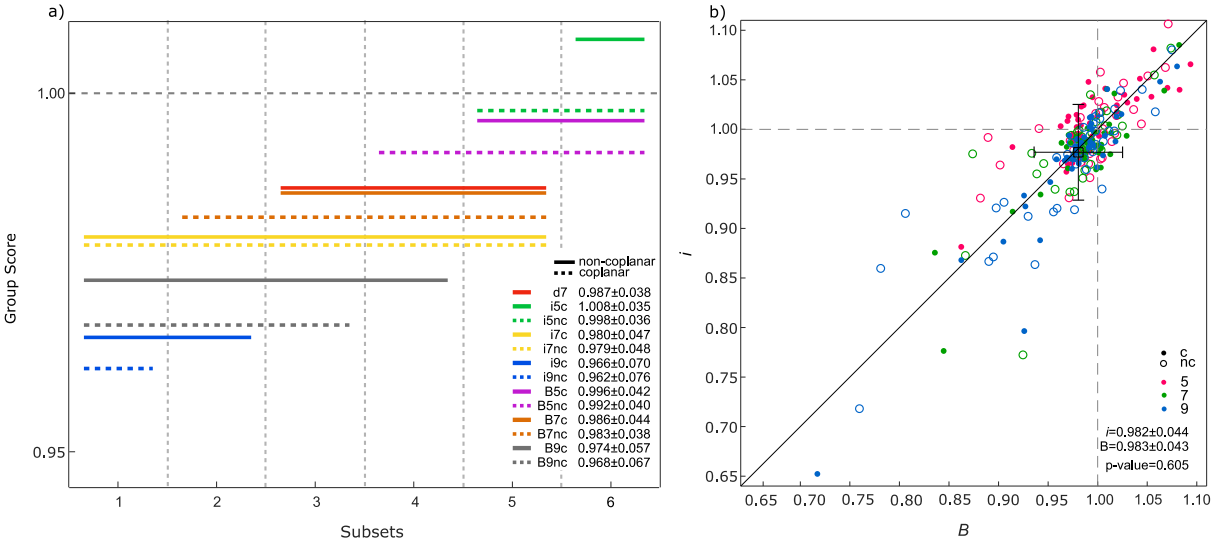

Figure S4.5 - a) Homogenous subsets resulting from post-hoc multiple comparisons test using the Tukey method with a level of significance of 5% of the mean Critical group scores for coplanar and non-coplanar BAO of algorithms *B* (*B5c*, *B7c, B9c, B5nc*, *B7nc* and *B9nc*) and *i (i5c, i7c, i9c, i5nc, i7nc* and *i9nc);* b) Mean Critical group score of all sets of algorithms *i* and *B*.

Optics group

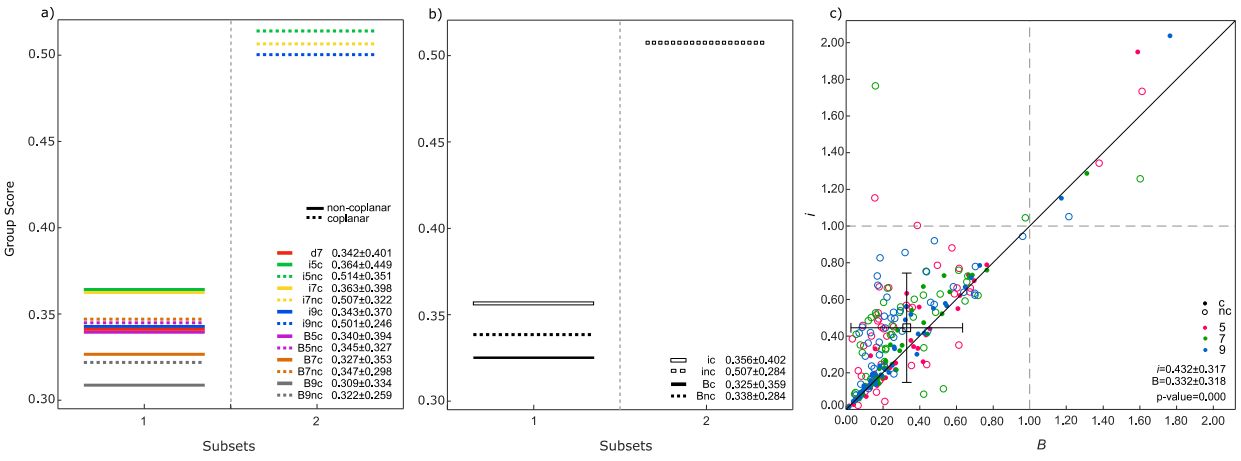

Figure S4.6 - Homogenous subsets resulting from post-hoc multiple comparisons test using the Tukey method with a level of significance of 5% of the mean Optics group scores for: a) coplanar and non-coplanar BAO of algorithms *B* (*B5c*, *B7c, B9c, B5nc*, *B7nc* and *B9nc*) and *i (i5c, i7c, i9c, i5nc, i7nc* and *i9nc);* b) the set of coplanar and non-coplanar algorithms *B* (*Bc* and *Bnc*) and *i* (*ic* and *inc*). c) Mean Optics group score of all sets of algorithms *i* and *B*.

DigestOral group

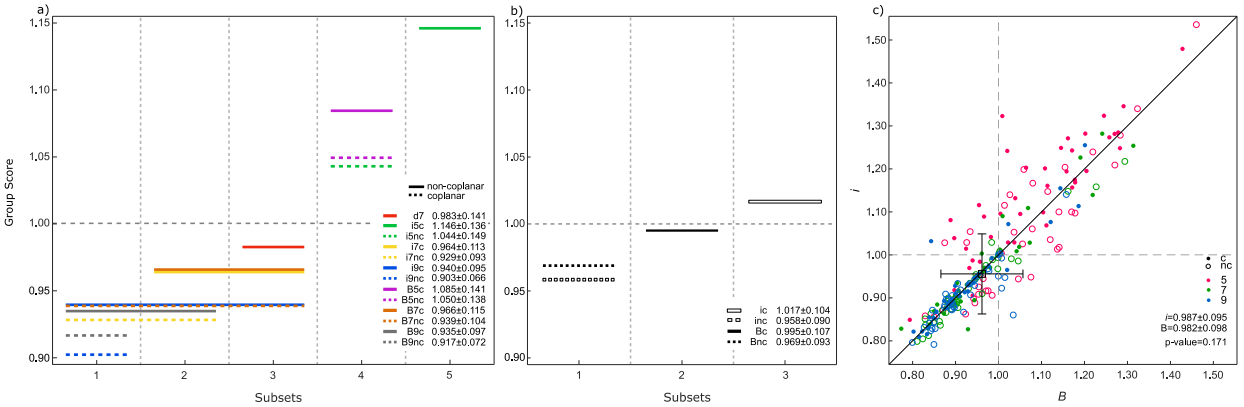

Figure S4.7 - Homogenous subsets resulting from post-hoc multiple comparisons test using the Tukey method with a level of significance of 5% of the mean DigestOtal group scores for: a) coplanar and non-coplanar BAO of algorithms *B* (*B5c*, *B7c, B9c, B5nc*, *B7nc* and *B9nc*) and *i (i5c, i7c, i9c, i5nc, i7nc* and *i9nc);* b) the set of coplanar and non-coplanar algorithms *B* (*Bc* and *Bnc*) and *i* (*ic* and *inc*). c) Mean DigestOral group score of all sets of algorithms *i* and *B*.

Bone group

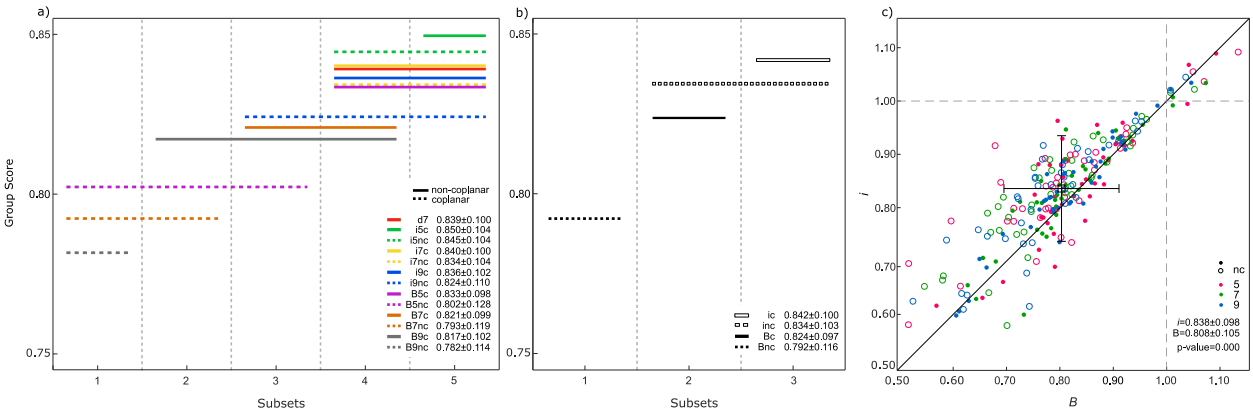

Figure S4.8 - Homogenous subsets resulting from post-hoc multiple comparisons test using the Tukey method with a level of significance of 5% of the mean Bone group scores for: a) coplanar and non-coplanar BAO of algorithms *B* (*B5c*, *B7c, B9c, B5nc*, *B7nc* and *B9nc*) and *i (i5c, i7c, i9c, i5nc, i7nc* and *i9nc);* b) the set of coplanar and non-coplanar algorithms *B* (*Bc* and *Bnc*) and *i* (*ic* and *inc*). c) Mean Bone group score of all sets of algorithms *i* and *B*.

chapter 5

Advantage of beam angle optimization in head-and-neck IMRT: patient specific analysis

IFMBE Proceedings, 2020, volume 76, pages 1256-1263

T Ventura[1,2,3], MC Lopes[1,2,3], H Rocha[3,4], BC Ferreira[3,5,6], J Dias[3,4]

[1]Physics Department, University of Aveiro, Aveiro, Portugal
[2]Medical Physics Department, IPOCFG, EPE, Coimbra, Portugal
[3]Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal
[4]Economy Faculty of University of Coimbra and Centre for Business and Economics Research, Coimbra, Portugal
[5]School Health Polytechnic of Porto, Porto, Portugal
[6]I3N, Physics Department, University of Aveiro, Aveiro, Portugal

Abstract

Radiation therapy (RT) main purpose is to eliminate, in a controlled way, all tumor cells sparing as much as possible the normal tissues. Intensity-Modulated Radiation Therapy (IMRT) is becoming the standard treatment technique in RT. Beam angle optimization (BAO) has potential to confer more quality to IMRT inverse planning process compared to manual trial and error approaches. In this study, the BAO advantages in head-and-neck patients are highlighted, using a patient specific analysis. Fluence optimization was done with Erasmus-iCycle multicriterial engine and BAO optimization was performed using two different algorithms: a combinatorial iterative algorithm and an algorithm based on a pattern search method. Plan assessment and comparison was performed with the graphical tool SPIDERplan. Among a set of forty studied nasopharynx cancer cases, three patients have been select for the specific analysis presented in this work. BAO presented plan quality improvements when beam angular optimized plans were compared with the equidistant beam angle solution and when plans based on non-coplanar beams geometries were compared with coplanar arrangements. Improvement in plan quality with a reduced number of beams was also achieved, in one case. For all cases, BAO generated plans with higher target coverage and better sparing of the normal tissues.

Keywords: Radiation therapy, beam angle optimization, plan assessment, head-and-neck cancer.

5.1. Introduction

Radiation therapy (RT) is one of the most important therapeutic options used in the battle against cancer. It makes use of ionizing radiation to eliminate in a controlled way the tumour cells, sparing as much as possible the adjacent normal tissues. Intensity-modulated radiation therapy (IMRT) is a radiation therapy technique that enables the generation of conformal dose distributions to the target volume by delivering non-uniform intensity fields from multiple directions. IMRT treatment planning uses inverse planning techniques, wherein an objective function containing the desired plan objectives guides the fluence map optimization (FMO) by scoring the goodness of the plan.[1] The FMO will determine the beam intensities for each of the selected angles. Most of the times, an empirical trial-and-error manual tuning of plan parameters (like weights, objectives or beam angles) is done until an acceptable plan is achieved. This planning framework is extremely dependent on the planner skills and experience and on the case complexity.[2] Moreover, it is not possible to guarantee that an optimal plan is found. The IMRT optimization process should be tackled by more reliable methods and algorithms such as multi-criteria optimization, beam angle optimization (BAO) and eventually by machine learning automated techniques.

BAO methods have contributed to the enhancement of the IMRT optimization process. A beam angle selection based on mathematical criteria may lead to important improvements in the quality of the plan dose distribution[3] that can be even more expressive if non-coplanar beam geometries are available in the optimization process.[4] The BAO problem mathematically described as a highly non-convex multi-modal optimization problem with many local minima[5] can be addressed separately or jointly with the FMO problem. For the coupled modality, the FMO solution guides the BAO problem along the optimization[6], where the best beam ensemble can be achieved using heuristic methods in exhaustive combinatorial searches over a discretized space search[6-8] or alternatively using pattern search methods[9] or multistart derivative-free optimization frameworks[4] that continuously explore the search space.

The purpose of this work is to highlight the advantages of BAO for the head-and-neck pathology following a patient specific analysis approach. From a set of forty nasopharyngeal (NPC) studied cases, particular patient cases were selected to show the improvements that can be obtained in plan quality when BAO plans are compared with the equidistant beam angle solution and when coplanar and non-coplanar BAO sets are confronted. To complete this patient specific analysis, a third case where it is possible to get improved plans diminishing the number of beams is also evaluated, to emphasize how the individual patient anatomy may influence the results.

5.2. Materials and methods

For this study, three cases were selected among a set of forty NPC clinical cases, previously studied. All cases had a simultaneous integrated boost prescription composed of two dose levels, where the tumour planning target volume (PTV) was prescribed with a dose of 70 Gy and the lymph nodes PTVs with a dose of 59.4 Gy. Spinal cord, brainstem, chiasm, optical nerves, retinas, lens, parotids, oral cavity, larynx, oesophagus, ear canals, temporal mandibular joints, mandible, brain, pituitary gland, thyroid and lungs were also contoured and defined as organs-at-risk (OAR).

IMRT optimization was performed by Erasmus-iCycle, a multicriterial dose calculation engine guided by a wish-list defined a priori, that automatically generates a single Pareto solution for a given set of beams.[10] All plans were initially optimized using the most clinically used beam angle configuration: 7 coplanar equidistant beams (*d7*). BAO was performed for coplanar and non-coplanar beam geometries of 5, 7 and 9 beams with two algorithms. The first (algorithm *i*) is included in

Erasmus-iCycle and it is based in a discrete iterative combinatorial approach.[10] The second algorithm, named as algorithm *B*, belongs to the continuous search space class and uses a pattern search method to find the best possible beam angle ensemble.[9]

For each selected NPC case, three plans were considered that highlight the BAO advantages. For the first case (patient #1), the *d7* plan was compared with the 9 non-coplanar beams plan of algorithm *B* (*B9nc*). For the second case (patient #2), the 9 coplanar beams plan and the 9 beams non-coplanar plan of algorithm *i* (*i9c* and *i9nc*, respectively) were compared side-by-side. For the third case (patient #3), the 5 non-coplanar beams plan of algorithm *B* (*B5nc*) and the 9 coplanar beams plan of algorithm *i* (*i9c*) were compared.

SPIDERplan, a graphical method based on a scoring approach,[11] was used for plan assessment and comparison. Targets and OARs are divided into groups. A score based on pre-defined planning objectives and relative weights defined according to the radiation oncologist clinical preferences is assigned to each structure. Global plan quality is computed as a weighted sum of all structures' scores. Graphically, a customised radar plot is used to represent all dosimetric information, where plan evaluation can be done visualizing all structures, in a Structures Plan Diagram (SPD), or just considering the defined groups, in a Group Plan Diagram (GPD). Each group can also be assessed with more detail in partial group radar plots named Structures Group Diagrams (SGD), where only the structures belonging to the group and a partial group score are displayed.

For this study, SPIDERplan configuration included six groups of structures: the PTV group composed by the PTVs, the Critical group constituted by the spinal cord and the brainstem, the Optics group including the chiasm, the optical nerves, the retinas and the lens, the DigestOral group composed by the parotids, the oral cavity, the oesophagus and the larynx, the Bone group composed of the temporal mandibular joint, the mandible and the ear canals and Other group considering the brain, the pituitary gland, the thyroid and the lungs. The groups were assigned with relative weights of 50%, 30%, 10%, 5%, 3.5% and 1.5%, respectively. The score of each structure is given by the ratio between the clinical tolerance criteria and the planned dose. A score value of one is achieved when the planned dose is equal to the structure tolerance criteria. Higher score values mean that these tolerances were surpassed and the best plans will have scores less than one, tending towards the centre of the radar plot.[11]

5.3. Results

Global plan comparison was done with SPIDERplan GPDs (left diagrams in Figures 5.1, 5.2 and 5.3), while individual group evaluation was performed with SPIDERplan SGDs (right diagrams in Figures 5.1, 5.2 and 5.3). The selection of the groups to be depicted in each figure was based on the visual analysis of the GPD and on the corresponding relative weight importance value.

For patient #1, *d7* and *B9nc* plans were compared, Figure 5.1. *B9nc* plan achieved the best global plan score with an expressive percent difference of -7% from *d7* score. The largest difference between the tolerance and the planned dose of these plans was obtained for the Optics group. For *d7* plan the score of both lenses was higher than one, meaning that the tolerance dose criteria was not accomplished. Also, for the Critical group, a percent score difference of 9% can be observed between the two plans.

For patient #2, the best global plan score was obtained with the non-coplanar plan *i9nc* with a percent difference score of -5% from the coplanar plan i9c, that presented a score of 1.004 (Figure 5.2). Also for *i9c,* score values higher than one were achieved by the Optics group, the DigestOral group and the Other group, while for *i9nc* only the Other group score was out of tolerance. The reason for these global results was found in the SGDs. Both plans presented scores very near the tolerance or out of tolerance for the lenses (Optics group SGD), for the parotids and for the oral cavity (DigestOral group SGD). Nevertheless, plan *i9nc* was much better than *i9c* plan, since for the

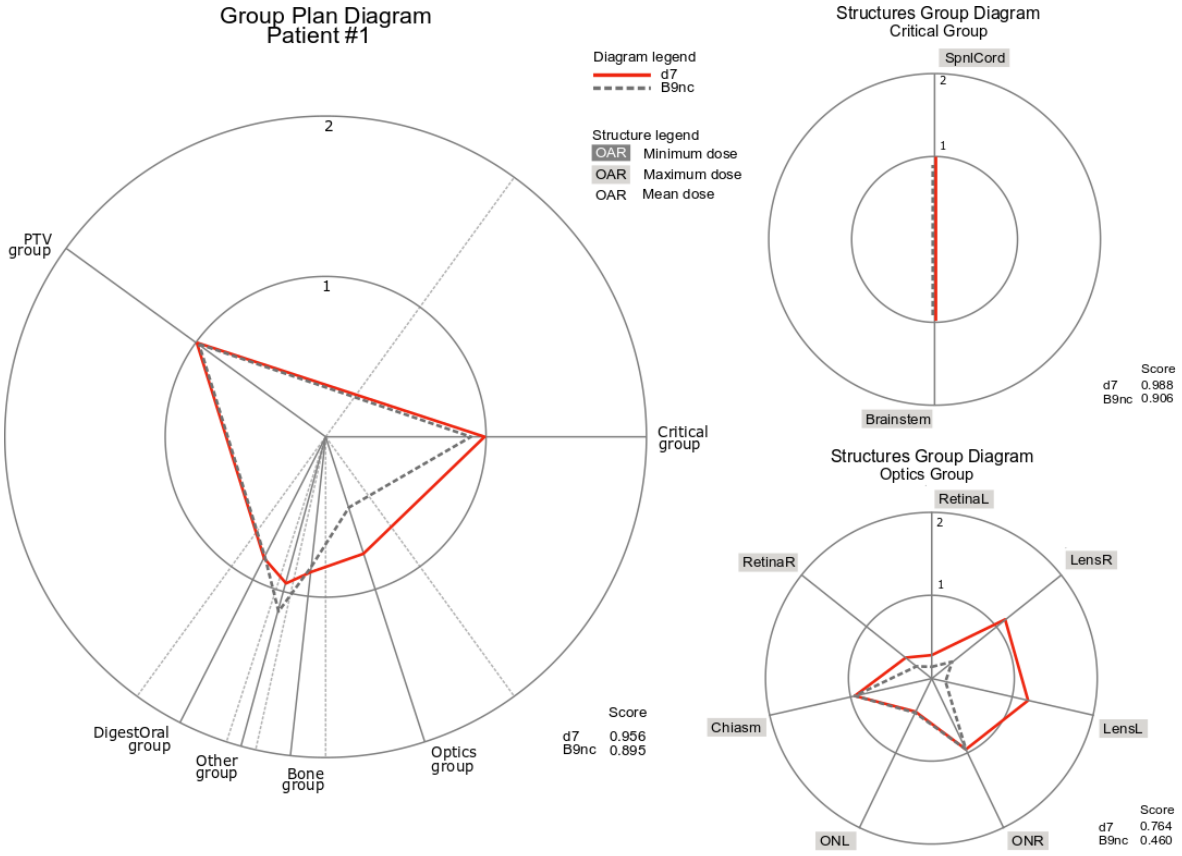

Figure 5.1 - SPIDERplan group plan diagram of patient #1 and structures group diagrams for Critical and Optics groups for 7 equidistant beam angles (*d7*) and 9 non-coplanar beams plan of algorithm B (*B9nc*).

mentioned structures it presented score values near one or just slightly higher than one, while *i9c* got score values well above the defined tolerances.

Figure 5.3 compares the quality of *B5nc* and *i9c* plans for patient #3. Globally, plan *B5nc* attained the best global plan score with a percent global score difference of -4% compared to *i9c* plan. This result is obtained even if for the DigestOral group *B5nc* plan is worse than plan *i9c,* due to the relatively lower weight of this group. The main reason for the best score of plan *B5nc* is the almost vanishing scores presented for the optics structures (dotted line at the Optics SGD centre). The topographic relation of the PTV and the optics structure, well separated for this patient, explains this result.

5.4. Discussion

In this study, patient specific plan assessment was used to highlight the advantages of BAO for head-and-neck cases.

Among a set of forty NPC clinical cases, three patients were chosen. For patient #1, important plan quality improvements were achieved when the plan using optimized directions was compared with the equidistant beam angle solution, the standard beam configuration used in the clinical routine. For this patient, the plan generated with BAO presented non-coplanar beam incidences and a higher number of beams than equidistant beam angle solution. This has improved the quality of dose distribution by further sparing important OARs.

The utilization of non-coplanar beams in RT is a very popular approach among medical physicists and mathematicians. There is an intrinsic perception that non-coplanar beam incidences

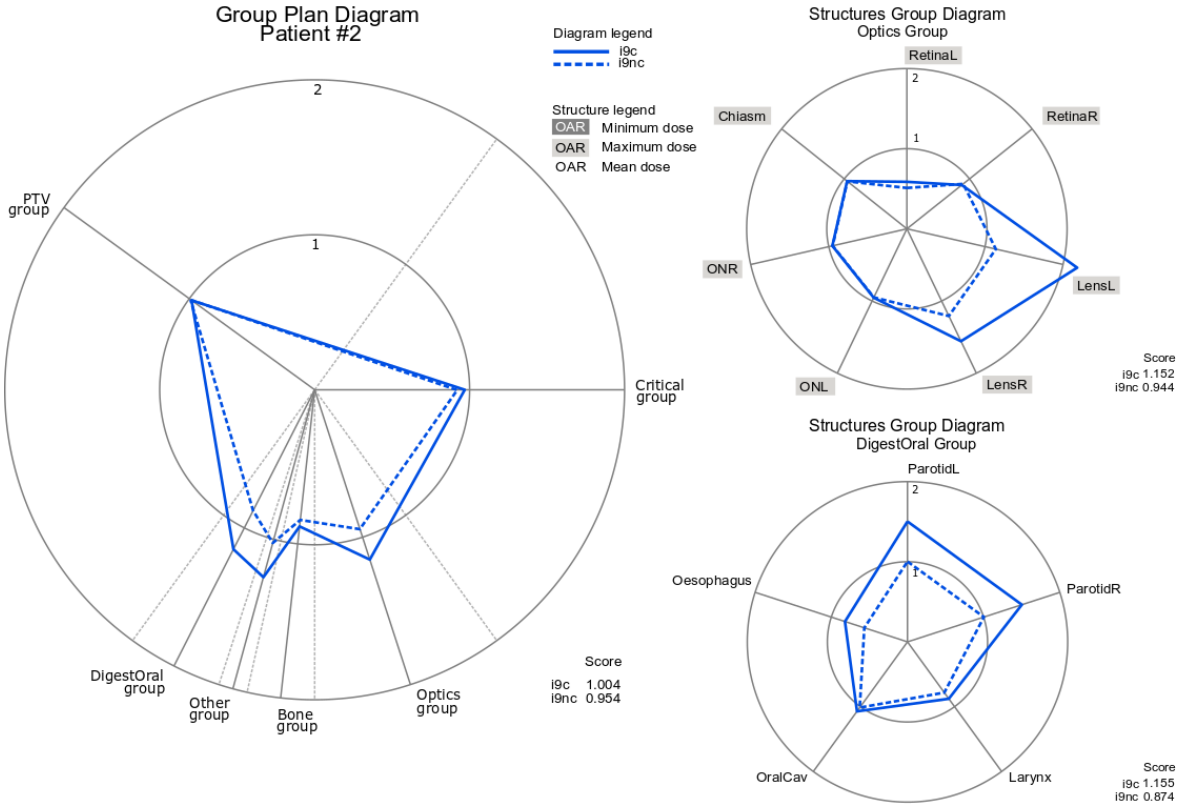

Figure 5.2 - SPIDERplan group plan diagram of patient #2 and structures group diagrams for Optics and DigestOral groups for algorithm i with 9 coplanar beams (*i9c*) and 9 non-coplanar beams (*i9nc*).

may enhance the quality of the plan. BAO has a crucial role in the selection process of the best beam arrangement, due to the high number of possible combinations and geometry complexity. Promising results with non-coplanar BAO have been achieved for different pathology sites such as: lung, liver, head-and-neck and intracranial tumours.[4,12-14] Nevertheless, depending on site complexity, a patient-by-patient assessment may be required to fully evaluate the real impact of the achieved improvements. For patient #2, i9c and i9nc plans were compared enhancing the advantage of non-coplanarity. The higher score group differences were found for the Optics group (lenses and left retina) and for the DigestOral group (oesophagus, parotids and larynx), where the structures were located very near or even contiguous to the PTVs.

The number of beams needed to achieve a 'good' dose distribution is also an interesting topic due to treatment delivery efficiency reasons. The planner intuition will always predict that plan quality will tend to increase with the number of beams available for optimization. However, the amount of fluence modulation caused by the addition of new beams has a physical threshold, where beyond that level no further plan quality improvements occur. BAO has then an important role to play, achieving better treatment plans with a smaller number of beams. This scenario occurred for patient #3 where plan B5nc was better than i9c, due to the significant improvement achieved in the optics structures that were completely spared when using fewer non-coplanar beam orientations.

For all patients, PTV group scores were always below one. The high target coverage index may be related with the quality of the FMO and to the definitions of the multicriterial IMRT optimization engine wish-list where a primordial importance was assigned to the PTVs. Also, for the PTV group, small score differences were obtained among the compared plans. For the OARs, the largest score group differences were obtained for the Critical group (patient #1), for the Optics group (patients #1, #2 and #3) and for the DigestOral group (patients #2 and #3), that contained structures of vital

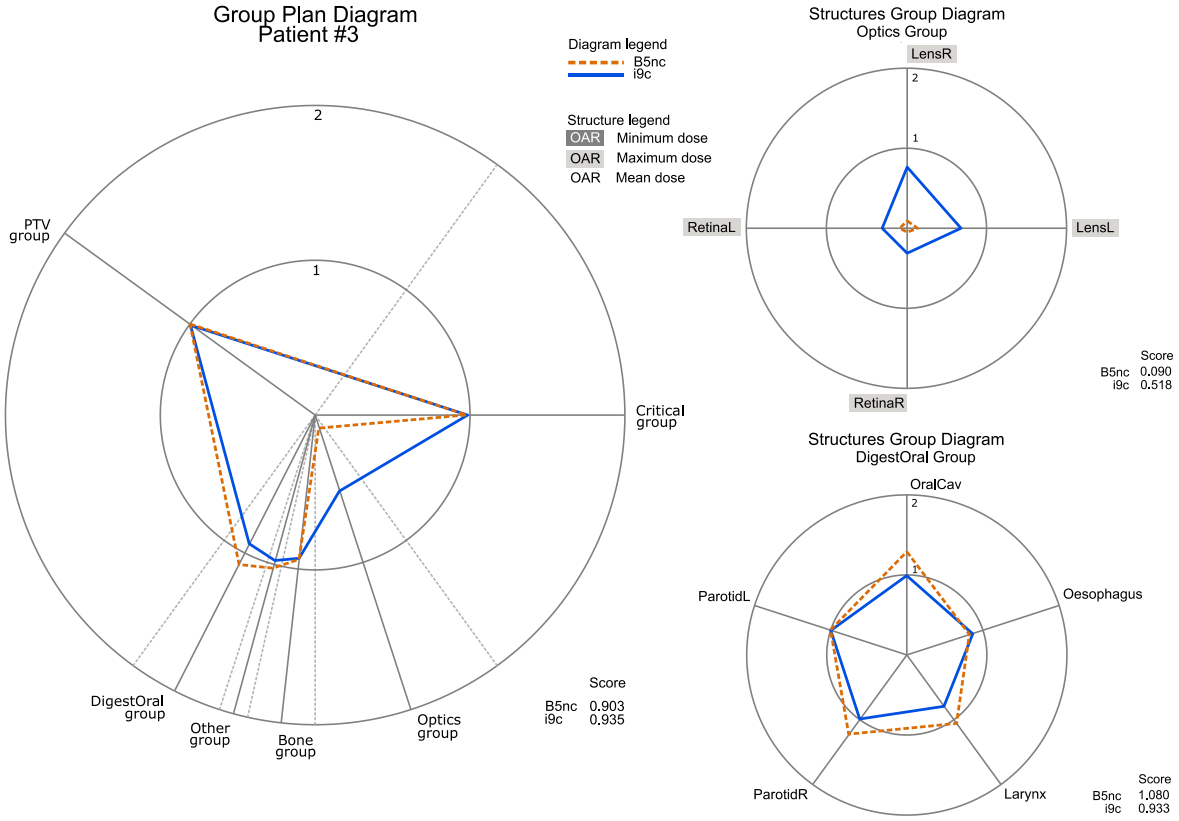

Figure 5.3 - SPIDERplan group plan diagram of patient #3 and structures group diagrams for Optics and DigestOral groups for algorithm B with 5 non-coplanar beams (*B5nc*) and for algorithm *i* with 9 coplanar beams (*i9nc*).

importance, as the spinal cord and the brainstem, and with great impact in the patient quality of life, as the lens, the optical nerves, the parotids or the oral cavity. These findings help to support two key aspects: the main contribution of BAO was in further sparing the OARs while maintaining good PTV coverage.

The use of BAO never worsens the treatment plan quality though for some patients, the improvements achieved may be negligible. However, if BAO can be done without the planner's intervention, then it is worth to be considered for all patients, since some of them will greatly benefit from this optimization.

The potential enhancements that an efficient BAO, made in reasonable computing times, may bring to the quality of static IMRT plans, can reopen the discussion about what is the most appropriate inverse planning technique for a given patient treatment.

5.5. Conclusions

In this study the benefits of BAO for head-and-neck patients were addressed. Patient specific analysis of NPC cases were used to highlight the improvements achieved by plans with BAO when compared with IMRT plans using equidistant beam angle directions. Furthermore, the advantages of non-coplanar over coplanar beam plans and the importance of BAO in the generation of plans with a small number of beams was highlighted. This study also showed that BAO did not compromise the PTV coverage and that the most important contribution was in the sparing of the normal tissues, taking into account the anatomy specificities of each patient.

5.6. Acknowledgments

The authors would like to express their gratitude to Sebastiaan Breedveld and Ben Heijmen for making available the multicriterial optimization of IMRT plans system Erasmus-iCycle and their valuable support and coaching along the work.

This work was supported by project grant POCI-01-0145-FEDER-028030 and by the Fundação para a Ciência e a Tecnologia (FCT) under project grant UID/Multi/00308/2019.

No potential conflict of interest nor any financial disclosures must be declared.

5.7. References

1. ICRU. International Commission on radiation units and measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Geneva; 2010. *J ICRU* 10(1):1–106
2. Thieke C, Küufer KH, Monz M. et al. A new concept for interactive radiotherapy planning with multicriteria optimization: First clinical evaluation. *Radiother and Oncol*. 2007;85(2):292-298 (2007).
3. Das SK, Marks LB. Selection of coplanar or noncoplanar beams using three-dimensional optimization based on maximum beam separation and minimized nontarget irradiation. *Int J Radiat Oncol Biol Phys.* 1997;38(3):643-55.
4. Rocha H, Dias J, Ventura T, Ferreira B, Lopes MC. A derivative-free multistart framework for an automated noncoplanar beam angle optimization in IMRT. *Med Phys*. 2016;43(10):5514-5526.
5. Ehrgott M, Holder A, Reese J. Beam selection in radiotherapy design. *Linear Algebra Appl*. 2008;428(5):1272:312.
6. Craft D. Local beam angle optimization with linear programming and gradient search. *Phys Med Biol. 2007;52(7):*127-135. (2007).
7. Dias J, Rocha H, Ferreira BC, Lopes MC. Simulated annealing applied to IMRT beam angle optimization: a computational study. *Phys Med*. 2015;31(7):747-756.
8. Bangert M, Unkelbach J. Accelerated iterative beam angle selection in IMRT. *Med Phys*. 2016;43(3):1073–82.
9. Rocha H, Dias J, Ferreira BC, Lopes MC. Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method. *Phys Med Biol*. 2013;58(9):2939-2953.
10. Breedveld S, Storchi P, Voet P, Heijmen B. iCycle: integrated, multi-criterial beam angle and profile optimization for generation of coplanar and non-coplanar IMRT plans. *Med Phys*. 2012;39(2):951-963.
11. Ventura T, Lopes MC, Ferreira BC, Khouri L. SPIDERplan: A tool to support decision-making in radiation therapy treatment plan assessment. *Rep Pract Oncol Radiother*. 2016;21(6):508-516.
12. Dong P., Lee P., Ruan D et al. 4π Non-Coplanar Liver SBRT: A Novel Delivery Technique. *Int J Radiat Oncol Biol Phys.* 2013;85(5):1360–1366.
13. Bangert M, Ziegenhein P, Oelfke U. Comparison of beam angle selection strategies for intracranial IMRT. *Med Phys*. 2013;40(1):011716
14. Lyu Q, Yu V, Ruan D. A novel optimization framework for VMAT with dynamic gantry couch rotation. *Phys Med Biol*. 2018;63(12):125013.

chapter 6

Non-coplanar optimization of static beams and arc trajectories for intensity-modulated treatments of meningioma cases

submitted to Physica Medica in October 2019

T Ventura[1,2,3], H Rocha[3,4], BC Ferreira[3,4], J Dias[3,4], MC Lopes[1,2,3]

[1]Physics Department, University of Aveiro, Aveiro, Portugal
[2]Medical Physics Department, IPOCFG, EPE, Coimbra, Portugal
[3]Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal
[4]Economy Faculty of University of Coimbra and Centre for Business and Economics Research, Coimbra, Portugal
[5]School Health Polytechnic of Porto, Porto, Portugal
[6]I3N, Physics Department, University of Aveiro, Aveiro, Portugal

Highlights

- Non-coplanar beam angular and arc trajectory optimizations were applied to meningioma
- Beam angle optimization was guided by SPIDERplan global plan score
- Automated arc trajectory selection may lead to improvements in plan quality

Abstract

Purpose: Two methods for non-coplanar beam direction optimization, one for static beams and another for arc trajectories, were proposed for intracranial tumours. The optimized plans were compared with the clinical ones.

Methods: Ten meningioma cases already treated with stereotactic intensity-modulated radiation therapy were selected for this retrospective planning study. Algorithms for non-coplanar beam angle optimization and arc trajectory optimization were used to generate the corresponding *BAO* and *ATO* plans, respectively. SPIDERplan plan quality score was used to guide the beam angle optimization process. The arc trajectory optimization algorithm was based on a two-step approach. For each patient, the clinical plans (*CLIN*) and coplanar VMAT plans (*VMAT*) were also generated. To harmonize plan comparisons, all plan optimizations were performed in an automated multicriterial optimization calculation engine and dosimetric plan quality was assessed.

Results: *BAO* and *ATO* plans presented, on average, moderate quality improvements over *VMAT* and *CLIN* plans. Nevertheless, while *BAO* and *CLIN* plans assured a more efficient OARs sparing, the *ATO* and *VMAT* plans presented a higher coverage and conformity of the PTV.

Conclusion: Globally, all plans presented high-quality dose distributions, including the clinical ones, corollary of the high skills and accumulated clinical experience of the planners. No statistically significant quality differences were found, on average, between *BAO*, *ATO* and *CLIN* plans. However, automated plan solution optimizations (*BAO* or *ATO*) saves the time cost associated with manual planning. In a patient-specific analysis, plan quality improvements were achieved with *ATO* plans, demonstrating the possible benefits of this automated optimized delivery technique.

Keywords: Beam angle optimization, non-coplanar, arc trajectory, VMAT

6.1. Introduction

In radiation therapy, non-uniform intensity field techniques are well-established for almost all cancer pathologies since they allow the delivery of highly conformal dose distributions to the target(s) while minimizing the injury to the organs-at-risk (OAR). The calculation of non-uniform beam intensities is done using inverse planning, where plan objectives are specified by means of physical or biological descriptors in an objective function that guides the fluence map optimization (FMO) process.[1] Usually, the planning optimization is performed through a trial-and-error manual tuning of plan parameters until an acceptable plan is obtained.

For conventional C-arm linear accelerators, this type of treatment techniques can be delivered through multiple modulated static or dynamic radiation fields (intensity-modulated radiation therapy - IMRT) or through continuously modulated radiation arcs combining the variation in dose rate, gantry speed and aperture shape (volumetric modulated arc therapy - VMAT). For most tumour sites, equivalent plan quality can be achieved by IMRT or VMAT. Nevertheless, VMAT treatments are usually more efficient requiring fewer monitor units and thus shorter delivery times.[2]

Most of IMRT and VMAT treatments are still performed using equidistant coplanar static beams or coplanar arcs. As these approaches typically obtain acceptable treatment plans, beam angle optimization methods are still not popular among the clinical community. However, when non-coplanar geometries are included in the optimization, improved normal tissue sparing, target conformity and steeper dose gradients can be achieved. Indeed, an appropriate beam assembly or arc trajectory selection may lead to improvements in the dosimetric quality of the plans.[3,4]

Beam angle optimization is complex, time-consuming and it often presents non-intuitive solutions. Mathematically, it is defined as a highly non-convex multi-modal optimization problem with many local minima,[5-7] requiring optimization methods that avoid being trapped in a local minimum. For IMRT, the beam angle optimization problem considering non-coplanar geometries has been extensively studied for brain,[3,8-10] head-and-neck,[10-12,14-16] lung,[17] gastric,[12] liver,[14,18,19], pancreas,[10] cervix[14] and prostate[10,12,13] sites. The reported beam angle optimization methods can be grouped into two classes. In the first class, beam angle selection and the FMO processes are independent and are addressed sequentially. The beam angle optimization process is normally driven by geometrical or dosimetric metrics or by methods that require some prior knowledge of the problem.[3,12,13] These methods are computationally efficient, but the resultant beam angle ensemble does not guarantee the optimality of the plan solution. In the second class of methods, the beam angle optimization and the FMO processes are simultaneously solved. The FMO is used to guide the beam angle optimization by assessing the goodness of the plan. The beam angle optimization problem can be formulated by considering a combinatorial search for the best ensemble over a discretized space search or by a continuous space search optimization. For the first approach, searches for the best beam combination can be done using heuristic methods,[8,11] iterative beam angle optimization methods,[9,10,14,17,18] or sparse optimization.[19] For the second approach, beam angle optimization can be done considering derivative-free optimization frameworks. [16,20]

In VMAT, optimization of non-coplanar beam geometries is considered, in most published works, for brain lesions,[4,22-30] and breast/chest wall irradiation.[27,28,31-33] Non-coplanar beam optimization for head-and-neck tumours,[21,34-36] lung,[25,29] liver[29] and prostate[27] has also been reported over the past years. The first planning studies using one or more arcs with static couch[23,31,33] or planner-defined arc trajectories[4,21,32] confirmed the benefits of non-coplanar incidences by OAR sparing. Recently, automated techniques have also been investigated. The simultaneous movement of the gantry and the couch while the beam is being modulated by the continuous movement of the multi-leaf collimator (MLC) grant to the plan optimization process additional degrees of freedom that

may result in promising improvements of plan quality. Similarly to the IMRT beam angle optimization problem, the arc trajectory optimization problem can also be divided into two classes of methods: those that decouple arc trajectory optimization from FMO and those that jointly address the two optimization problems. In the first class, geometric and/or dosimetric heuristics[22,24,26,35] are used to define feasible beam orientations. After that, the best delivery trajectory is determined. Beam grouping techniques[22,26] or graph search techniques, such as those proposed by Dijkstra[24] or the A* algorithms[35] that intend to solve the travelling salesman problem, are used to generate multiple sub-arcs (arcs with static couch or static gantry angles) paths or continuous gantry/couch angle paths, respectively. The VMAT plan is posteriorly optimized along the trajectory in a distinct optimization phase. In the second class of methods, fluence-based methods are used to guide the arc trajectory optimization problem. In some published works, non-coplanar beam angles, obtained from the IMRT fluence-based beam angle optimization problem, are used as anchor points for the path definition. The final arc trajectory is determined by solving the travelling salesman problem.[25,34] Although promising, these methods do not fully guarantee the optimality of the plan solution over the whole trajectory. Alternatively, techniques combining iteratively sparse solutions of feasible beams with graph search optimization for the trajectory definition[28] or applying Monte Carlo Tree Search algorithms[30] have been proposed. More recently, the anchor point concept was adapted to improve the dosimetric objectives over the whole arc trajectory, by including these optimal incidences in an iterative combinatorial beam angle optimization process that will add new anchor points until the beam path is completely defined.[36] Mixed approaches, that apply methods from both classes during the arc trajectory optimization phases, have also been recently presented.[27,29]

In previous works, the static beam angle optimization problem for head-and-neck pathologies has been addressed.[37,39] Two beam angle optimization algorithms belonging to the discrete and continuous space search approach optimization classes were compared using a dedicated plan assessment tool.[38] In the present work, the best direction/trajectory selection problem is addressed for intracranial tumour cases. For these cases, a high level of target coverage and conformity is required for plan approval. Non-coplanar beams or coplanar arcs combined with inverse planning optimization techniques are normally used in the clinical routine. In this study, the potential improvements of the automatic selection of the irradiation directions were investigated for a sample of ten meningioma cases. Algorithms for beam angular optimization and for arc trajectory optimization were applied. A global plan score, based on the dosimetric parameters of the anatomic structures and on the radiation oncologist clinical preferences,[38] was used to guide the non-coplanar beam angle optimization problem. For the arc trajectory optimization, a new two-step approach using optimized non-coplanar static beam directions as anchoring points of the arc path was proposed.

6.2. Materials and methods

6.2.1. Patient data

Ten meningioma cases already treated with stereotactic IMRT were selected for this study. All structures were delineated using two imaging modalities: computed tomography and magnetic resonance images that were conveniently fused. Apart from the planning target volume (PTV), the brainstem, the lens, the retinas, the optical nerves, the chiasm, the pituitary gland and the cochleas were also contoured by the radiation oncologist. The PTV was prescribed with doses of 50.4 Gy, 54.0 Gy, 59.4 Gy or 60.0 Gy delivered in 28, 30 or 33 fractions and the organs-at-risk (OAR) tolerance doses were established in agreement with the institutional protocol for the intracranial tumours treated with stereotactic IMRT (Table S6.1 in the Supplementary material).

6.2.2. Plan generation and optimization

The FMO was performed by Erasmus-iCycle IMRT multicriterial optimization framework,[14] that is guided by a wish-list using a constraint-based method (2pεc method) to generate a single Pareto solution in an automatic way.[40] A pencil-beam dose algorithm with equivalent path length inhomogeneity corrections is used to compute the dose distribution with a beamlet size of 2.5×5.0 mm^2 and with 10 mm of scatter radius. As no fluence segmentation is done during or after the optimization phase by Erasmus-iCycle, VMAT was simulated through no sequenced intensity-modulated static beams. Thus, the continuous gantry and MLC motions were approximated by 21 equidistant static beams distributed over the trajectory.[41]

The wish-list is composed of a set of clinical constraints and objectives. The constraints must be fulfilled by the multicriterial optimization algorithm and the objectives must be assigned with an optimization priority. For the meningioma cases, the wish-list was composed of six constraints and sixteen prioritized objectives divided into two optimization levels (Table S6.2 in the Supplementary material). The objective function associated with the PTV was the Logarithmic Tumour Control Probability (LTCP) function, regulated by a cell sensitivity parameter (α). An α value of 0.75 was applied to guarantee good coverage, i.e. that at least 95% of the PTV volume receives the prescription dose (D_p). The criteria considered for each OAR in the optimization levels were established according to the organ architecture. For the first optimization level, maximum dose objectives for the organs with serial architecture and mean dose objectives for the organs with parallel architecture were considered. For the second level, mean dose and maximum dose objectives were added for the organs with serial and parallel architectures, respectively. The generalized Equivalent Uniform Dose (gEUD) with a value of the tissue-specific parameter that describes the volume effect (*a*) equal to 15 and 6 was also used to minimize the maximum and the mean doses of the lenses and of the cochleas, respectively.

6.2.3. Beam angle optimization

The non-coplanar beam angle optimization of IMRT plans was performed using a derivative-free parallel multistart framework approach based on a continuous exploration of the search space to find the best beam ensemble.[20] To prevent possible collisions between the gantry and the treatment couch, avoidance beam orientations were defined.

The adopted beam angle optimization procedure takes advantage of relevant properties of the beam angle optimization search space. One of the main features is the symmetry of the solutions in the beam angle optimization search space which allows a drastic reduction of the space to be searched and, simultaneously, to define a strategy of partition into several sub-regions of this reduced search space for a parallel multistart exploration.[20] Each of the defined sub-regions is still a large search region of a highly non-convex problem with many local minima being, therefore, locally explored resorting to a derivative-free algorithm.[42] The measure used to compare different beam ensembles, and thus to drive the beam angle optimization search, was the SPIDERplan global plan score described in section 6.2.6. The parallel multistart framework using a derivative-free algorithm guided by this global score for the local search procedures is described in more detail in the Supplementary material.

6.2.4. Arc trajectory optimization

The non-coplanar arc trajectory optimization was done assuming that the gantry and the couch can rotate simultaneously with different rotation speeds, enabling the definition of highly non-coplanar trajectories. In this work, a two-step approach combining dosimetric considerations and geometric features, was followed. In the first step, the parallel derivative-free multistart framework was used to find feasible non-coplanar beam angles that will be defined as anchor points of the new arc trajectory (red dots in Figure 6.1a). In the second step, the anchor points were connected through

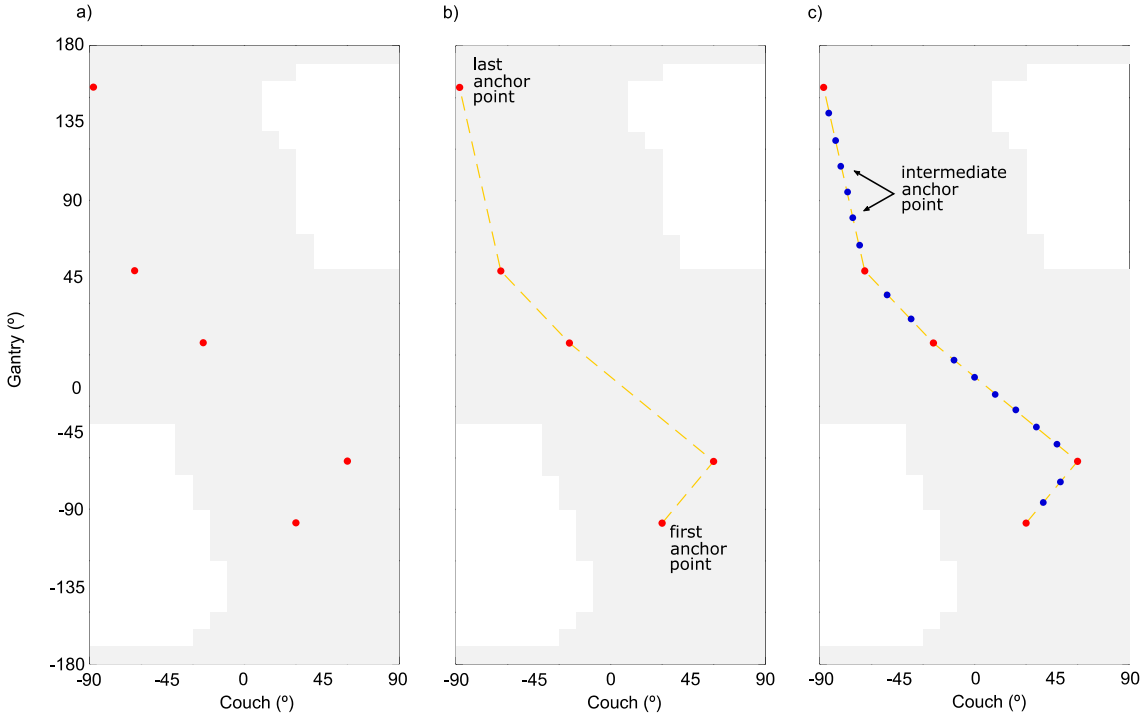

Figure 6.1 - Arc trajectory optimization algorithm phases. a) 5 non-coplanar beam angle optimized solution that defines the initial anchor points (red dots). b) linear trajectories (yellow dashed lines) between the anchor points. c) intermediate anchor points (blue dots) definition.

the definition of linear trajectories (yellow dashed lines in Figure 6.1b). New anchor points, placed equidistantly, were added to the trajectory so that the arc was divided into 21 arc sectors (blue dots in Figure 6.1c) to mimic a true VMAT technique.

The gantry and couch movements, defined by the linear paths passing through the anchor points, were outlined taking into account total delivery treatment time. It was established that the gantry trajectory should be always continuous without any inversions in its rotation and never exceeding a 360º arc. For the couch rotation movement, it was also defined that it could be reversed or even halted during the arc delivery. It was also decided that the trajectory must start at the anchor point closest to the search space coordinates (gantry=-180, couch=-90) and, when moving from one anchor point to another, the smallest distance between points must be considered. For anchor points placed inside the avoidance regions, a shift in the couch position to the nearest possible beam incidence was done.

6.2.5. Study design

All plans were automatically generated in iCycle-Erasmus. Based on the wish-list template, IMRT plans were generated employing the beam configurations used in the clinical plans (*CLIN*) and five non-coplanar beams optimized by beam angle optimization algorithm (*BAO*). All *CLIN* plans had been manually optimized in the clinical routine using 4-6 non-coplanar beams. The number of beams was defined for each case by the planner, considering case complexity. VMAT plans were produced using equivalent coplanar arcs (*VMAT*) and equivalent trajectories based in five initial non-coplanar anchor points optimized by the proposed arc trajectories optimization process (*ATO*). To further explore the potential advantage of arc trajectory optimization approach, an additional plan based on 9 initial anchor points (*ATO9*), was added to the initial plan sets library for a specific patient case.

6.2.6. Plan assessment and comparison

Plan assessment and comparison was performed with a graphical method named SPIDERplan.[38] This tool uses customised radar plots to graphically display a scoring approach that considers both target coverage and conformity and individual OAR sparing. Depending on their clinical importance, targets and OARs are divided into groups and a score based on the pre-defined planning objectives and relative weights is determined. A global plan score is calculated as a weighted sum of the structures' individual scores over all groups:

$$\text{Global plan score} = \sum_{i} w_{group(i)} \sum_{j} w_{struct(j)} \text{Score}_{struct(j)} \tag{6.1}$$

where $w_{struct(j)}$ and $\text{Score}_{struct(j)}$ are the relative weight and the score of structure j, respectively, and $w_{group(i)}$ is the relative weight of group i.

For the PTVs, the coverage and the conformity concepts, normally used by the radiation oncologist to assess the target's dose distribution for intracranial cases treated with stereotactic irradiation techniques, are included in the score:

$$\text{Score}_{PTV} = \frac{1}{2}\left(\frac{D_{TC,PTV}}{D_{P,PTV}} + \frac{0.6}{PCI}\right) \tag{6.2}$$

where $D_{TC,PTV}$ corresponds to the tolerance criteria for the PTV (in this case the dose in 95% of the PTV that should receive at least the prescribed dose, Table S6.1 in Supplementary material) and $D_{P,PTV}$ is the planned dose in the PTV. PCI is the Paddick[43] plan conformity index that, for conformal plans, should be above 0.6:

$$PCI = \frac{V_{PTV,100\%}^2}{V_{PTV}\, V_{External,100\%}} \tag{6.3}$$

where $V_{PTV,100\%}$ is the volume of the PTV covered by the isodose prescription, V_{PTV} is the volume of the PTV and $V_{External,100\%}$ is the volume of the body covered by the isodose prescription.

For the OARs, the score was set as:

$$\text{Score}_{OAR} = \frac{D_{P,OAR}}{D_{TC,OAR}} \tag{6.4}$$

where $D_{P,OAR}$ is the OAR planned dose and $D_{TC,OAR}$ is the tolerance dose for each OAR. For each objective, a value of one is expected if the dose for that structure is equal to the respective tolerance value. When a better organ sparing or target coverage is obtained, a score less than one will be obtained.

For this study, all delineated structures were grouped according to their location and clinical importance. Therefore, the PTV was assigned to the PTV group with a relative weight of 40%, the brainstem to the Critical group with a relative weight of 50%, the chiasm, the optical nerves, the

retinas and the lens to the Optics group with a relative weight of 7% and the cochlea and pituitary gland to Other group with a relative weight of 3%. Within each group, the same weight was attributed to all its structures (Table S6.1 in the Supplementary material). A partial group score based on the dose sparing of the structures that belong to that group was also calculated.

SPIDERplan analysis was complemented by the gradient index (GI) proposed by Paddick and Lippitz.[44] This index measures the steepness of the dose gradient outside the PTV providing information about the amount of irradiated healthy tissue. The GI is given by:

$$GI = \frac{V_{External,50\%}}{V_{External,100\%}} \tag{6.5}$$

where $V_{External,50\%}$ corresponds to the volume of healthy tissue covered by the isodose surface corresponding to half of the prescription dose. The lower the GI value the better the plan quality.

6.2.7. Statistical Analysis

Statistical comparisons of the global plan and the group scores were performed with IBM SPSS software, version 25. Statistically significant differences between the plan sets were assessed using a randomized block design ANOVA test and, if applicable, a post-hoc multiple comparison test using the Tukey method. For all statistical tests a level of significance of 5% was considered.

6.3. Results

The SPIDERplan global plan score values of *CLIN*, *BAO*, *VMAT* and *ATO* plans for all meningioma cases are shown in Figure 6.2a1. All plans presented global plan scores well below unity as a result of the high-quality level of the obtained dose distributions. The mean global plan scores over all patients ranged between 0.795 (for *BAO*) and 0.823 (for *VMAT*).

The statistical analysis applied to the global plan scores of the plan sets is summarized in Figure 6.2a2. Statistically significant differences between the global plan scores of the plan sets were found with the randomized block design ANOVA test (p-value = 0). Pairs of plan sets which do not statistically differ from each other were identified by the post-hoc multiple comparison test applied with the Tukey method. Two subsets, grouping the plan sets which did not present statistically significant differences were built. The first subset included the *BAO*, the *ATO* and the *CLIN* sets, meaning that the quality of these plans is statistically equivalent. The second subset grouped the *VMAT, the ATO and* the *CLIN* sets. Statistically significant differences in plan quality were only found between the *BAO* plans and the *VMAT* plans.

The results of the group scores are shown Figure 6.2b for the PTV group and in Figure 6.2c for the Optics group. On average, all sets presented mean group scores below one, corroborating the assessment results of the global plan score. For the PTV group, the best coverage and conformity indexes were achieved by the arc-based plans (*ATO* and *VMAT*), while for the Optics group the best sparing doses were reached by *BAO* and *CLIN* plans.

The evaluation of the steepness of the dose falloff outside the PTV was performed through the determination of GI (Figure 6.3). The arc-based plans *(ATO* and *VMAT)* presented the lower mean values of GI, while the plans optimized with static beams (*CLIN* and *BAO*) presented the highest mean values of GI. For similar levels of PTV coverage and conformity, the static beam plans irradiate, on average, the double volume of healthy tissues receiving doses comprehended between half of the prescription and the prescription dose. Only weak or very weak associations were found between the GI and the prescription dose or the global plan score. This confirms the granted finding

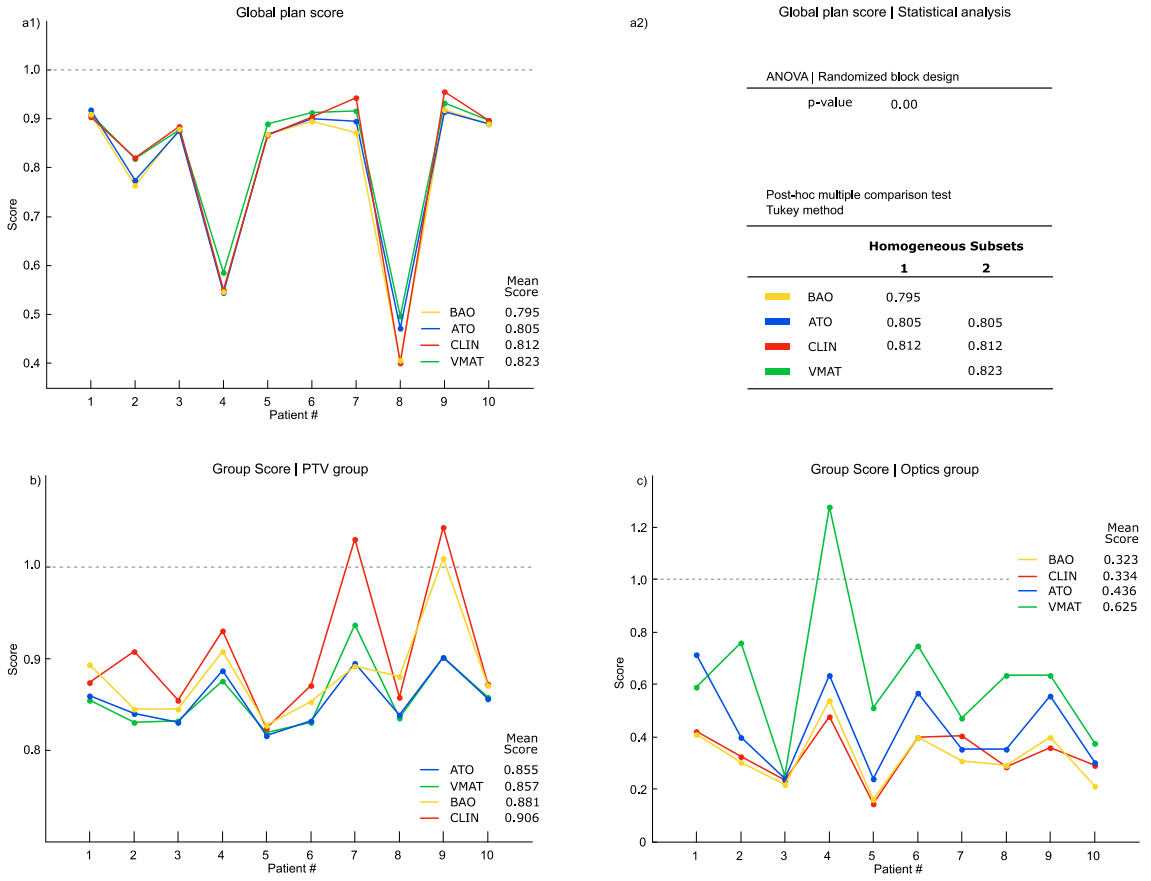

Figure 6.2 - SPIDERplan analysis of *CLIN*, *BAO*, *VMAT* and ATO plans for the 10 meningioma cases. a1) SPIDERplan global plan score. a2) Statistical analysis of the global plan score performed with ANOVA and the multiple comparisons tested using the Tukey method. b) SPIDERplan group scores for the PTV group. c) SPIDERplan group scores for the Optics group.

that the dose falloff outside the target is mainly determined by the irradiation technique and the technical specifications of the linear accelerators (e.g. leaf width of the MLC, speed of MLC, maximum dose rate etc.).

In radiation therapy, the inherent patient-specificity usually requires a careful evaluation of the available treatment options. Among the 10 cases of our sample, one (patient #9), presented notorious differences in SPIDERplan scores (global plan score and group scores) and in the gradient index values. For this patient, the PTV, located next to the chiasm, the left optic nerve and the brainstem, was prescribed with a dose of 60 Gy. In addition to the initial set of plans, an arc trajectory optimized plan based on 9 anchor points (*ATO9*) was also calculated. The evaluation of the quality of the plans calculated for patient #9 is presented in Figure 6.4. The best plans were achieved by techniques with direction/trajectory optimization (*ATO9*, *ATO* and *BAO*) and the worst by the *CLIN* plan. *ATO*-based plans achieved a high level of coverage and conformity of the PTV, while the static beam-based plans presented better performance for the OARs groups. Good results were also achieved by *ATO9.* For the Critical group, *ATO9* presented the best group score and improved the sparing of the Optics group compared to *ATO5*. In fact for this patient, the increase from 5 to 9 anchor points allowed an improvement of 10% in the global score plan. These results highlight not only the potential benefits that may arise from the optimization of direction/trajectory of the beams but also the need of investigating in more detail the influence of the number of anchor points on the quality of the dose distribution.

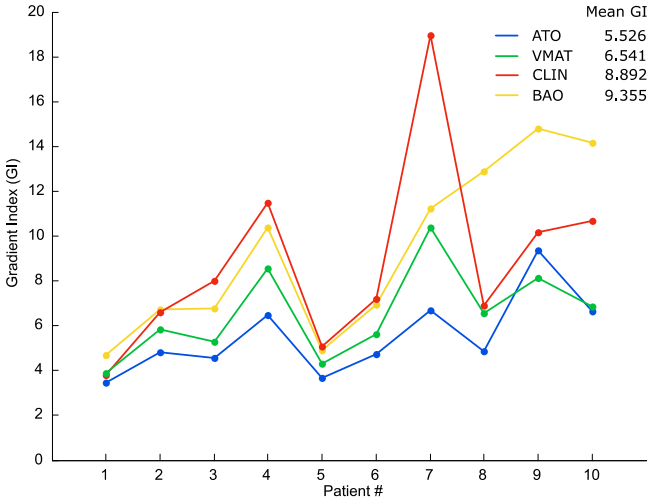

Figure 6.3 - Gradient index computed for the 10 patient cases

6.4. Discussion

In this work, algorithms for non-coplanar beam angular optimization and non-coplanar arc trajectory optimization were applied to 10 meningioma cases. Beam angular optimization was based on a multistart derivative-free optimization framework and guided by SPIDERplan global plan score. These non-coplanar static beam angle optimized incidences were used for non-coplanar VMAT plans generation, defining the anchor points of linear trajectories connecting consecutive points.

The fluence-based beam angle optimization methods are usually guided by the objective function values of the FMO problem, guaranteeing that reliable and high-quality plans are found. However, it cannot be assumed that a plan calculated by these optimization procedures will be

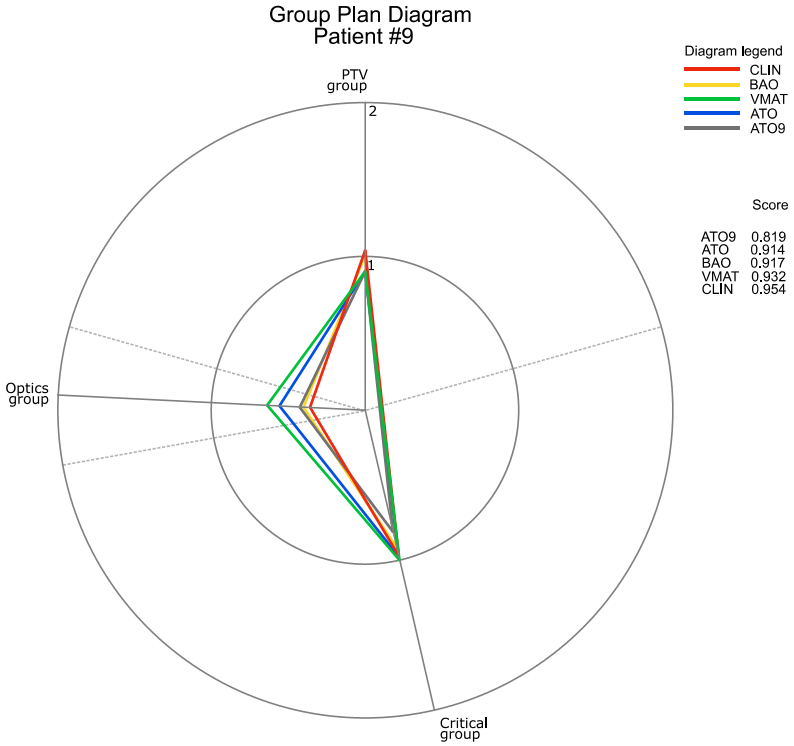

Figure 6.4 - SPIDERplan of patient case number 9

selected or even approved by the radiation oncologist. The integration of SPIDERplan global plan score in the beam angle optimization problem, as a measure of the quality of the beam angle set, intended to generate a plan solution that is optimal from the inverse planning optimization point of view. This solution should also be able to fulfil the clinical aims defined by the radiation oncologist. The referred methodology was firstly applied by Rocha *et al.*[20] to nasopharynx cancer cases, where the quality of the plans generated with and without the global plan score guidance was compared. The plans optimized with the global plan score presented a higher sparing of the OARs for the same PTV coverage than the ones optimized with the common objective function. In the present study, the global plan score was used to guide the optimization of the irradiation directions of the *BAO* and the *ATO* plans for intracranial tumours. Excellent levels of PTV coverage and conformity and high sparing of the OARs were obtained, in line with Rocha *et al.* findings.

The arc trajectory optimization problem is often mathematically described as a problem even more complex than the beam angle optimization problem.[25,29] The two-step method presented in this work intends to take advantage of the experience and the knowledge acquired with the beam angle optimization problem[16,20,36,37,39] contributing to the scientific debate of the arc trajectory optimization problem. The first step of the proposed arc trajectory method consisted in identifying the anchor points of the beam trajectory. The number and location of these points should guarantee a proper sampling of the space search and a smooth beam arc trajectory. A very high number of anchor points could result in complex or irregular trajectories when a connection method is applied while a very low number of anchor points may not be sufficient to define a trajectory likely to generate good quality plans. These initial configuration issues were handled by defining a fixed number of beams to be used in the optimization and by selecting a derivative-free optimization algorithm that considers a continuous search space for optimization. According to our recent work,[37] this algorithm presents a good performance on non-coplanar optimization geometries and a good beam coverage of the space search.

The number of beams defined for the non-coplanar beam angle optimization was based on clinical experience. For the ten meningioma cases considered for this study, the planners need, on average, 5 IMRT beams to achieve a satisfactory plan. In the second step of the arc trajectory algorithm, where the shape of the trajectory of the arc was defined, the anchor points were connected through linear paths. The rationale behind this choice was a literal interpretation of the term ‘anchor point’ and an option for simplifying the optimization algorithm. The irradiation directions that composed the linear trajectory between the anchor points were not fluence-based optimized. However, this drawback was overcome by using an efficient inverse planning optimizer as the Erasmus-iCycle multicriteria engine. The connection order of the linear paths was established by following the pre-condition that the gantry movement was continuous, inversions were not allowed and it was not possible to exceed a 360º arc. It was assumed that the gantry and the couch speeds may be different, and that the inertia associated with the change of velocity or direction of the couch is much lower than the inertia of the gantry. The implementation of these trajectory configuration options was only possible due to the regular and well disperse beam angle distribution of the anchor points in the search space that resulted from the non-coplanar beam angle optimization.

A VMAT optimization module was not available in Erasmus-iCycle at the time the arc trajectory optimization was performed. The continuous motion of the gantry and the MLC in the VMAT delivery were approximated by using 21 equidistant static beams distributed over the trajectory as was demonstrated by Bortfeld.[41] This approach implies that the beam modulation that occurred during each arc sector of ~17º is replaced by an intensity-modulated field placed in the centre of that arc sub-sector. The conversion of these theoretical fluence maps into deliverable MLC segments all over the arc trajectory will probably degrade the dose distribution quality. In the future, improvements to this arc trajectory optimization method should include endorsing the process of fully fluence-based methods like the one proposed by Rocha et al. in a preliminary study[36] and by using a VMAT optimization algorithm to calculate an improved dose distribution.

The plans optimized with non-coplanar irradiation directions, *CLIN, BAO and ATO*, presented higher dose distribution quality than the plans based on coplanar geometries (*VMAT*). The advantages of non-coplanar beams geometries over coplanar ones for brain cases were previously reported.[3,8,9] However, consensus about the best radiation therapy technique to treat these tumour lesions was not reached. Fogliata *et al.*[45] reported equivalent quality between non-coplanar IMRT plans, coplanar VMAT plans and helical Tomotherapy plans. Conversely, Panet-Raymond *et al.*[46] achieved equivalent PTV coverage with non-coplanar IMRT and coplanar VMAT plans, but higher OAR sparing with the former set of plans.

For non-coplanar geometries, all plans achieved high-quality (average SPIDERplan global scores were well below one). Although the best performance was obtained by *BAO,* no statistically significant score differences were found, on average, between *BAO*, *ATO* and *CLIN* plans. The *BAO* plans were most effective in sparing the OARs, while the *ATO* plans enabled higher PTV coverage and conformity. Furthermore, a steeper dose gradient outside the PTV was also possible with the *ATO* plans, due to the higher number of irradiation directions available with this technique. Previous studies with non-coplanar arc trajectory optimization algorithms applied to intracranial tumours have been published. Papp *et al.*[25] applied an iterative beam angle optimization method to define the anchor points and solved the travelling salesman problem to define the remaining trajectory of the arc. Langhans *et al*.[29] used an iterative method based on a 4π solution to find the feasible anchor points and defined the arc trajectory based on geometrical scoring evaluation of the available beam directions. Both works reported improved plan quality when non-coplanar VMAT plans were compared with non-coplanar IMRT plans (equivalent to *CLIN* plans). Although these conclusions are in line with the results of this study, our *ATO* plans have not brought any additional improvements to the quality of the dose distribution, when compared with *BAO* plans. This finding may be related with the high performance of the non-coplanar beam angle optimization algorithm and the multicriterial IMRT optimization engine, that are able to generate high-quality plan solutions with a low number of static beam directions (5 non-coplanar beams). Furthermore, the number of anchor points selected to build the beam trajectory may not be optimal. As was shown in the patient-specific analysis of patient #9, for some more complex situations, a higher number of anchor points can be advantageous. The ideal number of anchor points and its specific case applications are issues of interest for future research but are out of the scope of the present work.

The high quality of *CLIN* plans, and the fact that no statistically significant differences between the score of the *BAO* and the *ATO* plans were obtained, must be highlighted. The quality of *CLIN* plans resulted from the high skills and accumulated clinical experience of the planners. In the clinical routine, the plans for patients with intracranial tumours are usually manually optimized with 4-6 non-coplanar beams in a very time-consuming process, attempting to spare as much as possible the OARs and to fulfil the PTV coverage requirements. This planning strategy is clearly shown by the group score results that present similar behaviour to the *BAO* plans, i.e. better score in the OARs groups than in the PTV group. Even so, the inclusion of beam angle optimization into inverse treatment planning with minimal intervention from the user and with the guarantee of consistent generation of high-quality plans (such as the *BAO* and the *ATO* plans) should motivate a strong commitment towards introducing automated planning tools in the clinical practice.

6.5. Conclusions

In this work, the irradiation directions optimization problem was addressed for the intracranial pathology using ten meningioma cases. A beam angle optimization algorithm guided by a global score based on the dosimetric parameters of the structures and on the radiation oncologist clinical preferences and a new two-step arc trajectory optimization approach using non-coplanar beam angle optimized static beams as anchoring points of the arc path were used. A beam geometry with 5 non-coplanar incidences was chosen to run beam angle and arc trajectory optimization problem. *BAO*

plans presented, on average, a slightly better plan quality than *ATO* and *CLIN* plans, even if no significant statistical score differences were found between them. Moreover, the ATO plans assured a more efficient coverage and conformality of the PTV, while a higher sparing of the OARs was achieved by the *BAO* plans. This global analysis does not dismiss a patient specific analysis, since strong benefits may be obtained with 4π directions optimization for specific-patients.

6.6. Acknowledgements

In this work, a beam geometry with 5 non-coplanar incidences was chosen to run the beam angle and the arc trajectory optimization problem for the intracranial pathology using ten meningioma cases. *BAO* plans presented, on average, a slightly better plan quality than *ATO* and *CLIN* plans, even if no significant statistical score differences were found between them. The *ATO* plans assured a more efficient coverage and conformity of the PTV, while a higher sparing of the OARs was achieved by the *BAO* plans. This global analysis does not dismiss a patient specific analysis, where strong benefits may be obtained with 4π directions optimization in specific patients.

6.7. References

1. ICRU. International Commission on radiation units and measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Geneva; 2010. J ICRU 10(1):1–106.
2. Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *Br J Radiol*. 2011;84(1007):967-996.
3. Das SK, Marks LB. Selection of coplanar or noncoplanar beams using three-dimensional optimization based on maximum beam separation and minimized nontarget irradiation. *Int J Radiat Oncol Biol Phys*. 1997;38(3):643–655.
4. Podgorsak E, Olivier A, Pla M, Lefebvre P, Hazel J. Dynamic stereotactic radiosurgery. *Int J Radiat Oncology Biol Phys*. 1988;14(1):115-126.
5. Bortfeld T, Schlegel W. Optimization of beam orientations in radiation therapy: some theoretical considerations. *Phys Med Biol*. 1993;38(2):291–304.
6. Ehrgott M, Holder A, Reese J. Beam selection in radiotherapy design. *Linear Algebra Appl*. 2008;428(5):1272–1312.
7. Craft D. Local beam angle optimization with linear programming and gradient search. *Phys Med Biol*. 2007;52(7):127–135.
8. Bangert M, Ziegenhein P, Oelfke U. Comparison of beam angle selection strategies for intracranial IMRT. *Med Phys*. 2013;40(1):011716.
9. Popple R, Brezovich I, Fiveash J. Beam geometry selection using sequential beam addition. *Med Phys*. 2014;10(5):051713.
10. Bangert M, Unkelbach J. Accelerated iterative beam angle selection in IMRT. *Med Phys*. 2016;43(3):1073–1082.
11. Pugachev A, Li J, Boyer A et al. Role of beam orientation optimization in intensity modulated radiation therapy. *Int J Radiat Oncol Biol Phys*. 2001;50(2):551-560.
12. Potrebko P, McCurdy B, Butler J, El-Gubtan A. Improving intensity-modulated radiation therapy using the anatomic beam orientation optimization algorithm. *Med. Phys*. 2008;35(5):2170-2179.
13. Bangert M, Oelfke U. Spherical cluster analysis for beam angle optimization in intensity-modulated radiation therapy treatment planning. *Phys Med Biol*. 2010;55(19):6023-6037.
14. Breedveld S, Storchi P, Voet P, Heijmen B. iCycle: integrated, multi-criterial beam angle and profile optimization for generation of coplanar and non-coplanar IMRT plans. *Med Phys*. 2012;39(2):951-963.
15. Rocha H, Dias J, Ferreira B, Lopes MC. Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method. *Phys Med Biol*. 2013;58(9):2939-2953.
16. Rocha H, Dias J, Ventura T, Ferreira B, Lopes MC. A derivative-free multistart framework for an automated noncoplanar beam angle optimization in IMRT. *Med Phys*. 2016;43(10):5514-5526.
17. Dong P, Lee P, Ruan D et al. 4π Noncoplanar Stereotatic Body Radiation Therapy for Centrally Located or Larger Lung Tumours. *Int J Radiat Oncol Biol Phys*. 2013;86(3):407-413.

18. Dong P, Lee P, Ruan D et al. 4π Non-Coplanar Liver SBRT: A Novel Delivery Technique. *Int J Radiat Oncol Biol Phys*. 2013;85(5):1360-1366.
19. Liu H, Dong P, Xing L. A new sparse optimization scheme for simultaneous beam angle and fluence map optimization in radiotherapy planning. *Phys Med Biol*. 2017;62(16):6428-6445.
20. Rocha H, Dias JM, Ventura T, Ferreira BC, Lopes MC. Beam angle optimization in IMRT: are we really optimizing what matters? *Intl Trans in Op Res*. 2019;26(3):908-928.
21. Krayenbuehl J, Davis JB, Ciernik I. Dynamic intensity-modulated non-coplanar arc radiotherapy (INCA) for head and neck cancer. *Radiother Oncol*. 2006;81(2):151-157.
22. Yang Y, Zhang P, Happersett L et al. Choreographing couch and collimator in volumetric modulated arc therapy. *Int J Radiat Oncol Biol Phys*. 2011;80(4):1238-1247.
23. Audet C, Poffenbarger B, Chang P et al. Evaluation of volumetric modulated arc therapy for cranial radiosurgery using multiple noncoplanar arcs. *Med Phys*. 2011;38(11):5863-5872.
24. Smyth G, Bamber JC, Evans PM, Bedford JL. Trajectory optimization for dynamic couch rotation during volumetric modulated arc radiotherapy. *Phys Med Biol*. 2013;58(22):8163-8177.
25. Papp D, Bortfeld T, Unkelbach J. A modular approach to intensity-modulated arc therapy optimization with noncoplanar trajectories. *Phys Med Biol*. 2015;60(13):5179-5198.
26. MacDonald RL, Thomas CG. Dynamic trajectory-based couch motion for improvement of radiation therapy trajectories in cranial SRT. *Med Phys.* 2015;42(5):2317-2325.
27. Smyth G, Evans PM, Bamber JC et al. Non-coplanar trajectories to improve organ at risk sparing in volumetric modulated arc therapy for primary brain tumors. *Radiother Oncol*. 2016;121(1):124-131.
28. Dong P, Liu H, Xing L. Monte Carlo tree search-based non-coplanar trajectory design for station parameter optimized radiation therapy (SPORT). *Phys Med Biol.* 2018;63(13):135014.
29. Langhans M, Unkelbach J, Bortfeld T, Craft D. Optimizing highly noncoplanar VMAT trajectories: the NoVo method. *Phys Med Biol*. 2018;63(2):025023.
30. Lyu Q, Yu VY, Ruan D, Neph R, O'Connor D, Sheng K. A novel optimization framework for VMAT with dynamic gantry couch rotation. *Phys Med Biol*. 2018;63(12):125013.
31. Shaitelman SF, Kim LH, Yan D, Martinez AA, Vicini FA, Grills IS. Continuous arc rotation of the couch therapy for the delivery of accelerated partial breast irradiation: a treatment planning analysis. *Int J Radiat Oncol Biol Phys*. 2011;80(3):771-778.
32. Popescu CC, Beckham WA, Patenaude VV, Olivotto IA, Vlachaki MT. Simultaneous couch and gantry dynamic arc rotation (CGDarc) in the treatment of breast cancer with accelerated partial breast irradiation (APBI): a feasibility study. *J Appl Clin Med Phys*. 2013;14(1):161–175.
33. Fahimian B, Yu V, Horst K, Xing L, Hristov D. Trajectory modulated prone breast irradiation: a LINAC-based technique combining intensity modulated delivery and motion of the couch. *Radiother Oncol*. 2013;109(3):475-181.
34. Wild E, Bangert M, Nill S, Oelfke U. Noncoplanar VMAT for nasopharyngeal tumors: Plan quality versus treatment time. *Med Phys*. 2015;42(5):2157-2168.
35. Fix MK, Frei D, Volken W et al. Part 1: Optimization and evaluation of dynamic trajectory radiotherapy. *Med Phys*. 2018;45(9):4201-4212.
36. Rocha H, Dias J, Ventura T, Ferreira B, Lopes MC. An Optimization Approach for Noncoplanar Intensity-Modulated Arc Therapy Trajectories. In: Misra S, Gervasi O, Murgante B et al. (eds). Computational Science and Its Applications - ICCSA 2019. ICCSA 2019. Lecture Notes in Computer Science. Cham, Springer. 2019;11621:199-214.
37. Ventura T, Rocha H, Ferreira BC, Dias J, Lopes MC. Comparison of two beam angular optimization algorithms guided by automated multicriterial IMRT. *Phys Medica*. 2019;64:210-221.
38. Ventura T, Lopes MC, Ferreira BC, Khouri L. SPIDERplan: a tool to support decision making in radiation therapy treatment plan assessment. *Rep Pract Oncol Radiother*. 2016;21(6):508–16.
39. Ventura T, Lopes MC, Rocha H, Ferreira BC, Dias J. Advantage of Beam Angle Optimization in Head-and-Neck IMRT: Patient Specific Analysis. In: Henriques J, Neves N, de Carvalho P. (eds). XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON 2019. MEDICON 2019. IFMBE Proceedings. Cham, Springer. 2020;76;1256-1263.
40. Breedveld S, Storchi P, Keizer M, Heemik A, Heijmen B. A novel approach to multi-criteria inverse planning for IMRT. *Med Phys.* 2007;52(20):6339-6353.
41. Bortfeld T. The number of beams in IMRT: Theoretical investigations and implications for single-arc IMRT. *Phys Med Biol*. 2010;55(1):83-97.
42. Custódio AL, Rocha H, Vicente LN. Incorporating minimum Frobenius norm models in direct search. *Comput Optim Appl*. 2010;46:265-78.

43. Paddick I. A simple scoring ratio index the conformity of radiosurgical treatment plans. *J Neurosurg*. 2000;93:219–222.
44. Paddick I, Lippitz B. A simple dose gradient measurement tool to complement the conformity index. *J Neurosurg*. 2006;105:194-201.
45. Fogliata A, Clivio A, Nicolini G, Vanetti E, Cozzi L. Intensity modulation with photons for benign intracranial tumours: A planning comparison for volumetric single arc, helic arc and fixed gantry techniques. *Radiother Oncol.* 2008;89(3):254-262.
46. Panet-Raymond V, Ansbacher W, Zagvorodni S et al. Coplanar versus noncoplanar intensity modulated (IMRT) and volumetric-modulated arc therapy (VMAT) treatment planning for fronto-temporal high-grade glioma. *J Appl Clin Med Phys*. 2012;13(4):44-53.

6.8. Supplementary material

Table S6.1 - Groups, structures and weights considered for SPIDERplan processing.

Groups		**Structures**		
Name	*Group Weight*	*Name*	*Structure Weight*	*Tolerance criteria*
PTV	40%	PTV	100%	$V_{100\%} \geq 0.95$ COIN > 0.6
Critical	50%	Brainstem	100%	$D_{max} \leq 54Gy$
Optics	7%	Chiasm	14.3%	$D_{max} \leq 55Gy$
		Left optical nerve		$D_{max} \leq 55Gy$
		Right optical nerve		$D_{max} \leq 55Gy$
		Left retina		$D_{max} \leq 45Gy$
		Right retina		$D_{max} \leq 45Gy$
		Left lens		$D_{max} \leq 12Gy$
		Right lens		$D_{max} \leq 12Gy$
Other	3%	Right cochlea	33.3%	$D_{mean} \leq 45Gy$
		Left cochlea		$D_{mean} \leq 45Gy$
		Pituitary gland		$D_{max} \leq 60Gy$

D_{max} – maximum dose, D_{mean} – mean dose

Table S6.2 - General wish-list defined for meningioma cases.

Level	**Priority**	**Structure**	**Type**	**Goal**	**Sufficient**	**Parameters**
Constraints						
		PTV	maximum	$D_{p,107\%}$		
		Brainstem	maximum	54 Gy		
		Retinas	maximum	45 Gy		
		Optical nerves	maximum	55		
		Chiasm	maximum	55		
		Unspecified Tissue	maximum	D_p		
Objectives						
1	1	PTV	LTCP	1	0.5	D_p/0.95 Gy, α=0.75
	2	Brainstem	maximum	f_{10}DVH		
	3	Chiasm	maximum	f_{10}DVH		
	4	Optical nerves	maximum	f_{10}DVH		
	5	Retinas	maximum	f_{10}DVH		
	6	Lens	gEUD	f_{10}DVH		a=15
	7	Ring1 PTV_{Dp}	maximum	$D_{p,85\%}$		
	8	Ring2 PTV_{Dp}	maximum	$D_{p,70\%}$		
	9	Ring3 PTV_{Dp}	maximum	$D_{p,55\%}$		
2	10	Brainstem	mean	f_{10}DVH		
	11	Chiasm	mean	f_{10}DVH		
	12	Optical nerves	mean	f_{10}DVH		
	13	Retinas	mean	f_{10}DVH		
	14	Lens	gEUD	f_{10}DVH		a=6
	15	External ring	maximum	40%×D_p		
	16	Unspecified Tissue	mean	5 Gy		

Ring1 PTV - ring of 5 mm of thickness at 10 mm distance from PTV, Ring2 PTV - ring of 10 mm of thickness at 15 mm distance from PTV, Ring3 PTV - ring of 20 mm of thickness at 25 mm distance from PTV, External ring - ring of 20 mm thickness next to patient outer contour

Algorithm of Beam Angle Optimization

Initialization

- Set $k:=0$;
- Choose initial points, one for each sub-region, x_i^0, i=1,…,N;
- Compute $GPS(x_i^0)$, i=1,…,N in parallel;
- Set $x_i^*:=x_i^0$, i=1,…,N and $GPS_i^*=GPS(x_i^0)$, i=1,…,N;
- Set $Active_i:=1$, i=1,…,N, i.e. all sub-regions initially have active local searches;
- Choose $\alpha_i^0>0$, i=1,…,N and α_{min};

Iteration

1. Use a derivative-free algorithm to locally explore the sub-regions with active local searches;
2. For sub-regions i with active local search do

   If $GPS(x_i^k)<GPS(x_i^*)$ then

   If x_i^k is in sub-region i then

   $x_i^*:=x_i^k$;

   $GPS_i^*=GPS(x_i^k)$;

   Else

   $Active_i:=0$;

   Determine sub-region $j\neq i$ where x_i^k is;

   If $GPS(x_i^k)<GPS(x_j^*)$ then

   $x_j^*:=x_i^k$;

   $GPS_j^*=GPS(x_i^k)$;

   $Active_j:=1$;

   Else

   $\alpha_i^{k+1}:=\frac{\alpha_i^k}{2}$;

   If $\alpha_i^{k+1}<\alpha_{min}$ then

   $Active_i:=0$;

3. If there exist active sub-regions go to first step and set $k:=k+1$.

(GPS – Global Plan Score; α_{min} – threshold value)

chapter 7

Conclusions

The present work had as main purpose to contribute to the implementation of automated plan optimization methods supported by tools for the assessment of the treatment plans quality. For the assessment of the quality of the treatment plans a new graphical evaluation tool, named SPIDERplan, was developed. It takes simultaneously into account the dosimetric properties of the plan and the clinical aims of the radiation oncologist. SPIDERplan combines a weighted score definition and a graphical representation, allowing the decision-maker in a glance to select the best plan out of a set of plans. This tool is independent of the optimization algorithm, the treatment technique or the treatment planning system. Still, a clinical validation per pathology is required before its introduction in clinical practice.

SPIDERplan was successfully validated for nasal cavity and nasopharynx carcinomas. For the first pathology, a combinatorial weight approach was used. For the nasopharynx cases, different configurations of structures and weights were tested. SPIDERplan responses were compared against the choices of three radiation oncologists. For all configuration methods, SPIDERplan response was successfully linked with the ROs plan selections. Moreover, it was demonstrated that for the nasopharynx carcinomas, SPIDERplan evaluation presented an equivalent variability to that of the ROs and that any of the configuration options could be used without loss of accuracy.

SPIDERplan was the base for the comparison of nasopharynx plans optimized with coplanar and non-coplanar beam angle optimization in Erasmus-iCycle. In this study two beam angle optimization algorithms were compared: one based on a combinatorial iterative approach and the other on a continuous exploration of the search space. Both algorithms presented on average IMRT plans with high quality dose distributions. This fact probably justified that no statistically significant differences were found between the quality of the plans produced by either. Specific patient cases analyses were performed to highlight the improvement of beam angle optimization over the equidistant beam angle solution, the importance of the non-coplanar beam geometries in the beam angle optimization process and the usefulness of beam angle optimization in the generation of plans with a smaller number of beams.

SPIDERplan applicability was raised into another level in the non-coplanar beam angle optimization study carried out for intracranial tumours. Taking advantage from being able to choose a treatment plan according to the clinical preferences of the radiation oncologist, SPIDERplan was successfully incorporated in the optimization process for driving the beam angle optimization for meningioma cases. An arc trajectory algorithm based on the linear connection of the static beam anchor points defined by the beam angle optimization algorithms was proposed. For this case, a beam angle optimization algorithm able to spread the anchor points over the search space proved to ensure a smooth and feasible beam trajectory during treatment. Compared with plans with beam directions manually optimized by the planners, on average, no statistically significant improvements were obtained against the plans with beam angle and arc trajectory optimizations. Again, only a case-by-case analysis demonstrated the potential benefits of directions/trajectories optimization. The inclusion of beam angle or arc trajectory optimization algorithms in a systematic and automatic way ensure a consistent generation of high-quality plans that should motivate the adoption of automated planning procedures with minimal dependence from the planner interaction.

Fully personalized radiation therapy through on-line adaptive radiation therapy and machine learning techniques are totally dependent of fast plan automation techniques. The workflow and the professional roles in radiation therapy is changing to provide treatments with high and consistent quality adjusted to the patient daily condition. SPIDERplan can take place as an independent tool or part of an optimization framework for the quality assessment of the treatment plans or in the optimization guidance of inverse treatment planning algorithms.

appendixes

appendix 1

Evaluation of two arc trajectory optimization algorithms for intracranial tumours VMAT planning

abstract submitted to ESTRO 2020

T. Ventura[1,2,3], H. Rocha[3,4], B.C. Ferreira[3,5,6], J. Dias[3,4], M.C. Lopes[1,2,3]

[1]Physics Department, University of Aveiro, Aveiro, Portugal
[2]Medical Physics Department, IPOCFG, EPE, Coimbra, Portugal
[3]Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal
[4]Economy Faculty of University of Coimbra and Centre for Business and Economics Research, Coimbra, Portugal
[5]School Health Polytechnic of Porto, Porto, Portugal
[6]I3N, Physics Department, University of Aveiro, Aveiro, Portugal

A1.1 Purpose

To assess and compare plan quality and deliverability of two arc trajectory optimization methods for VMAT planning of intracranial tumours.

A1.2 Materials and Methods

For this study, four cases were selected among a set of meningioma cases previously studied. The arc trajectory optimization followed a two-step approach, where dosimetric and geometric features were combined. In the first step, non-coplanar beam angle optimization was used to search for feasible anchor points of the arc trajectory. In the second step, the anchor points were linearly connected. Two beam angle optimization algorithms (*B* and *i*) were used to optimize 5 non-coplanar beams. Algorithm *B* belongs to the continuous search space class and uses a pattern search method to find the best possible beam angle ensemble. Algorithm *i* is based in a discrete iterative combinatorial approach. For the arc trajectory definition, it was considered that the gantry and the couch could move simultaneously and with different speeds. All fluence map optimizations were performed using an automated multicriterial optimization calculation engine (Erasmus-iCycle). Plan quality was evaluated through SPIDERplan score and the Paddick gradient index. The arc trajectories were qualitatively evaluated.

A1.3 Results

Both algorithms produced plans with high dose distribution quality. Moderate quality plan enhancements were presented by algorithm *i,* due to the higher level of sparing of the optical structures. Nevertheless, the gradient index was lower for algorithm *B*, which means it presents steeper dose falloffs outside the PTV and a lower dose received by the normal tissues.

The two algorithms showed different arc trajectories (Figure A1.1). Algorithm *B* presented arc trajectories with higher gantry rotation amplitudes, covering almost all angles, and smooth couch displacements. In algorithm *i,* for patients #1 and #3, similar shape trajectories to algorithm *B* were generated. For patient #2 and #4, sharp couch rotations associated to reduced gantry amplitudes created winding arc trajectories that can compromise the treatment delivery and be very uncomfortable for the patient.

A1.4 Conclusion

The quality of the arc-based plans generated with algorithms *B* and *i* was very good with no significant differences between them. For algorithm *B*, the generated arc trajectories presented a regular pattern whereas for algorithm *i* the arc trajectory generation is irregular. In some cases, this should be well-considered before treatment plan approval.

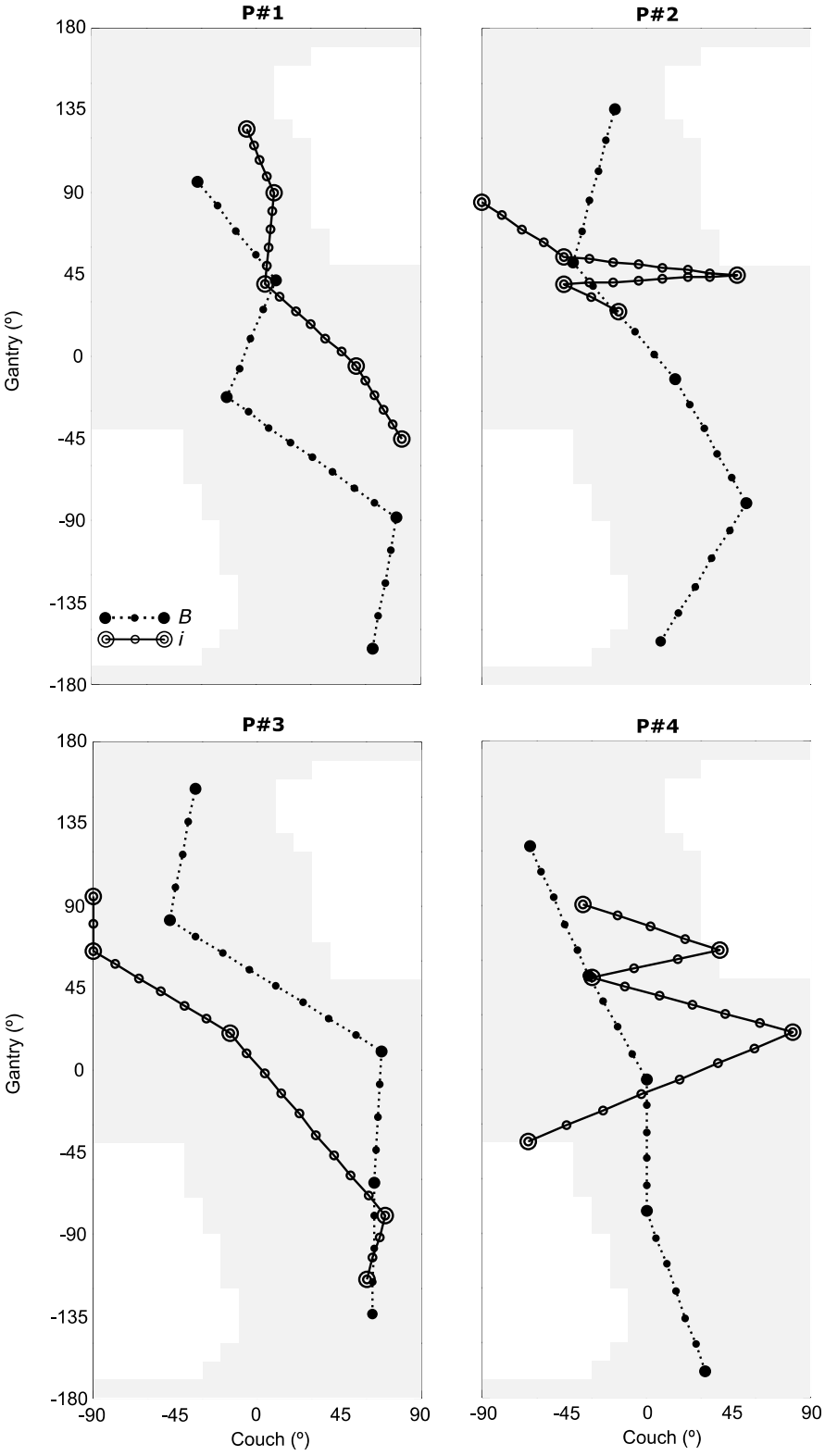

Figure A1.1 - Arc trajectories of algorithms *B* and *i* for the selected patients.