

Alejandro García Romero

Nuevas métricas para la verificación de tratamientos de IMRT

Departamento
Pediatría, Radiología y Medicina Física

Director/es
Hernández Vitoria, Araceli
Millán Cebrián, Pilar Esther

<http://zaguan.unizar.es/collection/Tesis>

© Universidad de Zaragoza
Servicio de Publicaciones

ISSN 2254-7606



Universidad
Zaragoza

Tesis Doctoral

NUEVAS MÉTRICAS PARA LA VERIFICACIÓN DE TRATAMIENTOS DE IMRT

Autor

Alejandro García Romero

Director/es

Hernández Vitoria, Araceli
Millán Cebrián, Pilar Esther

UNIVERSIDAD DE ZARAGOZA

Pediatría, Radiología y Medicina Física

2015



Universidad
Zaragoza

FACULTAD DE MEDICINA

DEPARTAMENTO DE PEDIATRÍA, RADIOLOGÍA Y MEDICINA FÍSICA

Tesis doctoral

**Nuevas métricas para la verificación de tratamientos
de IMRT**

Alejandro García Romero

Zaragoza, 2015

Doña Araceli Hernández Vitoria, Profesora Titular del Departamento de Pediatría, Radiología y Medicina Física de la Universidad de Zaragoza, y Dña P. Esther Millán Cebrián, Profesora Asociada en Ciencias de la Salud del mismo Departamento

Certifican:

Que D. Alejandro García Romero, licenciado en Ciencias Físicas, ha realizado bajo su dirección y en el Departamento de Pediatría, Radiología y Medicina Física de la Universidad de Zaragoza el trabajo "Nuevas métricas para la verificación de tratamientos de IMRT", que se recoge en esta memoria y constituye su Tesis Doctoral para optar al grado de Doctor por la Universidad de Zaragoza.

Y para que así conste, de acuerdo con la legislación vigente, firmamos el presente certificado.

Zaragoza, 15 de Octubre de 2015

Profa. Araceli Hernández Vitoria

Profa. P. Esther Millán Cebrián



Departamento de
Pediatría, Radiología
y Medicina Física
Universidad Zaragoza

Que por tu ventana sale, es tanta la claridad

Que por tu ventana sale, que dice la vecindad

Ya está la luna en la calle

Popular

Agradecimientos

A mis directoras, Araceli y Esther, y su esforzada revisión y empuje, su ejemplo y su valentía.

A todos mis compañeros que son y han sido, Pedro, Mari Ángeles, Pablo, Elena, Luisa, Aurora, Rosa, Tina, Mamen, y una larga lista de residentes, secretarias y técnicos que pasaron por el Servicio de Física del HCU de Zaragoza. También a un residente de disposición transitoria, D. Javier Gracia, él sabe por qué. Este agradecimiento se hace extensivo a los médicos, técnicos, enfermeras y secretarias del Servicio de Oncología Radioterápica. De todos aprendí y de todos tengo algo, porque sin permiso tomé prestado un montón de cualidades que me han ayudado y moldeado. Mis defectos ya los traía de serie.

En especial, este trabajo tiene parte de Sheila, Vero y Sonia, ellas lo saben, y espero que en el futuro muchas veces sigamos colaborando.

A Óscar Ripol, porque nos une un EPID Perkin-Elmer, el tozal de la cobeta y sus brillantes ideas.

A Victor Laliena, porque es un espejo en el que mirarse, porque tuyas son algunas gráficas, y porque tuvo la iniciativa que permitió desarrollar SMLC.

A Lola Lardies, por la estrecha colaboración de muchos años, y a Miguel Canellas, porque tuyos son los kernels sagrados y tuya es la paciencia que siempre ha tenido conmigo.

A Santos y a Unai, porque el Oncor no se paró (del todo).

En el plano espiritual, muchas gracias a Anders Ahnesjö, padre y madre del cono colapsado, y a al señor Rogers y todos aquellos que desarrollaron BEAMnrc, tal vez nunca lo sepan, pero su trabajo iluminó mi camino.

A mi madre, mi esforzada y querida madre, la mejor madre que he tenido, porque ha habido otras también muy buenas. A nadie puede hacerle más feliz ver esta tesis que a mi madre, y quizá nadie tenga más que ver con ella.

A Carmen, por todo lo que dejé de jugar contigo estas tardes, seguro que me lo perdonas.

Y la última, mi señora de las alturas, la que ha tenido que sacrificarse tanto, la mejor persona que conozco. Es como si la firmara ella, y así constará en nuestra ficha TESEO particular.

Glosario de términos y acrónimos

3DCRT: Radioterapia conformada 3D

ALE: Acelerador lineal de electrones

ALPO: Promedio de la apertura entre pares de láminas

ASCII: Código americano estándar para el intercambio de información. Del inglés “American Standard Code for Information Interchange”.

AUC: Área bajo la curva ROC

BEAMnrc: Código Monte Carlo para la simulación de aceleradores lineales de electrones procedente del National Research Council de Canadá

BEV: Vista desde el punto de vista del haz (del inglés “Beam’s Eye View”)

CBCT: Tomografía computarizada de haz cónico, del inglés “Cone Beam Computed Tomography”

CMPC: Algoritmo de cálculo en el sistema de medida Compass a partir del plan transferido desde el sistema de planificación y la estimación de la fluencia esperada

CMPM: Algoritmo de cálculo en el sistema de medida Compass a partir de medidas experimentales de la fluencia

CTKMAX: Número máximo de cortes con los que trabaja el programa CTCREATE

Dosis: Se entenderá en el texto la palabra dosis como la magnitud dosis absorbida (en agua o en el medio según se especifique)

CRI_D: Cristalino derecho

CRI_I: Cristalino izquierdo

CTV: Volumen objetivo clínico (del inglés “Clinical Target Volume”)

Curva ROC: Curva característica operativa del receptor

Dosis: Se entenderá en el texto la palabra dosis como la magnitud dosis absorbida (en agua o en medio según se especifique)

DTA: Distancia al punto más cercano con la misma dosis (del inglés "Distance to agreement")

DM: Dosis media que recibe un determinado volumen

DX: Dosis que recibe el X por ciento del volumen

DXcc: Dosis que reciben X centímetros cúbicos del volumen

EPID: Dispositivo electrónico de imagen portal (del inglés "Electronic Portal Imaging Device")

FI: Función incumplimiento

FIP: Función incumplimiento ponderado

FGO: Función global de diferencias en dosis por objetivos

FG: Función global de diferencias en dosis

FWHM: Anchura a mitad de altura (del inglés "Full Width at Half Maximum")

gEUD: Dosis equivalente uniforme generalizada (del inglés "Generalized Equivalent Uniform Dose")

HCUZ: Hospital Clínico Universitario "Lozano Blesa" de Zaragoza

HDV: Histograma dosis-volumen

IMRT: Radioterapia de intensidad modulada (del inglés "Intensity Modulated Radiation Therapy")

IMRT estática ("step and shoot"): Técnica de tratamiento de IMRT que permite modular la intensidad cambiando la forma del campo para una misma orientación de brazo de manera sucesiva. El acelerador sólo irradia cuando las láminas del colimador están en situación estática. Cada una de las posiciones estáticas se denomina "segmento".

IR: Índice de robustez. Valor entre 0 y 1 que resume desde el punto de vista radiobiológico, dado un plan calculado o reconstruido con un algoritmo determinado, la distancia global al punto de máxima pendiente de las curvas de respuesta.

MC: Algoritmo Monte Carlo para el cálculo de dosis absorbida en pacientes.

MCesc: Algoritmo MC escalado de manera proporcional para obtener unos histogramas dosis volumen que incluyan el resultado de la medida con el sistema Compass.

MLC: Colimador multilamina (Optifocus en el presente trabajo)

MED: Médula

NSEG: Número de segmentos del plan de IMRT

NTCP: Probabilidad de complicación en tejido sano (del inglés “Normal Tissue Complication Probability”)

OAR: Órgano de riesgo o estructura crítica (del inglés “Organ at Risk”)

PAR_D: Parótida derecha

PAR_I: Parótida izquierda

PC: Punto de corte de la curva ROC

PDP: Porcentaje de dosis en profundidad

PPBC: Tanto por ciento del promedio de picos proveniente del test bandas contiguas del colimador MLC.

PTV: Volumen blanco de planificación (del inglés “Planning Target Volume”)

PTVp: PTV principal del plan de tratamiento

PTVs: PTV secundario del plan de tratamiento

RDR: Radiografía digital reconstruida

REC: Recto

ROC: Curva de sensibilidad frente a 1 – especificidad para un sistema clasificador de casos (del inglés “Receiver Operating Characteristic”)

SCC: Algoritmo de superposición de cono colapsado de PCRT3D

SMLC: Módulo de BEAMnrc que describe el colimador Optifocus de Siemens.

SPB: Algoritmo de cálculo de dosis basado en convolución con haces de tipo pincel (del inglés “Single Pencil Beam”)

SPSS: Programa de análisis estadístico (del inglés “Software Package for Statistical Analysis”)

Nuevas métricas de verificación IMRT

SPT: Sistema de planificación de tratamientos

SRV: Sistema de registro y verificación

TC: Tomografía computarizada

TCP: Probabilidad de control tumoral (del inglés "Tumor Control Probability")

TERMA: Energía total liberada por unidad de masa debida a fotones primarios

TPR: Razón tejido-maniquí

UM: Unidades de monitor

VARMLC: Módulo de BEAMnrc que describe un colimador genérico tipo VARIAN.

VEJ: Vejiga

VOP_D: Nervio óptico derecho

VOP_I: Nervio óptico izquierdo

VX: Volumen que recibe el X por ciento de la dosis

Resumen

Un paso previo a aplicar un tratamiento de IMRT a un paciente, con un plan calculado con un sistema de planificación (SPT), es su verificación con un medio alternativo a dicho cálculo. Para esto se pueden usar diferentes métodos, bien de medida, bien de cálculo y las métricas usadas para decidir si un plan es aceptado o no pueden ser también distintas, siendo la más frecuentemente usada un análisis gamma de la distribución de dosis en un plano significativo, dentro de un maniquí.

Tras la puesta en marcha de la técnica de IMRT en el Hospital Clínico Lozano Blesa de Zaragoza, la experiencia demostró que con los distintos métodos de verificación que habían estado disponibles a lo largo el tiempo, los resultados en la aceptación/rechazo de planes, podían ser contradictorios en algunos casos, incluso aunque se hubiera usado la misma métrica. Este problema también ha sido señalado en la literatura.

En este trabajo se ha tratado de paliar este problema, buscando una métrica que condujera a una aceptación/rechazo del plan de forma más fiable, fundamentalmente desde el punto de vista de la relevancia clínica de las variaciones encontradas. Para ello, además de las métricas tradicionales, se han propuesto otras nuevas que se han evaluado de distintos modos y se ha establecido la correlación existente entre ellas. Las métricas están basadas en el análisis de la información presente en los histogramas dosis-volumen, y, a su vez, pueden basarse en diferencia de dosis, cumplimiento de objetivos y diferencias en los índices radiobiológicos.

Con este fin se han analizado 58 planes de tratamiento de pacientes con distintas localizaciones tumorales. Todos ellos se han verificado previamente al tratamiento, usando métricas tradicionales, y, además, posteriormente al mismo, de forma retrospectiva, con los nuevos métodos de verificación que obtienen los histogramas dosis volumen a partir de medidas realizadas en la unidad de tratamiento. Se investiga si los nuevos métodos conducen al mismo resultado de aceptación/rechazo del plan que el obtenido inicialmente.

El punto de partida para este trabajo ha sido disponer de un buen algoritmo de cálculo de dosis en el paciente en el SPT, por lo que se comenzó poniendo a punto un algoritmo de superposición de cono colapsado (SCC), cuya exactitud está suficientemente avalada en la literatura.

Del mismo modo, es bien reconocida la gran aceptación de los algoritmos de cálculo por simulación Monte Carlo (MC), por lo que también se ha desarrollado uno de este tipo en este trabajo. Para ello ha sido necesaria la adecuación del programa BEAMNRC al colimador

multilamina utilizado, que no existía en la literatura, y la elaboración de una aplicación informática que permitiera calcular la dosis en el paciente a partir del fichero RTplan y el conjunto de imagen CT. Esta aplicación también produce una salida de datos que puede ser importada en el SPT para poder realizar una comparación de los resultados entre SCC y MC. SCC y MC han sido verificados experimentalmente y los resultados han sido satisfactorios.

Se ha concluido en este trabajo que el algoritmo de cálculo utilizado en la reconstrucción de la dosis de verificación en el paciente tiene influencia en el resultado final de la verificación, teniendo una incertidumbre asociada que debe ser tomada en cuenta.

Se ha visto que las métricas de diferencias en dosis pueden ser representadas convenientemente por un parámetro único: la función global de diferencias en dosis y con un punto de corte óptimo que puede obtenerse del análisis ROC de la métrica.

En un intento de correlacionar métricas basadas en parámetros radiobiológicos con las basadas en diferencias de dosis, se ha observado que no siempre es posible. La razón es que, en numerosas ocasiones, las dosis de restricción impuestas a un plan están lejos de las dosis que suponen cambios significativos en los índices radiobiológicos, y, por tanto, estos son insensibles a diferencias grandes de dosis que puede haber entre la dosis de verificación y la de referencia. Como alternativa se propone en este trabajo utilizar métricas radiobiológicas modificadas que conduzcan a una mejor correlación con las de diferencias de dosis.

En todo caso, es preferible la elección de planes robustos desde el punto de vista radiobiológico, lo que puede ser tenido en cuenta antes de aceptar un plan. En particular, se ha propuesto en este trabajo un índice de robustez, con valores entre 0 y 1, que puede ser calculado para cada plan y que es un indicativo de la probabilidad de que un plan incumpla métricas relacionadas con los índices radiobiológicos modificados.

Se ha concluido también que las métricas basadas en la función gamma 3D global, para cada uno de los volúmenes considerados en la planificación, tienen una clara correlación con las métricas de diferencias en dosis, y permiten predecir con mucha sensibilidad qué planes resultaran rechazados en las verificaciones.

Como resumen acerca de la verificación de planes, se propone un análisis que parta de planes que sean lo más robustos posibles, evaluándose posteriormente la función global de diferencias en dosis, conjuntamente con la investigación, si es el caso, de las circunstancias que producen el rechazo del plan en dicha función.

Abstract

As a previous step to deliver an IMRT treatment to a patient, with a plan designed by a treatment planning system (TPS), it is required to verify the calculation with an alternative system. In order to do that, we may employ calculation-based approaches or measurement-based approaches. To decide whether the plan is acceptable or not, we can use different metrics, the gamma analysis of a planar dose in a significant plane inside a phantom being one of the most popular.

After the IMRT commissioning at the Hospital Clinico Universitario “Lozano Blesa” in Zaragoza, our experience has shown that using the different verification methods available during the first years, the results in plan acceptance or rejection could be contradictory in some cases, even if the same metric was used. This issue has also been shown in the literature.

The aim of this work has been to offer solutions to this problem, searching for new metrics that could give more reliable acceptance/rejection criteria from a clinically relevant point of view. Therefore, apart from conventional metrics we have proposed new ones that have been evaluated with new tools, and we have established what correlations can be found among them. The metrics are based on the processing of the dose-volume histograms information and this process is in turn based on dose differences, constraints compliance and biomathematical treatment outcome models.

With this purpose we have analyzed 58 treatment plans concerning several patient locations. All of them have been verified previously to the treatment, using conventional metrics, and in a retrospective manner after the treatment, with the new metrics which provide dose volume histograms from measurements obtained at the treatment unit. We wonder if they can retrieve the same acceptance/rejection plan distribution as obtained initially.

The starting point has been to develop a state-of-the-art calculation algorithm for the TPS. This algorithm is a collapsed cone superposition algorithm (CCS), which is well benchmarked, as can be found elsewhere, as an appropriate algorithm to calculate dose inside the patient in IMRT treatments.

In the same way, the Monte Carlo (MC) algorithm is widely accepted in radiotherapy plan calculations, so it has been employed as a reference redundant calculation. To use MC as intended, it has been necessary to adapt the BEAMNRC package to our particular multileaf collimator (MLC), developing a new component module, and a new software application to

perform the patient dose calculation, using the RTplan file and the patient CT image set. This application produces an output that can be imported from the current TPS.

Both calculation methods, CCS and MC, have been successfully validated using measurements.

We have concluded that the calculation algorithm used in the verification patient dose reconstruction influence the verification results, having a non-negligible uncertainty that has to be taken into account.

The dose difference metrics can be condensed in a unique parameter: the dose difference global function, with an optimal cutoff that can be extracted from a metric ROC analysis.

When trying to correlate differences in biomathematical treatment outcome models with dose difference metrics we have found that this is not always possible. This is due to the fact that the constraint dose is often away from the significant dose that can affect the radiobiological model and therefore they are insensible to big dose differences between the verification system and the TPS. As an alternative we propose to employ modified radiobiological models that can drive to a better correlation.

In any case, it is better to choose robust plans from a radiobiological point of view, and this can be taken into account before plan acceptance. In particular, we have proposed a robustness index in this work, which can take values from 0 up to 1, and can be calculated for every plan. This index is a good predictor of the probability of plan rejection according to metrics derived from modified radiobiological modes.

The global 3D gamma-based metric calculated for each plan volume, has a clear correlation with the dose difference metrics and allow to predict with good sensibility what plans could be rejected.

In brief, regarding the plan verification process, we propose a procedure that first selects the most robust plans that can be obtained from the TPS, evaluating subsequently the dose difference global function, in conjunction with the investigation, if relevant, of the circumstances that produce the plan rejection under that function.

Índice

1. INTRODUCCIÓN	19
1.1. EL PROCESO RADIOTERÁPICO	19
1.2. LA RADIOTERAPIA DE INTENSIDAD MODULADA (IMRT)	22
1.3. ESTADO DE LA CUESTIÓN DE LAS MÉTRICAS EN LA ACEPTACIÓN DE TRATAMIENTOS	24
1.4. OBJETIVOS DEL TRABAJO	26
1.5. ESTRUCTURA DE LA TESIS	28
2. PACIENTES SELECCIONADOS Y TÉCNICA DE IMRT	29
3. SISTEMAS DE MEDIDA Y VERIFICACIÓN	32
3.1. MEDIDAS DE VERIFICACIÓN CONVENCIONAL	32
3.1.1. MEDIDAS DE DOSIS EN UN PUNTO CON CÁMARA DE IONIZACIÓN	32
3.1.2. MEDIDAS DE DISTRIBUCIÓN DE DOSIS EN UN PLANO CON EPID	34
3.1.3. MEDIDAS EN PLANOS EN EL INTERIOR DE UN MANIQUÍ.	36
3.2. MEDIDAS CON MANIQUÍ DELTA4	38
3.3. MEDIDA CON SISTEMA COMPASS Y RECONSTRUCCIÓN DE DOSIS EN EL PACIENTE	39
3.4. CONTROL DE LA INFLUENCIA DEL MLC EN LAS MEDIDAS	44
4. ALGORITMO DESARROLLADO PARA EL CÁLCULO EN EL SISTEMA DE PLANIFICACIÓN	47
4.1. DESARROLLO DEL ALGORITMO SCC	47
4.1.1. OBTENCIÓN DE ESPECTROS ENERGÉTICOS EFECTIVOS	47
4.1.2. CONTAMINACIÓN ELECTRÓNICA.	49
4.1.3. CÁLCULO DEL TERMA	49
4.1.4. CONO COLAPSADO	50
4.1.5. FACTOR GEOMÉTRICO	51
4.1.6. DETERMINACIÓN DE LA FLUENCIA ENERGÉTICA DE CADA UNIDAD DE TRATAMIENTO E INTRODUCCIÓN DE MODULADORES DEL HAZ	52
4.1.7. TRATAMIENTO DE HETEROGENEIDADES	53
4.1.8. CÁLCULO DE UM	54
4.2. VERIFICACIÓN DEL ALGORITMO SCC	54
5. CÁLCULO MONTE CARLO DE DISTRIBUCIONES DE DOSIS	58
5.1. PARTICULARIDADES DE ADAPTACIÓN A BEAMNRC Y DOSXYZNRC Y UTILIDADES DE MCVERIF	59
5.2. DESARROLLO DEL “COMPONENT MODULE” SMLC Y AJUSTE DE LA UNIDAD DE TRATAMIENTO	63
5.2.1. AJUSTE DE LA UNIDAD DE TRATAMIENTO	63
5.2.2. DESARROLLO DE SMLC	64
6. MÉTRICAS A UTILIZAR, SISTEMAS CLASIFICADORES Y HERRAMIENTAS DE DISCRIMINACIÓN	67
6.1. PUNTO DE PARTIDA DE LA COMPARACIÓN ENTRE HDV Y DEL ESCALADO DE HISTOGRAMAS	67
6.2. HERRAMIENTAS DE EVALUACIÓN	69
6.2.1. GAMMA GLOBAL 2D Y 3D ENTRE DISTRIBUCIONES DE DOSIS.	69

6.2.2.	FUNCIONES PROPUESTAS Y VARIABLES QUE EVALÚAN LA DIFERENCIA ENTRE LOS HDV OBTENIDOS POR DOS MÉTODOS.	71
6.2.3.	FUNCIONES RADIOBIOLÓGICAS	75
6.2.4.	CORRELACIÓN ENTRE SISTEMAS CLASIFICADORES	80
6.2.5.	ANÁLISIS ROC DE LOS SISTEMAS CLASIFICADORES.	82
6.3.	RESUMEN DE MÉTRICAS DE COMPARACIÓN Y VARIABLES BINARIAS	85
6.4.	PROCESADO DE LOS HDV	91
7.	RESULTADOS	93
7.1.	VERIFICACIÓN DEL ALGORITMO DE SUPERPOSICIÓN DE CONO COLAPSADO	93
7.2.	AJUSTE DE LA UNIDAD DE TRATAMIENTO EN EL CÁLCULO POR MONTE CARLO	98
7.3.	INFLUENCIA DEL ESTADO DEL ACELERADOR EN EL RESULTADO FINAL	101
7.3.1.	INCERTIDUMBRE EN LA DOSIS POR UM DEL ACELERADOR	101
7.3.2.	INCERTIDUMBRE EN EL POSICIONAMIENTO DEL MLC	101
7.3.3.	EVALUACIÓN DE LA VALIDEZ DEL ESCALADO PROPORCIONAL ENTRE CMPC Y CPM	108
7.4.	ANÁLISIS ROC DE LAS MÉTRICAS EVALUADAS	111
7.4.1.	ANÁLISIS DE LAS FUNCIONES FI, FIP, FGO Y FG	111
7.4.2.	ANÁLISIS DE LA COMPLEJIDAD DE PLANES	115
7.4.3.	ANÁLISIS DE LOS ÍNDICES RADIOBIOLÓGICOS	116
7.4.4.	ANÁLISIS DEL ÍNDICE GAMMA 3D	122
7.4.5.	CAPACIDAD PREDICTIVA DE LAS MEDIDAS CON CÁMARA	123
7.5.	CORRELACIÓN ENTRE MÉTRICAS DE ACEPTACIÓN	124
7.5.1.	CORRELACIONES PARA DELTA4	127
7.5.2.	CORRELACIÓN ENTRE MC ESCALADO Y CPM	129
7.5.3.	CORRELACIONES PARA LAS MEDIDAS CONVENCIONALES	131
7.5.4.	ÍNDICE DE ROBUSTEZ	136
7.6.	PROPUESTA DE VERIFICACIÓN DE PLANES DE IMRT BASADA EN LAS MÉTRICAS ANALIZADAS	137
8.	CONCLUSIONES	142
8.1.	CONCLUSIONES GENERALES	142
8.2.	CONCLUSIONES PARTICULARES	144
9.	REFERENCIAS	146

1. Introducción

1.1. El proceso radioterápico

La radioterapia es una técnica de tratamiento de la enfermedad oncológica que utiliza la radiación con el objeto de conseguir el control local de la enfermedad. La radioterapia externa o teleterapia utiliza como unidad de tratamiento el acelerador lineal de electrones (ALE), que puede generar tanto haces de rayos-X como de electrones. En el primer caso se interpone un blanco de alto número atómico para producir un espectro continuo de fotones. Hoy en día la mayoría de los tratamientos de teleterapia se realizan con fotones de rayos-X. Las energías típicas de los tratamientos oscilan entre los 6MV y los 18MV, donde MV indica que la energía expresada es cercana a la máxima del espectro en MeV, existiendo un espectro continuo con energías menores, con un máximo en torno a la tercera parte del valor nominal. La cantidad de radiación que genera un ALE se mide en las denominadas unidades de monitor (UM), que están relacionadas con la dosis absorbida que se mide en condiciones de referencia en un maniquí de agua, que se expresa en Gy por UM (Gy/UM).

Antes de que los pacientes se sometan a una sesión de radioterapia, su tratamiento se diseña con un sistema de planificación de tratamientos (SPT). Dicho sistema tiene como entrada series de imágenes del paciente, por norma general de tomografía computarizada (TC), aunque también pueden ser de otras modalidades de imagen. Esas imágenes son tomadas en posición de tratamiento y, por norma general, usando una TC o serie de imagen de simulación específica diseñada para planificar sobre ella. Para el SPT, el paciente es un conjunto de vóxeles estático, cada uno con una determinada densidad electrónica. Sobre las imágenes de simulación del paciente se diseñan los volúmenes de tratamiento. Por un lado se diseña el volumen blanco de planificación (PTV), generado a partir del tumor o de la zona donde se considera que hay determinada densidad de clonógenos, células tumorales con capacidad de proliferación. Por otro lado se contornean o segmentan los órganos de riesgo (OARs) en los que se considera de interés conocer la dosis que recibe determinado volumen. Puede haber varios PTVs así como un número variable de OARs, dependiendo de la

localización y de la estadificación de la enfermedad. En cada PTV se prescribe una dosis que se da en varias fracciones siguiendo el principio básico de la radioterapia: el tejido sano sufre menos daño en comparación con el tumoral si la misma dosis se da fraccionada, dada la mayor capacidad de recuperación del daño del primero.

Una vez diseñados los volúmenes se planifica el tratamiento. En función de la localización y el propósito del tratamiento la complejidad de la planificación es variable. Los SPT disponen de una serie de herramientas que permiten escoger las orientaciones del brazo de la unidad de tratamiento más apropiadas, los giros de colimador, la forma particular de cada campo de tratamiento y si se va a usar algún tipo de modulación del haz que produzca diferencias de intensidad del mismo a lo largo del tamaño de campo. El principio fundamental que guía una planificación de tratamientos de radioterapia es la sencillez, es decir, buscar siempre la manera más simple de conseguir los objetivos clínicos de la planificación. Hoy en día, los objetivos clínicos casi siempre se expresan en términos de restricciones en un análisis estadístico de los volúmenes implicados que se denomina histograma dosis volumen (HDV). Un HDV es un histograma de frecuencias de dosis que analiza cada volumen vóxel a vóxel. Se habla por tanto de HDV del PTV y de HDV de cada OAR. La manera más intuitiva de entender el HDV como histograma de frecuencias es lo que se denomina HDV diferencial. El tratamiento ideal es aquel en el que el HDV diferencial del PTV tiene un solo pico en la dosis prescrita y todos los OARs tienen un solo pico en 0, es decir, todos los vóxeles del PTV están tratados con la dosis que se quiere impartir y el tejido sano recibe dosis 0. En la práctica siempre se produce un compromiso o balance entre restricciones en el PTV y en los OAR y la situación ideal nunca se cumple, siendo necesario asumir una serie de riesgos para el tejido sano. Para ello, el HDV en su expresión integral es la representación más práctica, donde cada punto expresa la dosis que recibe un determinado volumen del PTV o OAR considerado o bien una dosis superior. Se habla de DX y VX como dosis que recibe un volumen X o volumen que recibe una dosis X, y las restricciones del tratamiento se expresan en estos términos. Para los PTV conviene siempre que el 100% del volumen reciba la dosis prescrita (DP), $V_{100} = DP$, pero a su vez que la dosis que recibe un volumen pequeño, supongamos el 1%, D_1 , sea también la dosis prescrita $D_1 = DP$, es decir, que la dosis en el PTV sea lo más homogénea posible. En los OAR las restricciones se suelen dar en términos de que un determinado volumen (V_L) no supere una

dosis límite X , $VX < V_L$, o bien que una determinada dosis límite D_L no llegue a más de un determinado volumen X en porcentaje o centímetros cúbicos, $DX < D_L$.

Dadas unas restricciones, una planificación es satisfactoria cuando el plan de tratamiento diseñado las cumple y además es el más sencillo posible, tanto en términos de operación de la unidad de tratamiento como de tiempo que ha de permanecer el paciente en la unidad de tratamiento.

Una vez que el plan es satisfactorio, el paciente es llamado para ser tratado, y el primer día es posicionado con especial cuidado para verificar que todo lo planificado es correcto y no interfiere con características de la unidad que no se hayan tenido en cuenta. Hoy en día es frecuente realizar una o varias imágenes del paciente antes de dar cada sesión por medio de un dispositivo electrónico de imagen portal (EPID), que son comparadas con las imágenes previstas en determinadas orientaciones de brazo generadas por el SPT. Algunas unidades de tratamiento son capaces de realizar una tomografía computarizada de haz cónico (CBCT) que se puede comparar con la TC de planificación antes y después de realizar el tratamiento. Todo el proceso de imagen tiene como objetivo asegurar que el paciente y sus volúmenes se localizan donde estaba previsto y en su caso que los volúmenes planificados coinciden con los extraídos de la imagen de la fracción de tratamiento considerada. Una vez que el posicionamiento es correcto se imparte el tratamiento. El plan generado por el SPT es transferido automáticamente a la unidad de tratamiento y cada sesión queda registrada en una base de datos, lo que permite hacer un seguimiento de las sesiones impartidas.

Por lo general, la planificación realizada por el SPT se basa en un cálculo sobre el que, con otro sistema diferente, se realiza una verificación paralela o redundante de las UM por campo de tratamiento, pero que no es comprobado experimentalmente porque se considera innecesario. Sin embargo, en función de la complejidad de los tratamientos, determinadas técnicas sí que son comprobadas experimentalmente de alguna manera antes de que se produzca el tratamiento del paciente. Este el caso de la técnica que se describe en el apartado siguiente y a la que se refiere el presente trabajo.

1.2. La radioterapia de intensidad modulada (IMRT)

Dentro de los tratamientos que se realizan en radioterapia externa, la técnica de radioterapia de intensidad modulada (IMRT) es una técnica compleja que consiste en la generación con un haz de rayos-X del ALE, para cada orientación de brazo seleccionada, de una fluencia no homogénea optimizada previamente, generalmente a través de un algoritmo denominado planificación inversa, que tiene en cuenta las restricciones y objetivos del tratamiento planteado. Para generar estas fluencias no homogéneas se utiliza el colimador multilámina (MLC).

Existen técnicas para tratamiento con rayos-X más simples que la IMRT. Utilizan haces de fluencia homogénea o alterada mediante moduladores sencillos denominados cuñas. Es lo que se denomina la radioterapia 3D conformada (3DCRT). En su verificación basta comprobar que la imagen portal que se hace del paciente en la unidad de tratamiento usando el EPID coincide con la que prevé el SPT usando la proyección del campo de tratamiento desde el punto de vista del haz (BEV) y una radiografía digital reconstruida (RDR). Además, usando un software independiente se realiza un cálculo redundante de las UM que proporciona el SPT, usando la dosis calculada en uno o varios puntos. La IMRT se utiliza en aquellos casos en que la ventaja dosimétrica para los OARs es clara, dado que encierra mayor complejidad. Las distribuciones de dosis que se obtienen con IMRT son más heterogéneas y contienen importantes gradientes, incluso dentro de los volúmenes a tratar. Se consigue conformar mejor el volumen y con ello proteger mejor los OAR, generándose incluso concavidades en la distribución de dosis. Por ello se considera, en general, que la IMRT no puede ser verificada con el mismo enfoque que la 3DCRT dado que los planes están compuestos por superposiciones de segmentos de tamaño pequeño y/o elongado y se establecen gradientes de dosis absorbida muy próximos a las estructuras críticas. El uso de la modulación en fluencia permite que cada orientación de brazo sea independiente del tumor y de las estructuras críticas u OARs, permitiéndose que los haces penetren a través de ellas y dejando luego que sea esa modulación la que limite la dosis, no siempre de una manera intuitiva. Por tanto las imágenes portales del EPID no bastan para verificar que un determinado OAR está a salvo; los esfuerzos han de concentrarse en comprobar cuál es el resultado global de la irradiación a través de la dosis absorbida en el paciente durante todo

el tratamiento de la sesión (1). Además, la dosis que recibe el paciente se ve afectada por el estado del acelerador y, en particular, del sistema de colimación, en mayor medida que cuando se realizan tratamientos de 3DCRT (2). El posicionamiento de las láminas del MLC ha demostrado ser crítico y, como consecuencia, existen recomendaciones acerca de una exactitud para dicho posicionamiento inferior al milímetro (3). El control de calidad rutinario es necesario pero la comprobación del tratamiento mediante medidas pre-tratamiento proporciona un método de verificación condensado (4) que tiene en cuenta posibles errores en toda la cadena, desde el modelado de las unidades de tratamiento, los algoritmos de cálculo y hasta el comportamiento del ALE.

Así pues, desde el comienzo de la verificación de tratamientos de IMRT, se ha venido utilizando una sistemática de comparación pre-tratamiento entre los resultados de las medidas realizadas y los del cálculo en el SPT que estaba basada principalmente en la comparación de dosis en puntos y en la comparación de distribuciones de dosis en planos. Ambos tipos de comprobaciones se llevaban a cabo sobre un maniquí preparado para poder realizar medidas en su interior, bien fuera con cámaras de ionización, placas radiográficas o radiocrómicas o matrices de diodos o de cámaras de ionización (5). Las comprobaciones de distribuciones planares pueden ser tanto del tratamiento global como de cada campo individual que lo conforme. Denominaremos a este tipo de estrategias “verificación convencional”.

Actualmente existen varios sistemas que permiten utilizar las medidas pre-tratamiento o durante el tratamiento para alimentar diversos algoritmos de cálculo que permiten obtener una estimación suficientemente precisa de los HDV reales de tratamiento, con el objeto de poder comparar si los objetivos clínicos iniciales del mismo se cumplen (6–9) . Por el hecho de ofrecer resultados directamente relacionados con la calidad clínica del tratamiento, este tipo de métricas derivadas de medidas experimentales constituyen una estimación mucho más realista acerca de si, en el momento de realización del tratamiento, siguen cumpliéndose los objetivos previstos. Esta estrategia será denominada “verificación basada en HDV”. Existen varias posibilidades que permiten estimar ese cumplimiento de objetivos, entre las que cabe mencionar, el recálculo de distribuciones de dosis en paciente a partir de:

- Mapas de fluencia medidos con una matriz bidimensional de detectores (6,7,10).

- Distribuciones de dosis medidas en planos o cilindros medidos con un detector con geometría de aspa o circular en el interior de un maniquí (11).
- Mapas de fluencia medidos con un (EPID) (12).
- Distribuciones de dosis planares a partir de planos de distribución dosis de transmisión medidos con un EPID (13–15).
- Los parámetros geométricos y dosimétricos medidos por el acelerador automáticamente y recogidos en archivos de registro (16).

Un debate abierto y que se resolverá en los próximos años concierne al hecho de que la verificación pre-tratamiento sea sustituida por una verificación durante el tratamiento, basada precisamente en los dos últimos puntos del párrafo precedente y en un control de calidad exhaustivo de los parámetros relacionados con el MLC, los movimientos de la unidad y la tasa de dosis producida por el ALE. Dado que la verificación durante el tratamiento involucra procedimientos de análisis complejos y que requieren un nivel alto de experiencia, las verificaciones pre-tratamiento continúan siendo muy populares y en muchos centros de radioterapia se siguen considerando como el estándar (6).

1.3. Estado de la cuestión de las métricas en la aceptación de tratamientos

En los últimos años numerosos grupos han investigado el uso de nuevas métricas en el análisis de la verificación de planes de tratamientos de IMRT con objeto de obtener resultados con una mayor relevancia clínica (8,9,11,17–19). Se entenderá por métrica una función que procede de la comparación entre dos distribuciones de dosis y que proporciona como resultado tanto una variable continua como una variable binaria.

La métrica tradicional, que en este trabajo denominaremos “convencional”, y que se ha utilizado hasta el presente en la mayoría de ocasiones proviene de la medida de dosis en un punto en el interior de un maniquí con una cámara de ionización y su comparación con el valor previsto por el SPT. A esto se añade la medida de dosis en uno o varios planos en el interior de un maniquí convenientemente preparado y su comparación con los cálculos realizados por el sistema de planificación de tratamientos (5). Para ello se emplean

programas que calculan la similitud entre dos distribuciones de dosis, a través de una métrica en el espacio dosis-coordenadas denominada gamma (Γ) (20–22) que se verá con más detalle en el apartado 6.2. Esta función calcula la distancia en ese espacio punto a punto y mide la longitud del vector más pequeño que se pueda trazar normalizada en unidades de la tolerancia en dosis o distancia. De este modo, si la longitud de dicho vector es inferior a 1, la distribución de referencia y la distribución analizada cumplen las tolerancias establecidas para el punto que se esté analizando. Al final se obtiene un histograma gamma que refleja el número de puntos que cumplen la tolerancia y aquellos que no la cumplen, expresado como porcentaje de puntos. Previo al análisis es necesario establecer un umbral de significación en dosis por debajo del cual no se analizan puntos, y tener las dos distribuciones a comparar, la medida y la calculada con el SPT, normalizadas, bien en un punto donde se conozca que la dosis es la misma, o bien en unidades de dosis para un mismo número de unidades de monitor (UM).

Cuando la medida con cámara de ionización discrepa de lo previsto en una cantidad mayor a un porcentaje establecido o bien cuando el porcentaje de puntos que no cumplen el criterio gamma fijado supera un valor preestablecido (típicamente 5-10%), la verificación resulta insatisfactoria y, por tanto, hay que investigar las causas de la discrepancia y ulteriormente, rechazar el tratamiento planificado. Esta sistemática puede conducir a lo que se denominan falsos positivos y falsos negativos, ya que se da el caso de distribuciones que discrepan en un número de puntos significativo, y sin embargo clínicamente puede no tener relevancia, y también circunstancias en las cuales los puntos que no cumplen el criterio gamma son pocos pero afectan a alguna zona crítica que no debería pasarse por alto (23–26). La primera circunstancia deriva inevitablemente en un gasto innecesario de recursos tanto en la replanificación del tratamiento como en una nueva verificación, mientras que los falsos negativos permiten tratar al paciente con distribuciones de dosis que incumplen los objetivos clínicos y por tanto podrían ser potencialmente peligrosas para él. Adicionalmente, cuando las distribuciones de dosis son tridimensionales, el criterio gamma a aplicar también es tridimensional, lo que puede conducir a resultados distintos que los que se obtienen utilizando distribuciones planares (25–27). Además, ha sido acreditado que distintos dispositivos de medida con distinta geometría usados en las mismas verificaciones conducen a resultados diferentes, perdiéndose la correlación entre ellos (11).

Las nuevas métricas que se proponen en la literatura tratan de evitar estos inconvenientes mediante una evaluación de parámetros clínicos en vez de una mera comparación de la similitud de las distribuciones. Las aproximaciones que suponen son variadas y de momento no está clara su exactitud, su validez y sus rangos de utilización (17). Se trata de valorar por un lado métricas basadas en las diferencias entre medida y cálculo en los HDV, que se obtienen en la verificación y extraídos del SPT respectivamente, y, por otro lado, otras que tienen en cuenta, además, modificaciones en los índices radiobiológicos calculados a partir de la matriz de dosis en el paciente (19).

Para realizar un cálculo redundante al del SPT, también existen una serie de herramientas basadas en el cálculo Monte Carlo (28–30) que permiten calcular las distribuciones de dosis en pacientes de IMRT (31,32). Estas herramientas permiten obtener una distribución de dosis que se podría considerar como de referencia siempre y cuando la unidad de tratamiento esté bien modelada y el algoritmo utilizado esté validado para las condiciones en las que se está usando (33).

Los dispositivos electrónicos de imagen portal (EPID) han sido caracterizados también en los últimos años con objeto de poder medir distribuciones de dosis tanto en modo de transmisión como en modo de no-transmisión (12–14,34–38). Con ellos es posible tanto evaluar los mapas de fluencia como la dosis en un plano en el interior del paciente a partir de la medida de la transmisión a través del mismo.

A la hora de evaluar la eficiencia de un método que clasifique planes de IMRT, calculados con el SPT, entre aceptables e inaceptables, se necesitan lo que se denominan sistemas clasificadores. Estos sistemas pueden ser comparados entre sí con el objeto de extraer conclusiones sobre cuál es más conveniente y cómo están relacionados.

Todos estos elementos permiten diseñar una estrategia de análisis encaminada a redefinir las herramientas que en el futuro van a servir para realizar la verificación de tratamientos de IMRT.

1.4. Objetivos del trabajo

El objetivo principal de este trabajo es la propuesta y evaluación de nuevas métricas que permitan correlacionar los resultados de las medidas pre-tratamiento realizadas en

tratamientos de IMRT con el impacto clínico en el paciente de la distribución de dosis absorbida, indicando cuáles son las mejores y si pueden resumirse en un número pequeño de indicadores.

Se desarrollará un método para la evaluación de las métricas y su clasificación, así como una utilidad informática que permita el procesado de los datos. Las métricas empleadas corresponderán tanto a variables continuas como a variables binarias. Se compararán las métricas que provienen de evaluar diferencias en dosis con aquellas que provienen de la evaluación de los índices radiobiológicos a partir de los volúmenes definidos en cada plan de tratamiento. También se buscará si existe una buena correlación entre la clasificación de planes realizada por métodos convencionales y las métricas basadas en el análisis de HDV. Una vez establecidas las métricas más apropiadas se establecerá una sistemática en la verificación de planes de IMRT con la tecnología disponible.

Un objetivo paralelo es comparar los resultados que proporcionan varios algoritmos y sistemas de medida diferentes que generan distintos conjuntos de HDV. Estos conjuntos se obtienen tanto a partir de medidas en la unidad de tratamiento como a partir de algoritmos de cálculo redundante que importan los datos que exporta el SPT.

A la hora de comparar resultados, y tal como se menciona en el apartado anterior, se puede considerar en dosimetría una buena referencia los cálculos realizados por el método de Monte Carlo (MC), por lo que, un objetivo previo de este trabajo ha sido establecer dicha forma de cálculo adaptada al ALE con el que se llevan a cabo los tratamientos, que no existía en la literatura (39).

Asimismo, otro objetivo previo ha sido el desarrollo y verificación de un algoritmo de cálculo basado en superposición de cono colapsado, que es el que utiliza el SPT con el que se planifican los pacientes que se van a estudiar (40). Este algoritmo se ha desarrollado como mejora del algoritmo que denominaremos “preciso”, con el que se trabajaba previamente.

Un objetivo particular del trabajo es evaluar las incertidumbres involucradas en el proceso de verificación y separar aquellas que se deben al comportamiento de la unidad de tratamiento de aquellas que se deben a los algoritmos de cálculo utilizados. Se estudiará el comportamiento del MLC del que se dispone y su influencia en el resultado de la verificación.

1.5. Estructura de la tesis

Esta tesis está estructurada en 9 capítulos. El primero corresponde a la introducción con la presentación de los objetivos del trabajo y la problemática asociada a la cuestión planteada. En el segundo se describe el grupo de pacientes seleccionado para el análisis y la técnica de tratamiento empleada. En el tercer capítulo se presentan los sistemas de medida empleados en la verificación de los tratamientos. En los capítulos 4 y 5 se resumen las características de dos algoritmos de cálculo desarrollados por el autor, y que son utilizados para calcular las distribuciones de dosis por el sistema de planificación de tratamientos y para realizar una verificación redundante basada en HDV por el método de Monte Carlo. En el capítulo 6 se describen las métricas que se van a emplear en la verificación, para posteriormente, en el 7, presentar todos los resultados asociados a dichas métricas así como la verificación de la validez de los algoritmos desarrollados. En el capítulo 8 se muestra una relación de las conclusiones obtenidas, tanto generales como particulares, distinguiéndose de esta manera aquellas que no dependen de los sistemas de medida, cálculo y tratamiento empleados en el trabajo como aquellas que sí lo hacen. Por último, en el capítulo 9 se presentan las referencias utilizadas.

2. Pacientes seleccionados y técnica de IMRT

Se ha creado una base de datos que consta de 58 planes, tratados en un ALE presente en el Hospital Clínico Universitario “Lozano Blesa” de Zaragoza (HCUZ) con la técnica denominada IMRT estática. Los planes cubren las siguientes localizaciones:

- Cervical
- Cerebro
- Próstata y ganglios linfáticos asociados
- Pelvis ginecológica

En la tabla 2.I pueden verse los casos estudiados y la prescripción de dosis utilizada, indicándose el número de planes por prescripción.

Tabla 2.I Número de planes por localización y prescripción de dosis.		
Localización	Número total de planes	Planes por prescripción Dosis prescrita (cGy)→ número de planes
Cervical	32	7000 → 3 6880 → 2 6790 → 1 6600 → 25 6000 → 1
Cerebro	10	4600 → 1 4200 → 2 6000 → 3 5400 → 1 5220 → 1 5000 → 2
Próstata	13	7600 → 1 7800 → 4 7400 → 2 5600 → 2 5000 → 1 3200 → 1 2800 → 1 4600 → 1
Pelvis	3	5126, 5000 y 5100

Los pacientes escogidos tienen diferente número de volúmenes blanco de tratamiento (PTVs), siendo además la prescripción de dosis en numerosas ocasiones diferente para cada uno de los volúmenes. En tratamientos de cabeza y cuello se suelen utilizar de 1 a 3 PTVs con dosis equivalentes de 70 Gy, 60 Gy y 50 Gy en un esquema de 2Gy por fracción, no estando siempre presentes todos en función de la extensión de la enfermedad del paciente en cuestión. En tratamientos de próstata se suelen delimitar dos volúmenes, uno de pelvis ganglionar que se irradia hasta un equivalente de 46 o 56 Gy a 2 Gy por fracción y otro de próstata que suele llegar hasta 78 Gy a 2 Gy por fracción. En tumores cerebrales la dosis que suele prescribirse al tumor alcanza los 60 Gy utilizándose en ocasiones un volumen exterior que se irradia hasta 42 Gy a 2 Gy por fracción. En el caso de la pelvis ginecológica las zonas ganglionares suelen llegar hasta 46-50 Gy a 2 Gy por fracción y los tumores hasta 60Gy.

Los órganos de riesgo principales en cada localización son: parótidas, médula, nervios ópticos, quiasma, tronco cerebral, cristalinos, laringe, cavidad oral, mandíbula y tiroides para cabeza y cuello, y recto, vejiga, cabezas femorales, riñones y médula para zonas pélvicas y abdominales. Los órganos de riesgo suelen clasificarse entre aquellos que se pueden considerar en serie o en paralelo, en función de que el daño grave en una de sus unidades suponga la inutilización completa del órgano (en serie) o no (en paralelo). Ningún órgano es completamente serie o paralelo, salvo quizás la médula, pero, en función de cómo se pueda interpretar el órgano o el efecto radiobiológico que se quiera investigar interesará más documentar la dosis media en el caso de los órganos en paralelo o las dosis altas a pequeños volúmenes en el caso de órganos en serie.

La unidad de tratamiento ha sido en todos los casos un acelerador Siemens Oncor Impression Plus (Siemens Medical Systems) dotado con auto-secuenciación de campos y capacidad para hacer IMRT estática, con un colimador OPTIFOCUS de 41 pares de láminas opuestas de tungsteno de 7,5 cm de espesor que proyectan una anchura de 1 cm en el isocentro. Las energías de fotones de las que dispone la unidad son 6 y 15MV y con ambas es posible realizar la técnica de IMRT.

El sistema de planificación con el que fue diseñado su tratamiento fue PCRT3D 6.0 (Técnicas Radiofísicas). Todos los pacientes siguen un proceso de obtención de imágenes de planificación, consistente en una TC realizada con un escáner GE HiSpeed II. Posteriormente,

sobres los cortes de la TC, se realiza el contorneo de las estructuras críticas y se delinear los volúmenes a tratar.

El siguiente paso, una vez contorneadas las imágenes del paciente que va a seguir un curso de IMRT, es el diseño del plan usando planificación inversa más una segmentación posterior de las fluencias ideales, de acuerdo con las restricciones de optimización introducidas por el usuario para cada caso concreto, que involucran al PTV y a los OAR. De este modo se obtienen una serie de segmentos por cada orientación de brazo con un peso proporcional a las UM que tendrá el segmento en el acelerador. El campo de tratamiento para una determinada orientación de brazo está formado por la suma de todos los segmentos que lo constituyen. El módulo de optimización del SPT contiene la planificación inversa y la segmentación, siendo ésta transferida posteriormente al módulo de cálculo definitivo. Durante la planificación inversa se buscan de manera iterativa las fluencias ideales que cumplen los objetivos introducidos por el usuario. Estas fluencias tienen que ser reproducidas por el MLC en forma de una serie de segmentos discretos que son las que componen una fluencia real que debe parecerse a la ideal pero que siempre degrada un poco los histogramas conseguidos con la planificación inversa. Una vez hecha la segmentación, se pasa al cálculo en el paciente donde el usuario recalcula la distribución mediante un algoritmo de superposición de cono colapsado (40) y puede seguir modificando la distribución, si lo desea, cambiando la forma de los segmentos y recalculando sucesivamente. PCRT3D permite la exportación del plan, las estructuras y la matriz de dosis en formato DICOM-RT (RTPlan, RTStruct y RTDose). También es posible exportar los HDV en formato ASCII (del inglés “American Standard Code for Information Interchange”).

3. Sistemas de medida y verificación

3.1. Medidas de verificación convencional

Durante muchos años se ha considerado el estándar en la verificación de IMRT el empleo de la medida conjunta de la dosis absorbida en un punto que no esté situado en una zona con alto gradiente de dosis más la medida de dosis relativa en uno o más planos dentro de un maniquí de geometría similar al paciente. Se ha considerado un enfoque equivalente el empleo de la dosis absorbida en un punto más la comprobación de la distribución de dosis por medio de un detector planar en un plano perpendicular al haz, para cada haz de manera independiente y sin que exista un maniquí que simule el paciente.

Las medidas presentadas en este apartado fueron realizadas en su día para la aceptación clínica de los planes de tratamiento y el conjunto de ellas proporciona el resultado de la verificación convencional. En los apartados 3.3 y 3.4 se presentan nuevos sistemas de medida que han servido para revisar los planes de manera retrospectiva.

3.1.1. Medidas de dosis en un punto con cámara de ionización

La cámara de ionización empleada para medidas puntuales dentro de un maniquí en este trabajo ha sido una cámara PinPoint CC01 de IBA dosimetry, con su correspondiente certificado de calibración para dosis absorbida en agua. Las medidas se han realizado siempre dentro de un maniquí de láminas de agua sólida (RW3) denominado IMRT Phantom (IBA dosimetry) (figura 3.1)

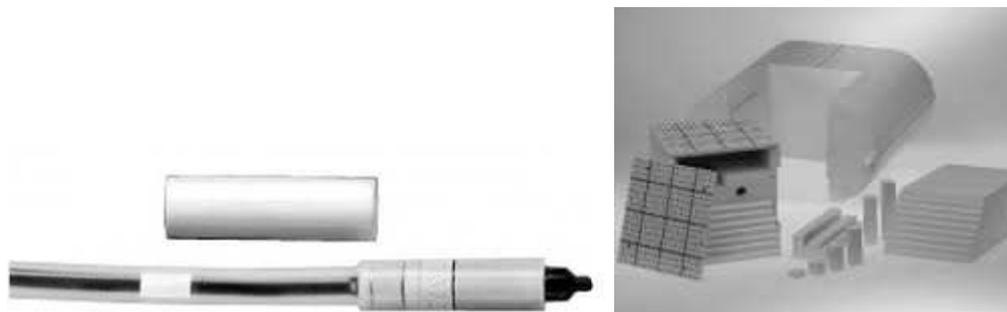


Figura 3.1.- Cámara CC01 y maniquí IMRT Phantom con insertos para la cámara y láminas que permiten su colocación en diferentes puntos.

Respecto a la cámara de ionización, su volumen es pequeño (0,01 cm³) y permite la medida de dosis en zonas de alto gradiente, con una señal suficiente, siendo su sensibilidad del orden de 4·10⁻¹⁰ C/Gy. Por tanto para 2 Gy la lectura es del orden de 800 pC. En las medidas se ha empleado un electrómetro UNIDOS (PTW) para recoger la carga, seleccionándose el rango bajo para obtener la lectura de cada campo de IMRT. La carga recogida se transformó a dosis absorbida para cada campo utilizando la siguiente expresión:

$$D_{p,campo} = \frac{N_{D,w} \cdot (L_Q - \frac{t_{medida}}{t_{fuga}} (L_{fuga_1} + L_{fuga_2}) \cdot 0,5) \cdot F_{P,T} \cdot k_q \cdot F_{estabilidad}}{h_{RW}}$$

(ec 3.1)

siendo

$N_{D,w}$ el coeficiente de calibración de dosis absorbida en agua en cGy/nC para la cámara CC01

L_Q la Lectura en nC

t_{medida}/t_{fuga} el tiempo de medida de la señal dividido por el tiempo de medida de la fuga del conjunto cámara electrómetro

$L_{fuga_1} + L_{fuga_2}$ la suma en nC de las fugas medidas al principio y al final de la sesión

$F_{P,T}$ el factor de presión y temperatura

k_Q el factor específico de la cámara para la calidad del haz empleada

$F_{estabilidad}$ la corrección por la diferencia en dosis absorbida por unidad de monitor introducida en el SPT con respecto a la dosis absorbida por unidad de monitor en el día de la medida

h_{RW} el factor de corrección por el empleo del material RW3 y sus diferencias de equivalencia en agua. Para 6MV es 0,989 y para 15MV es 0,975

El maniquí IMRT Phantom puede configurarse, según la localización del plan, con una geometría cúbica (16 cm x 16 cm x 16 cm) o bien como si fuera una pelvis añadiendo sendas piezas con forma de semicilindro en los laterales. La cámara puede ser colocada en cualquier posición dentro del cubo en saltos de resolución de 1 cm y en las tres direcciones del

espacio. Esto es útil en aquellos casos en los que el punto de medida no coincide con la zona central del cubo. En general siempre se busca que el punto de medida tenga a su alrededor un gradiente no muy pronunciado de dosis. Cuando se prepara el plan de verificación, se comprueba que en el entorno del punto escogido la dosis no varíe más del 3% en 2 mm en cualquier dirección.

En la verificación con este tipo de medidas se busca que la medida de dosis total del plan coincida con la calculada por el SPT en la misma situación en el entorno del 3%.

3.1.2. Medidas de distribución de dosis en un plano con EPID

Desde que los dispositivos electrónicos de imagen portal (EPID) fueron desarrollados y comercializados junto con aceleradores lineales de electrones con el propósito de verificar la posición del paciente, o para monitorizar los movimientos de la estructuras anatómicas de mismo durante el tratamiento, ha existido el interés en investigar su uso como dosímetros. Tras la aparición de los dispositivos basados en silicio amorfo (aSi), se ha estudiado su comportamiento para fines dosimétricos de forma muy extensa, con buenos resultados en general de reproducibilidad, linealidad, efectos de memoria y otras características (35,41–43)

El principio de funcionamiento se ilustra en la figura 3.2 y es análogo al de los detectores usados en radiología digital. El primer paso es convertir los rayos X incidentes en fotones de luz visible utilizando un conversor de rayos X hecho con una placa metálica de aproximadamente 1 mm de grosor y una pantalla que actúa como centelleador. Estos fotones de luz visible son detectados por un fotodiodo de silicio, que actúa como un condensador e integra la señal hasta su lectura. Se trata, pues, de una detección indirecta de los rayos X. Los sistemas basados en detección directa se encuentran actualmente en fase de desarrollo y aún no están incluidos en los ALE.

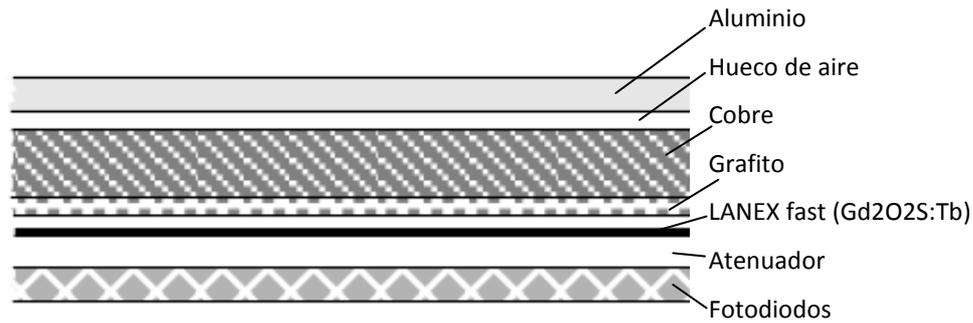


Figura 3.2.- Corte longitudinal de un dispositivo electrónico de imagen portal. La lámina de LANEX es un centelleador que emite luz visible que recogen los fotodiodos, tras llegar electrones procedentes de la lámina de cobre.

El objeto de la lámina metálica de cobre es generar electrones de alta energía y absorber los rayos X de baja energía, ya que son indeseables debido al aumento de la respuesta de los EPID a los fotones de baja energía. Todos los EPID de silicio amorfo (aSi) tienen un centelleador de Gd₂O₂S:Tb.

El EPID utilizado en el acelerador Oncor es un Optivue 1024, que a su vez es un modelo Perkin-Elmer 1640 AG9. El tamaño de pixel es de 0,34 mm y utiliza 7 fotogramas por segundo aproximadamente.

Se ha desarrollado un modelo (44) basado en el modelo de Nijsten (14) para la dosimetría con el EPID en modo de “no transmisión”.

El proceso que se sigue para verificar los planos calculados en el SPT es el siguiente: se calcula en el SPT cada campo por separado sobre un maniquí plano, se normaliza a una dosis adecuada para que el número de UM por segmento no sea demasiado grande y se exporta un plano en el isocentro a 2 cm de profundidad en el maniquí, ya que este es el espesor de la zona de acumulación estimado para el EPID.

El plan se exporta al sistema de registro y verificación (SRV) y se irradia desplegando el EPID para cada orientación de brazo. Al final se obtienen tantas imágenes como orientaciones de brazo. Cada imagen se compone de las imágenes individuales de los segmentos que a su vez se componen de varios fotogramas.

La utilidad desarrollada se denomina JD2D (Java Dicom to Dose) y lee los ficheros con extensión .dcm generados por el acelerador para cada imagen, los reinterpola para obtener una matriz de 512x512 píxeles y realiza las correcciones incluidas en la ecuación de

conversión de valor de gris a dosis (44), incluida la deconvolución con el kernel de campo. Para deconvolucionar se utiliza el método de Richardson-Lucy, que permite construir la deconvolución a partir de convoluciones iterativas partiendo de una señal inicial estimada. La resolución del EPID permite obtener mapas de fluencia-dosis en maniquí muy detallados, que reflejan el efecto de machihembrado y problemas en el posicionamiento de las láminas. En la figura 3.3 se puede observar el detalle de la medida (línea verde), frente al cálculo (línea roja) para un perfil perpendicular al movimiento de las láminas de un campo concreto de IMRT.

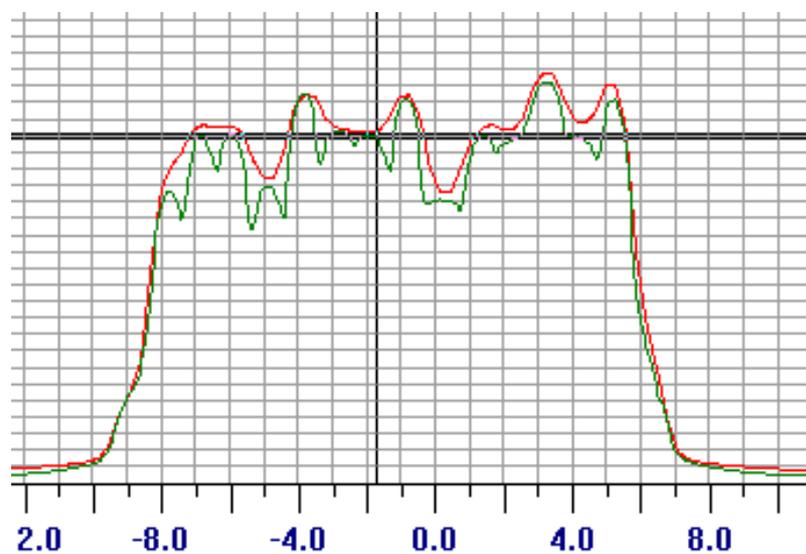


Figura 3.3.- Perfil de dosis medida con el EPID (línea verde) frente a calculada (línea roja).

La evaluación de la verificación con EPID de un plan de IMRT se hace a través de la función gamma 2D que se describe en el apartado 6.2

La asignación del 100% se realiza de manera arbitraria, pero se busca que sea una dosis representativa para cada campo. El criterio gamma fijado es 3%-3mm, buscándose que el % de puntos que tengan $\Gamma > 1$ sea inferior al 10%

3.1.3. Medidas en planos en el interior de un maniquí.

Se dispone de dos maniqués para realizar la verificación convencional de planos de distribución de dosis.

Uno de ellos es el mismo que se describe en el apartado 3.1.1 para las medidas con cámara de ionización (IMRT phantom), ver figura 3.1. En él se pueden insertar placas radiocrómicas de tamaño 16 cm x 16 cm que, tras un proceso de calibración, generan un plano de dosis a comparar con el mismo que se exporta del sistema de planificación con el plan calculado en el mismo maniquí. El procedimiento de calibración y uso de las placas radiocrómicas EBT2 está descrito en muchas publicaciones (45–48), y se empleó una macro del software de libre disposición ImageJ, desarrollada en el HCUZ, para su procesado. Se pueden medir con placa una o varias distribuciones planares de dosis en el interior del maniquí, y estos planos pueden ser axiales, sagitales o coronales. En general, se recomiendan planos sagitales o coronales ya que la orientación del colimador en la mayoría de pacientes es cercana a 0 grados, y, si se mide un plano axial apenas entran un juego unos pocos pares de láminas en la verificación. De los 58 planes estudiados en este trabajo se ha empleado este tipo de placas en 8. La resolución que proporcionan las placas es superior a la de la matriz de cálculo de los planes, por lo que el número de puntos por plano donde se verifica el plan depende de esta última. Una vez se puso en marcha el sistema de medida de planos de dosis con EPID, descrito en el apartado anterior, dejaron de usarse placas, dada la mayor comodidad del nuevo sistema.

En algunos pacientes se ha empleado como método de verificación redundante un detector planar en el interior del maniquí Multicube (figura 3.4). La matriz utilizada es una MatriXX (IBA dosimetry) con sensor de ángulos. Consta de 1020 cámaras de ionización, con una distancia entre centros de 7,62 mm. Las cámaras son cavidades cilíndricas de 5 mm de altura y 4 mm de diámetro. Su resolución es de 7,62 mm en ambas direcciones y cubre un campo de 24,4 cm si se coloca en el isocentro. En las condiciones utilizadas en la verificación convencional proporciona un plano de dosis con una resolución más pobre que la de la matriz 3D de cálculo de los planes con el SPT, que suele tener una resolución de entre 3 y 4 mm. La matriz permanece estática sobre la mesa en el interior del Multicube, y es necesario aplicar correcciones angulares a la respuesta de la cámara (49). La razón de emplear este sistema para la verificación de la dosis en un plano ha sido comprobar algunos planes que estaban en el límite aceptable de la verificación realizada con cámara y EPID o placas radiocrómicas. En concreto la configuración MatriXX + Multicube se ha empleado en 7 pacientes.

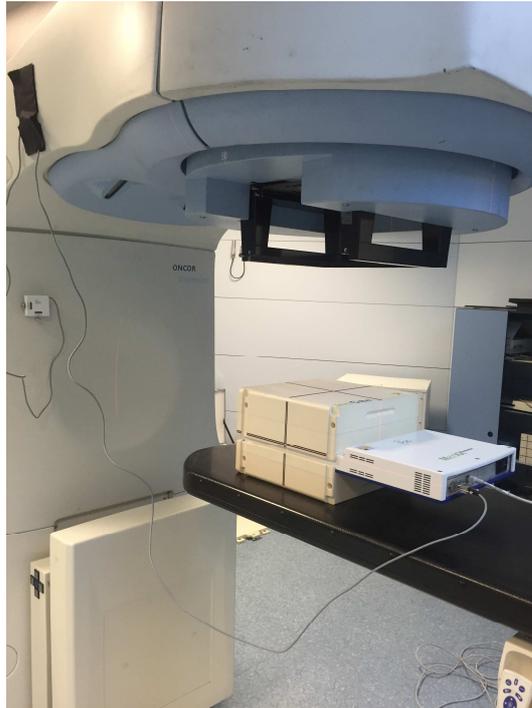


Figura 3.4.- Detector MatriXX en el interior del maniquí Multicube. Se aprecia el sensor de ángulos adherido al brazo de la unidad de tratamiento.

Los planos de dosis medidos tanto con placa, como con MatriXX, como con EPID, pueden ser importados en Omnipro l'MRT (IBA dosimetry), que es el software que permite compararlos y cuantificar las diferencias existentes entre ellos y con respecto a los del SPT.

3.2. Medidas con maniquí Delta4

Otro sistema de verificación utilizado en el presente trabajo es el software Delta 4 (ScandiDos) y su detector asociado, formado por el conjunto de dos matrices de 1069 diodos distanciados 5 mm en el área central (6 cm x 6 cm), y 10 mm en el resto de los 200 mm x 200 mm de superficie detectora. Estas matrices se encuentran situadas en forma de aspa, y cuales se cruzan dentro de un cilindro de 22 cm de metacrilato (figura 3.5). Lleva incorporado un inclinómetro que registra la posición del brazo en cada momento. Los haces recogidos por el detector son solamente aquellos que interaccionan al menos en una de las matrices de detectores. La reconstrucción 3D de la dosis absorbida medida se hace utilizando la distribución de dosis calculada por el SPT, que es escalada en función de las medidas que se obtienen en cada punto del sistema detector, usando la trayectoria de cada

haz y después sumando todas las trayectorias para todas las angulaciones y para todos los tos obtenidos. Esta circunstancia es importante tenerla en cuenta porque implica que el

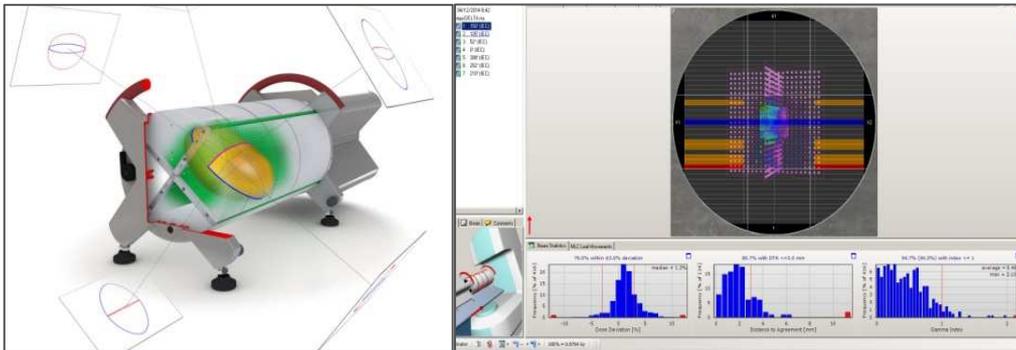


Figura 3.5.- Detector Delta4 (izq) y software de reconstrucción de dosis (dch). Las medidas tomadas en los planos en forma de X sirven para reconstruir la dosis en 3D. Es posible dibujar histogramas y hacer un análisis gamma de cada volumen. Los HDV pueden ser exportados para su análisis posterior.

punto de partida de la verificación es la propia matriz del SPT, justo aquella que se quiere verificar.

Aunque el software del Delta 4 tiene la posibilidad de introducir la TC de planificación del paciente para reconstruir la dosis medida sobre ésta y así comparar sobre el volumen real la planificación frente a la medida (módulo “Anatomy”), no se tenía configurada esta opción, así que los resultados obtenidos son respecto del maniquí cilíndrico con los PTV (volúmenes de tratamiento) y los OAR (órganos de riesgo) insertados en su interior. Se pueden comparar los HDV de los OAR y de los PTV para la fracción medida y la calculada por el SPT. También es posible realizar un análisis gamma para cada volumen.

Con Delta4 no se han evaluado todos los pacientes del trabajo, por motivos de disponibilidad del equipo. La muestra de pacientes es de 36 de los 58 totales.

3.3. Medida con sistema Compass y reconstrucción de dosis en el paciente

El sistema Compass para reconstrucción de dosis en el paciente parte de las distribuciones planares medidas con la matriz de cámaras de ionización.

Consta de dos partes bien diferenciadas. Por un lado se utiliza como hardware el detector MatriXX descrito en 3.1.3, colocado solidario al brazo en un plano que contiene al isocentro

y es perpendicular al eje del haz (figura 3.6). Se sitúa en dicho punto por medio de un soporte que va acoplado al brazo, y en las condiciones de medida del presente trabajo se colocan 2 cm de espesor de acumulación. Una calibración previa permite establecer la relación entre el número de unidades de monitor impartidas por el acelerador y la dosis absorbida en la matriz.

La segunda parte del sistema es el software Compass 3.0 (IBA dosimetry), que permite calcular la distribución 3D de dosis sobre el paciente virtual establecido a partir de la TC de planificación y las estructuras contorneadas, utilizando las fluencias medidas por la MatriXX de todos los haces contenidos en el plan previsto de tratamiento (6,7,10).



Figura 3.6.- Detector MatriXX acoplado al acelerador con un soporte de desplazamiento milimétrico. Se observa el sensor de ángulos acoplado al brazo y un espesor de acumulación de 2 cm colocado sobre el detector.

Cuando se realiza el cálculo en un paciente a partir de la fluencia medida con el detector, la resolución del mismo (7.62 mm) es una limitación, especialmente comparada con las que se pueden conseguir con sistemas como una película o el EPID. Esta limitada resolución es inicialmente una traba del sistema para reconstruir adecuadamente la fluencia, más si se tiene en cuenta que para pequeños segmentos de IMRT se necesita una frecuencia espacial de al menos $0,4 \text{ mm}^{-1}$ para representar la distribución de dosis verdadera (7). El sistema Compass soluciona este problema usando un modelo que permite predecir la fluencia del acelerador y corrige dicha predicción de acuerdo a las medidas que se obtienen con la matriz de cámaras de ionización. Este procedimiento se basa en un núcleo de corrección y permite

aumentar la resolución de la matriz de fluencia medida. En la figura 3.7 se aprecia un esquema del método utilizado para obtener la dosis 3D en la geometría TC del paciente.

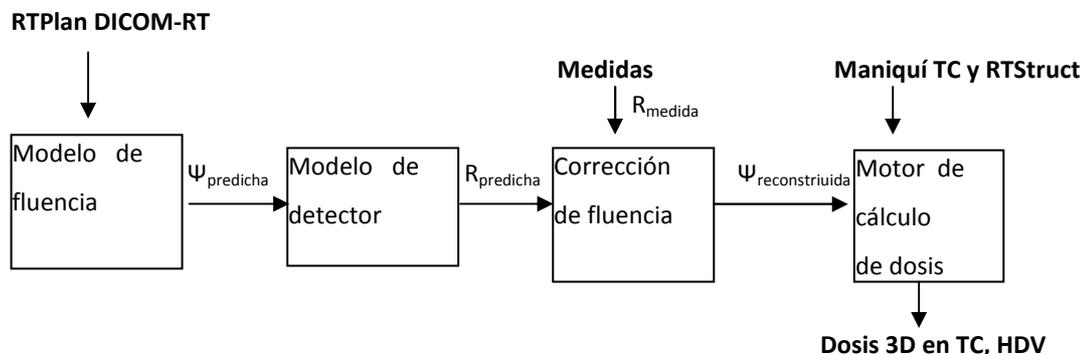


Figura 3.7.- Diagrama de flujo del algoritmo de obtención de la dosis reconstruida a partir de las medidas con el detector MatriXX. La entrada es el plan de tratamiento y las medidas (R_{medida}) se usan para corregir la fluencia predicha ($\Psi_{predicha}$) y generar una respuesta predicha ($R_{predicha}$) hasta generar una fluencia reconstruida ($\Psi_{reconstruida}$) con resolución suficiente para calcular en la TC del paciente. En negrita se resaltan las entradas y salidas del sistema.

Típicamente, la resolución de la fluencia predicha ($\Psi_{predicha}$), construida a partir del Plan DICOM RT es 2 mm x 2 mm. La clave del proceso es que la fluencia reconstruida ($\Psi_{reconstruida}$) se calcula a partir de la diferencia de la respuesta ($R_{predicha}$) que debería tener la fluencia predicha ($\Psi_{predicha}$) en el detector y la respuesta que se ha medido (R_{medida}). Una vez que se calcula la fluencia reconstruida se lanza sobre la TC del paciente usando el algoritmo de cono colapsado y con una resolución típica que va de 2 mm x 2 mm x 2 mm a 4 mm x 4 mm x 4 mm.

Las propiedades dosimétricas del detector han sido descritas por Herzen et al. La función de respuesta del detector a un campo de 1 mm de anchura puede ser aproximada por una gaussiana con 5,8 mm de anchura a mitad de altura (50) (ver figura 3.8).

Las ecuaciones que generan la fluencia reconstruida son las siguientes:

$$R_{predicha} = \sum_{i=1}^{N_{CI}} f_i(x, y) \bullet \Psi_{predicha} \quad (\text{ec 3.2})$$

donde la respuesta predicha del detector se genera usando la inversa del núcleo de convolución $f_i(x,y) = k_{\psi_i}(x,y)^{-1}$, la fluencia predicha $\Psi_{predicha}$ y sumando la contribución para todas las cámaras.

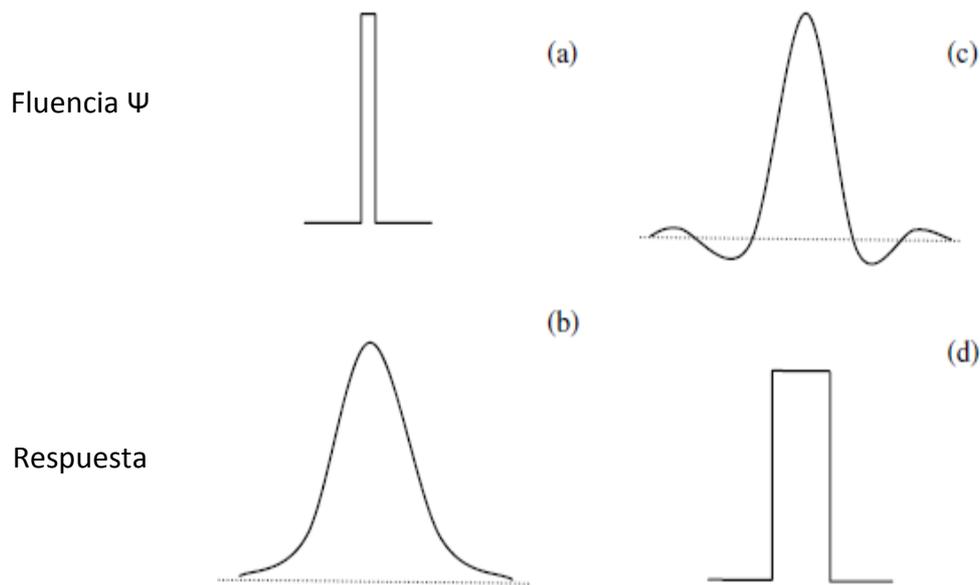


Figura 3.8.- Una fluencia elemental en un elemento de la matriz de fluencia (a) da una respuesta en el detector de acuerdo con la función de respuesta del detector f (b). El núcleo de corrección de la fluencia $k_{\psi_i}(x,y)$ (c) es la fluencia teórica que produce una señal exclusivamente en una sola cámara de ionización (d). En la práctica, el núcleo de corrección $k_{\psi_i}(x,y) \rightarrow$ (c) es proporcional a la diferencia en señal entre la respuesta predicha y la medida. Nótese que las partes negativas en el núcleo de corrección de fluencia son necesarias para cancelar las respuestas procedentes de las cámaras de ionización en la vecindad.

Usando las respuestas predicha y medida de cada cámara i se obtiene R_i

$$R_i = A_{escala} R_{i,predicha} - R_{i,medida} \quad (ec\ 3.3)$$

R_i es la denominada respuesta residual para la cámara i y se calcula utilizando un factor de escala A_{escala} que corrige por las posibles diferencias de amplitud de la señal en promedio entre todas las cámaras, es decir, tendría en cuenta posibles diferencias en las dosis/UM del acelerador o en las UM impartidas en el momento de la medida con respecto a las del plan verificado. El valor de A_{escala} lo optimiza Compass cuando compara las dos respuestas quedando situado en aquel que mejor acuerdo produce. A_{escala} se multiplica por la fluencia

predicha y al resultado se le aplica una perturbación de segundo orden a la fluencia por medio de la respuesta residual y el núcleo de convolución.

$$\Psi_{reconstruida} = A_{escala} \Psi_{predicha} + \sum_{i=1}^{N_{CI}} R_i \bullet k_{\Psi_i}(x, y) \quad (\text{ec 3.4})$$

De este modo se obtiene la fluencia reconstruida. Nótese que $f_i^{-1}(x, y) = k_{\Psi_i}(x, y)$ y que la convolución y deconvolución permiten el paso de matrices planares de resolución 2 mm x 2 mm a 7,62 mm x 7,62 mm que es la resolución del sistema de cámaras. Esta transformación matemática y el uso del núcleo de convolución es lo que permite pasar de una matriz de respuesta de resolución pobre, la medida con el detector MatriXX, a una fluencia reconstruida con suficiente resolución para el cálculo en el paciente.

El algoritmo de reconstrucción a partir de medidas de Compass se va a denominar en este trabajo con el acrónimo “CMPM”.

El sistema Compass, además, también puede calcular la distribución de dosis en un paciente a partir de sus propios datos de la unidad de tratamiento y el plan de tratamiento de IMRT obtenido con el SPT, como si de un sistema de planificación redundante se tratara. Por tanto es capaz de generar una fluencia esperada que puede ser comparada con la medida, y después calcular una matriz de dosis que permita comparar los HDV obtenidos a partir de ella tanto con los calculado en el SPT como con los reconstruidos a partir de la distribución de dosis medida.

El algoritmo de cálculo en Compass, a partir del plan del SPT, con una fluencia esperada se va a denominar en este trabajo “CMPC”. Tanto CMPM como CMPC utilizan un algoritmo parecido al de PCRT3D para calcular en el paciente, superposición de cono colapsado¹. La puesta en marcha de Compass incluye una sistemática parecida a la de un sistema de planificación. Para cada energía se introducen en el sistema a través de la utilidad “Beam Commissioning” porcentajes de dosis en profundidad (PDP), perfiles a diferentes profundidades y factores campo (figura 3.9). Como salida del ajuste de la unidad, Compass proporciona un espectro y una fluencia efectivos que permiten el cálculo de dosis con el

¹ Compass User’s Guide 3.0a. IBA Dosimetry

algoritmo de superposición de cono colapsado. La utilidad “Beam Commissing” contempla también la posibilidad de realizar modificaciones en el comportamiento del MLC en cuanto a tamaño de campo efectivo, ya que incorpora un ajuste automático de la posición de la lámina en función de los perfiles medidos en la dirección del colimador. Este ajuste se puede cambiar por uno manual si el usuario lo decide. Además, la definición del MLC contiene un parámetro que permite caracterizar el efecto de machihembrado. La transmisión interlámina no está modelada. Estos parámetros fueron ajustados de acuerdo con los datos experimentales recogidos.

La utilidad adicional del algoritmo CMPC es que, al ser comparado con los resultados de CMPM, permite establecer qué parte de las discrepancias entre CMPM y SPT se deben a la medida que se realiza cada día. Compass se ajusta mensualmente a la dosis por unidad de monitor y a la geometría de la máquina, realizándose la calibración del dispositivo de medida con esta periodicidad. Por tanto las medidas de los pacientes recogidas en este trabajo, pueden estar influenciadas por la estabilidad del acelerador en el día de la medida y por el comportamiento del MLC que, como se analizará en el apartado siguiente, tiene su relevancia en el resultado de la comparación.

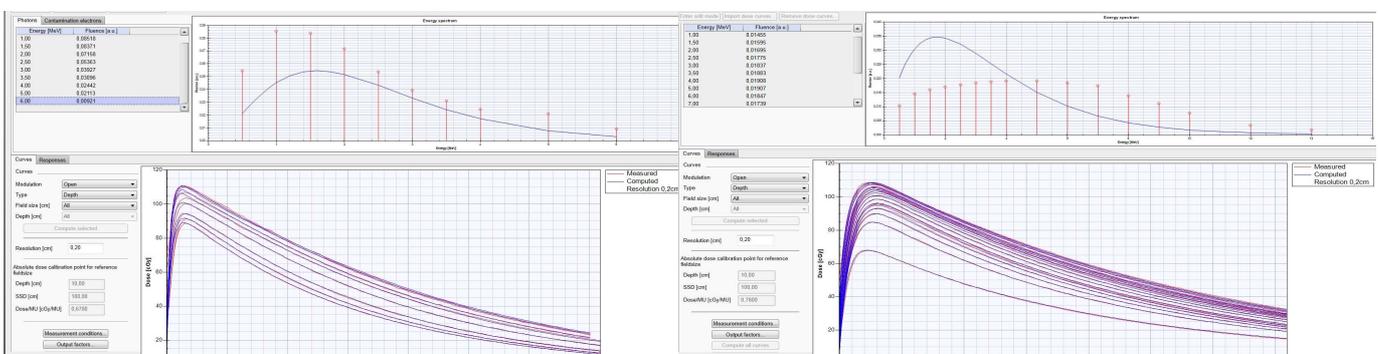


Figura 3.9.- Ajuste de PDP para 6MV (izquierda) y 15MV (derecha). Los factores campo generan la diferencia de altura entre los PDP. La comparación entre curvas medidas y calculadas permite comprobar si el ajuste es correcto.

3.4. Control de la influencia del MLC en las medidas

Para estudiar la variación de la posición de las láminas a lo largo del tiempo y periódicamente establecer la influencia del estado del MLC en la medida, se realiza un test sistemático de bandas contiguas inspirado en el informe nº72 de la AAPM (5), que consiste

en una irradiación de 10 bandas adyacentes de 2 cm x 28 cm, sin huecos entre ellas, con el mismo número de UM cada una. La imagen obtenida con el EPID se transforma a distribución de dosis planar con el mismo programa JD2D descrito en 3.1.2 y se analiza mediante el programa OmniPro IMRT de IBA Dosimetry (Figura 3.10). Se obtienen unos perfiles longitudinales, paralelos a las bandas irradiadas, precisamente en la zona de unión de éstas. De estos perfiles se hace un muestreo de forma que se obtiene el valor del perfil en la posición que ocupa el centro de la unión de cada una de las láminas. Posteriormente se calcula el valor medio y la desviación estándar de esos valores para de cada uno de los perfiles. El promedio en porcentaje de estos valores medios se utiliza como parámetro de caracterización del posicionamiento de las láminas del MLC. Se denominará a este parámetro promedio de picos del test bandas contiguas (PPBC). El valor de la fluencia se normaliza en la zona central del campo en el centro de la banda más cercana. Las características del colimador con el que se realizan los tratamientos de IMRT obligan a hacer un seguimiento continuado de su comportamiento. Por ejemplo, se ha observado una tendencia a cerrarse las láminas que se soluciona realizando una autocalibración semanal del MLC.

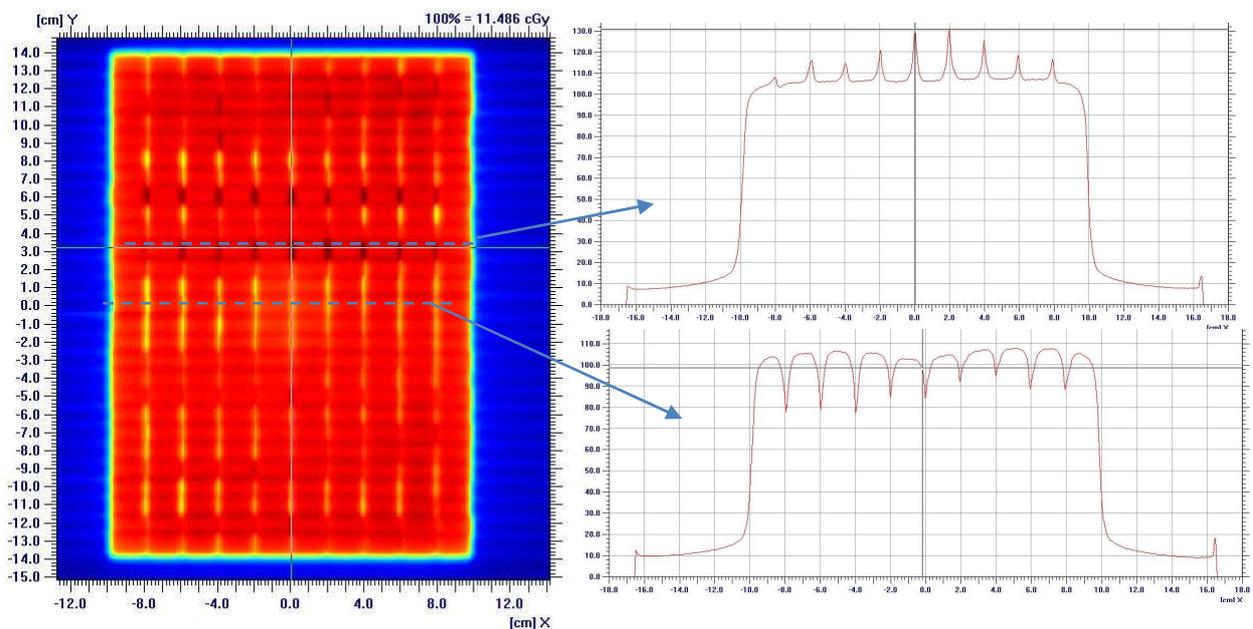


Figura 3.10.- Test bandas contiguas. De la imagen trasformada a dosis se extraen perfiles longitudinales a lo largo de la línea de unión de las bandas. En la imagen se muestran perfiles transversales para poder apreciar los picos y valles que serán evaluados.

En una primera fase se ha estudiado la reproducibilidad de dicha prueba tanto en una misma jornada como en días diferentes. Posteriormente se han provocado variaciones conocidas en la posición de las láminas con el fin de estimar la sensibilidad del test. En concreto se modificó la posición de todo el banco X1 de láminas +2mm y -1mm y se estudió la variación de PPBC en estas situaciones. También se ha verificado la reproducibilidad de los resultados así como la variación de PPBC al variar el ángulo de giro del brazo. Para ello, se ha irradiado en las 4 orientaciones principales, 3 veces consecutivas, y este mismo proceso se ha repetido en días diferentes. Una vez establecida la sensibilidad del test y las variables que influyen se ha procedido a estudiar si existe una correlación entre el valor que proporciona y las métricas definidas a partir de los HDV obtenidos a partir de las distribuciones de dosis que se miden con el mismo estado de MLC. A partir de estos resultados se puede estimar cuál es la influencia del estado el MLC en los resultados que se obtienen al medir con los sistemas de verificación. En particular, se estudiará cómo influye PPBC en la diferencia en dosis media en el PTV principal que se encuentra al comparar los HDV que provienen de CPM y de SCC.

4. Algoritmo desarrollado para el cálculo en el sistema de planificación

Se ha implementado en el SPT el método de superposición de cono colapsado (SCC) siguiendo el trabajo inicial de Anhesjö (51). Se basa en la aproximación de **cono colapsado** para reducir el tiempo de cálculo con respecto a una superposición pura de núcleos de dispersión. El acrónimo SCC se usará a lo largo del presente trabajo exclusivamente para el algoritmo implementado en PCRT3D. Tal como está descrito, el método de superposición utiliza unos núcleos polienergéticos pregenerados para cada unidad de tratamiento que son escalados de acuerdo con las heterogeneidades del paciente y sumados adecuadamente para dar la dosis en cada vóxel. Estos núcleos son pesados por el TERMA, que es la energía total liberada por unidad de masa en cada punto por fotones primarios. De manera resumida, el método desarrollado consta de las siguientes características que vemos a continuación.

4.1. Desarrollo del algoritmo SCC

4.1.1. Obtención de espectros energéticos efectivos

Los espectros efectivos son calculados a partir de datos experimentales de las unidades de tratamiento (razones tejido-maniquí (TPR)) y magnitudes tabuladas (coeficientes de atenuación másicos para blanco y filtro, núcleos de dispersión monoenergéticos precalculados por MC, secciones eficaces de producción de radiación de frenado para el blanco) (52).

Se ajusta un espectro efectivo siguiendo el modelo dado por Anhesjö, de modo que la fluencia energética Ψ_E se calcula según la siguiente expresión:

$$\Psi_E(E_0, t, f, Z) = E \cdot e^{-\mu_t f} \int_0^t \frac{d\sigma}{dE}(\bar{E}) \cdot e^{-\mu_f(t-x)} \cdot dx \quad (\text{ec 4.1})$$

donde t , f y Z son los parámetros del ajuste, esto es: t espesor del blanco de producción de rayos X del acelerador, f espesor del filtro aplanador y Z es el número atómico efectivo del blanco. μ_t y μ_f son los coeficientes de atenuación lineal para el filtro aplanador y para el

blanco respectivamente. $\frac{d\sigma}{dE}(\bar{E})$ es la sección eficaz de producción de radiación de frenado para la energía media de los electrones a cada profundidad dentro del blanco.

El método de ajuste es iterativo. El rango de energía considerado es desde 0,1 a 25 MeV. Este rango se divide en 25 intervalos. El espectro es calculado para cada uno de esos intervalos y las iteraciones tratan de minimizar la cantidad

$$\delta(\Psi_E(E_0, t, f, Z)) = \frac{1}{\sum_{k,j} \xi_{jk}} \sum_{k,j} \xi_{jk} \left| a_k \sum_i d_{ijk} \cdot \Psi_i - b_k \cdot TPR_{jk}(z_j, S_k) \right| \quad (\text{ec 4.2})$$

que es una diferencia en valor absoluto entre unos TPR calculados y los TPR experimentales. El cálculo de TPR se realiza a partir del espectro y de la dosis d_{ijk} por unidad de fluencia energética a una profundidad z_j en un campo paralelo de tamaño S_k . Los subíndices están dispuestos así: i se mueve en energías, j recorre profundidades y k recorre campos. Esto quiere decir que d_{ijk} expresa la dosis en un campo S_k para una profundidad z_j dentro de un haz monoenergético de energía correspondiente al intervalo i . $\xi_{j,k}$ anula las profundidades a las cuales hay contaminación electrónica, que, al estar presente en los datos experimentales no puede utilizarse para la optimización dado que el cálculo de los elementos d_{ijk} presupone sólo dosis debida a fotones incidentes. Esto permite luego extraer la contaminación electrónica presente en cada campo. a_k y b_k son valores de normalización que se fijan para que el valor calculado y experimental coincidan para cada campo a la profundidad de 20 cm.

Para minimizar la cantidad de la ecuación 4.2 se utiliza un método simplex (53)(algoritmo de optimización no lineal) donde t, f y Z efectivas son las salidas del mismo y como datos de entrada se utilizan valores cercanos a los típicos de las mismas.

El cálculo de los elementos d_{ijk} se realiza de acuerdo con Ahnesjö y Andreo (54) utilizando núcleos monoenergéticos de deposición de energía que han sido calculados previamente para cada intervalo de energía de acuerdo usando el método Monte Carlo, obligando a un fotón a interactuar un gran número de veces en un mismo punto y contabilizando la deposición de energía (55,56) alrededor del punto de interacción. En este caso el código utilizado para generar los núcleos de dispersión puntuales fue PENELOPE.

Una vez obtenido el espectro efectivo por el método iterativo descrito, se suman, pesados por él los núcleos monoenergéticos para obtener el polienergético. Dicho núcleo

polienergético $k_{poli}(r,\theta)$ se ajusta, creándose 48 biexponenciales, una por cada dirección polar θ , de acuerdo con la ecuación:

$$k_{poli}(r,\theta) = \frac{A_{\theta} \cdot e^{-a_{\theta}r} + B_{\theta} \cdot e^{-b_{\theta}r}}{r^2} \quad (\text{ec 4.3})$$

siendo A_{θ} , a_{θ} , B_{θ} y b_{θ} los parámetros resultantes del ajuste de cada biexponencial, de tal modo que una unidad de tratamiento queda determinada a partir de 48x4 números que definen su núcleo de superposición polienergético, A.

4.1.2. Contaminación electrónica.

Se modela y calcula de forma similar al modelo de Ahnesjö (57). El sistema calcula a partir de los datos del espectro los TPR para cada uno de los campos y los compara con los TPR medidos de tal forma que la diferencia es la contaminación electrónica que es función del tamaño de campo equivalente, de la energía y del diseño del cabezal del ALE. Esta contaminación será añadida al cálculo final de dosis, y está normalizada con respecto a la dosis calculada en condiciones de calibración. Para cada tamaño de campo A de la unidad se ajusta una función dependiente de la profundidad z, de modo que la dosis debida a la contaminación electrónica tiene la siguiente forma:

$$D_{contaminacion}(A, z) = k \cdot \alpha \cdot e^{-\beta \cdot z} \frac{\pi}{\gamma} \text{Erf}^2\left(\sqrt{\gamma} \frac{A}{2}\right) \quad (\text{ec 4.4})$$

donde k es la constante de normalización y α , β y γ son las constantes del ajuste, que no dependen del campo.

4.1.3. Cálculo del TERMA

Inicialmente se calcula el TERMA en cada punto del paciente por la que irán multiplicadas las contribuciones de dosis en el cálculo de superposición. Para calcular el TERMA es necesario conocer a partir de datos experimentales la distribución de fluencia energética lateral (ver 4.1.6). Se asumirá que esta distribución viene dada relativamente por el perfil de intensidad calculado según un método de convolución de “single pencil beam” (SPB) a partir de perfiles experimentales para diferentes tamaños de campo (58,59) multiplicado por un factor geométrico para tener en cuenta la anchura de la fuente y por las

funciones de modulación correspondientes a las cuñas, bloques o multilamina, en el caso de que los hubiere, como veremos más adelante.

4.1.4. Cono colapsado

El método de superposición exige la suma de las dispersiones producidas en todos los puntos del paciente teniendo en cuenta las densidades electrónicas que están presentes entre el punto de dispersión y el de cálculo, para todos los puntos de la matriz de cálculo. Para simplificar este cálculo se utilizará la aproximación de cono colapsado que realiza el cálculo en un cono de vóxeles a partir de un elemento dispersor suponiendo todo el cono colapsado sobre su eje. La ecuación final del cálculo de superposición de cono colapsado para determinar la dosis en el punto caracterizado por un vector de posición \vec{r} es:

$$D(\vec{r}) = \frac{\eta(\vec{r})}{\rho(\vec{r})} \frac{1}{d^2 u} \sum_m \sum_n [a_m R_{mn}^p(\vec{r}) + b_m R_{mn}^s(\vec{r})] \quad (\text{ec 4.5})$$

donde la suma en m y n se extiende sobre todas las direcciones discretas de energía radiante. m corre en ángulo polar theta y n en el ángulo azimutal phi. $\eta(\vec{r})$ es el número de electrones por unidad de volumen en el punto caracterizado por \vec{r} , que depende del nºCT del vóxel, $\rho(\vec{r})$ es la densidad másica en \vec{r} y también depende del nºCT del vóxel. Los dos sumandos entre corchetes corresponden a la energía radiante primaria:

$$R_{mn}^p(r_i) = R_{mn}^p(r_{i-1}) \cdot e^{-a_m \eta_i \Delta r} + T_i \rho_i \Omega_{mn} d^2 u \frac{A_m}{\eta_i a_m^2} (1 - e^{-a_m \eta_i \Delta r}) \quad (\text{ec 4.6})$$

y a la energía radiante secundaria

$$R_{mn}^s(r_i) = R_{mn}^s(r_{i-1}) \cdot (1 - b_m \eta_i \Delta r) + T_i \rho_i \Omega_{mn} d^2 u \frac{B_m}{b_m} \Delta r \quad (\text{ec 4.7})$$

Estas expresiones indican que a lo largo de cada eje de cono (rayo) se pueden calcular las dos energías radiantes de manera recursiva.

A_m, B_m, a_m, b_m son las constantes que provienen de ajustar los núcleos polienergéticos a una biexponencial de acuerdo con la ecuación 4.3. Los elementos T_i corresponden al terma polienergético de acuerdo con

$$T_i = \frac{r^2_0}{r^2} \frac{\mu(\vec{r})}{\rho(\vec{r})} \Psi(\vec{r}_0) e^{-\int_{r_0}^r \mu(l) dl} \quad (\text{ec 4.8})$$

Cada elemento sub- i se ha precalculado para cada vóxel teniendo en cuenta la fluencia energética $\Psi(\vec{r}_0)$ de referencia para cada rayo y el coeficiente de atenuación másico $\frac{\mu(\vec{r})}{\rho(\vec{r})}$ en el vóxel e integrando el coeficiente de atenuación a lo largo del rayo que llega al punto en cuestión para obtener un factor de atenuación promedio.

Las líneas de cono colapsado se disponen en el espacio de tal manera que cada vóxel sea atravesado solo una vez en cada dirección m, n . Para una red de conos optimizada se recorren todos los puntos generadores de conos depositando la energía dispersada en ellos de acuerdo con el método recursivo a lo largo del rayo descrito. El número de ángulos polares por defecto es 48 (entre 0 y 180°) y el número de ángulos azimutales (en la dirección en la que existe simetría del núcleo) es 10, por lo que se generan inicialmente 480 conos por punto. La dosis puede ser calculada en todos los puntos usando un número de operaciones proporcional a $M \cdot N^3$ donde M es el número de conos para cada punto y N es el número de vóxeles³.

4.1.5. Factor geométrico

El cálculo de superposición se realiza para una fuente puntual. Para tener en cuenta la anchura de la fuente introducida en la unidad de tratamiento el perfil de intensidad se multiplica por una función de la forma

$$F(x, y, z) = \frac{1}{1 + e^{\frac{k \cdot d_{\text{borde}}(x, y, z)}{2 \cdot \sigma_{\text{eff}}(z)}}} \quad (\text{ec 4.9})$$

donde $d_{\text{borde}}(x,y,z)$ es la distancia al elemento más cercano de colimación proyectado a la profundidad del punto x,y,z y $\sigma_{\text{eff}}(z)$ es el radio de la fuente proyectado por ese elemento a la profundidad z . La distancia d_{borde} es negativa cuando el punto está dentro del haz útil y positiva cuando el punto está fuera del mismo. k es un factor positivo de ajuste.

4.1.6. Determinación de la fluencia energética de cada unidad de tratamiento e introducción de moduladores del haz

La fluencia energética relativa viene modelada por el perfil de intensidad calculado para el algoritmo SPB desarrollado paralelamente. Esta fluencia que luego va a incidir sobre el paciente se modula adecuadamente para tener en cuenta el efecto de cuñas físicas, dinámicas, MLC, bloques etc..., usando la calculada en el algoritmo de convolución cuando se genera una unidad². Unida al factor geométrico descrito anteriormente modula el TERMA que se utiliza en el cálculo de superposición. Para deconvolucionar el perfil de intensidad se utiliza un "pencil beam" construido a partir de datos experimentales siguiendo el método desarrollado por Storchi (58–60). De acuerdo con la ecuación

$$P_{\text{int}} = \frac{1}{\|K_{\text{SPB}}\|} P_e \otimes^{-1} K_{\text{SPB}} \quad (\text{ec 4.10})$$

donde K_{SPB} es el núcleo que representa al pencil beam y P_e es el perfil envolvente calculado a partir de perfiles medidos en agua para diferentes tamaños de campo y profundidades. Como resultado se obtiene una matriz de fluencia, normalizada en el centro de campo e independiente del paciente. P_{int} se identifica con $\Psi(\vec{r}_0)$ en la ecuación 4.8.

En el algoritmo SSC el MLC y las mordazas se añaden multiplicando el perfil de intensidad P_{int} por valores que tienen en cuenta determinados aspectos del mismo:

- Transmisión del MLC y trasmisión de las mordazas, dada por $e^{-\mu_{\text{bl}} \cdot x}$ donde x es el espesor de la lámina o de la mordaza y μ_{bl} es el coeficiente de atenuación del material correspondiente.

² Manual de PCRT3D. Técnicas Radiofísicas SL.

- Segmentos de IMRT. Cada punto de P_{int} se ve afectado por segmentos abiertos y cerrados, ya que la suma de todos los segmentos compone la fluencia total para una orientación dada. Cada elemento del colimador es modificado de acuerdo con el número de segmentos que lo tienen abierto. Además se utiliza el factor de colimación de cada segmento para componer la suma.

De este modo el valor de T_i que se encuentra en las ecuaciones 4.6 y 4.7 se ve modificado por P_{int} multiplicado a su vez por los factores mencionados en función de que el punto de cálculo se encuentre dentro o fuera de campo o debajo de bloque o bien exista cuña dinámica o física.

4.1.7. Tratamiento de heterogeneidades

Como se ha mencionado anteriormente el algoritmo SCC requiere que cada vóxel venga caracterizado no sólo por la densidad electrónica, sino por un coeficiente de atenuación μ y una densidad másica ρ . A partir de valores publicados (NIST)³ se ha construido una tabla de tejidos presentes en el cuerpo asociándoles su densidad electrónica y con ello los otros dos valores que caracterizan el vóxel, siendo μ construido para cada espectro efectivo ajustado a partir de una suma ponderada usando los intervalos de energía (descritos más arriba) cuando se genera la unidad de tratamiento y para cada densidad electrónica. Durante el cálculo, el algoritmo explora los vóxeles del paciente, lee la densidad electrónica, asigna un coeficiente de atenuación lineal de acuerdo con la relación precalculada para cada unidad de tratamiento y asocia la densidad másica correspondiente. Los materiales extraídos de NIST, con intención de cubrir todo el rango posible, han sido: aire, pulmón, tejido adiposo, agua, músculo, hueso cortical, titanio, hierro y tungsteno. Los valores del coeficiente de atenuación son interpolados para todas las densidades electrónicas intermedias y guardados como un dato más de la unidad de tratamiento. Los valores de densidades asociadas a los materiales escogidos se encuentran en la tabla 4.I, así como los coeficientes de atenuación másicos correspondientes a dos espectros de 6 y 18MV obtenidos de acuerdo con la ecuación 4.1.

³ Hubbell JH and Seltzer SM. Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients. Disponibles on-line en: <http://physics.nist.gov/PhysRefData/XrayMassCoef/cover.html>

Tabla 4.I Materiales de referencia, densidades y coeficientes de atenuación que se asignan a los vóxeles del paciente

MATERIAL	DENSIDAD	DENSIDAD	$\mu/\rho(\text{cm}^2/\text{g})$	$\mu/\rho(\text{cm}^2/\text{g})$
	(g/cm^3)	ELECTRÓNICA	6MV	18MV
		(nº elec/ cm^3)		
Aire	0,001205	$3,623 \cdot 10^{20}$	0,0810	0,0572
Pulmón	0,258	$8,554 \cdot 10^{22}$	0,0438	0,0287
Tejido	0,95	$3,8 \cdot 10^{23}$	0,0462	0,0299
Agua	1	$3,343 \cdot 10^{23}$	0,0465	0,0305
Músculo	1,05	$3,478 \cdot 10^{23}$	0,0462	0,0302
Hueso	1,92	$5,95 \cdot 10^{23}$	0,0459	0,0314
Titanio	4,54	$2,26 \cdot 10^{24}$	0,0479	0,0366
Hierro	7,874	$3,98 \cdot 10^{24}$	0,0572	0,0451
Tungsten	19,3	$8,43 \cdot 10^{24}$	0,0642	0,0562

4.1.8. Cálculo de UM

La dosis por UM en el punto de normalización dentro de un paciente dado por el vector \vec{r} ,

$\frac{D(\vec{r})}{UM}$, se calcula mediante la expresión:

$$\frac{D(\vec{r})}{UM} = \frac{\Psi_0}{UM_0} d(\vec{r}) \quad (4.11)$$

donde

$$\frac{\Psi_0}{UM_0} = \frac{(D/UM)_{meas}^{calib}}{d(\vec{r}_{calib})} \quad (4.12)$$

es la fluencia por unidad de monitor (UM_0) en las condiciones de referencia, extraída como cociente entre la dosis de referencia de la unidad de tratamiento y la dosis calculada por el algoritmo SCC $d(\vec{r}_{calib})$ en la situación de referencia. De este modo toda la matriz de cálculo $d(\vec{r})$ queda expresada en unidades de dosis por unidad de monitor.

4.2. Verificación del algoritmo SCC

El algoritmo SCC fue verificado utilizando cálculos Monte Carlo y comparando las distribuciones de dosis obtenidas en maniqués complejos usando campos convencionales de radioterapia conformada (40). Para ello se ha utilizado el código BEAM, que es un

conjunto de programas para simular el transporte de radiación. Sus programas principales son:

- BEAMnrc (28) que simula el transporte de radiación (fotones, electrones y positrones) a través del cabezal de un ALE y proporciona fluencia de partículas en planos perpendiculares al eje del haz (ficheros de espacio de fase).
- DOSXYZnrc, que simula el transporte a través del paciente voxelizado y calcula la dosis en cada vóxel
- CTCREATE, que genera un paciente voxelizado a partir de las imágenes TC y asocia a cada vóxel un material y una densidad en función del número CT.

El código MC que utiliza todo el sistema BEAM es EGS4nrc (61). Este código está considerado internacionalmente como bien referenciado y su comportamiento en los problemas de radioterapia externa se considera como perfectamente validado en un gran número de publicaciones. Como paso previo es necesario que los parámetros de la simulación que corresponden a la caracterización de la unidad de tratamiento previa hayan sido ajustados para reproducir los datos experimentales. Esta validación experimental parte fundamentalmente de comprobar la coincidencia de PDP y perfiles de dosis medidos en un maniquí de agua con detectores adecuados. Cuando se validó el algoritmo SCC, las unidades de tratamiento que se emplearon fueron haces de 6 y 18 MV correspondientes a un acelerador Primus (Siemens).

El proceso de ajuste de las unidades de 6MV y 15MV del acelerador Oncor usadas en los planes de IMRT de este trabajo está descrito en los apartados 5.2.1 y 7.2. El mismo procedimiento se empleó para las dos energías de fotones del ALE Primus.

Gran parte de la mejora en la exactitud del cálculo con SCC frente a otro tipo de algoritmos menos sofisticados se centra en el cálculo en presencia de heterogeneidades. En los tratamientos de radioterapia dichas heterogeneidades son fundamentalmente pulmón y hueso. Por ello, una vez ajustadas la unidades, se pasó a comparar los resultados del cálculo realizado con los algoritmos SCC y MC en maniquíes específicos, con heterogeneidades en su interior. Los maniquíes empleados están en la figura 4.1. Las características de los materiales pueden encontrarse en las tabla 4.IIa y 4:IIb

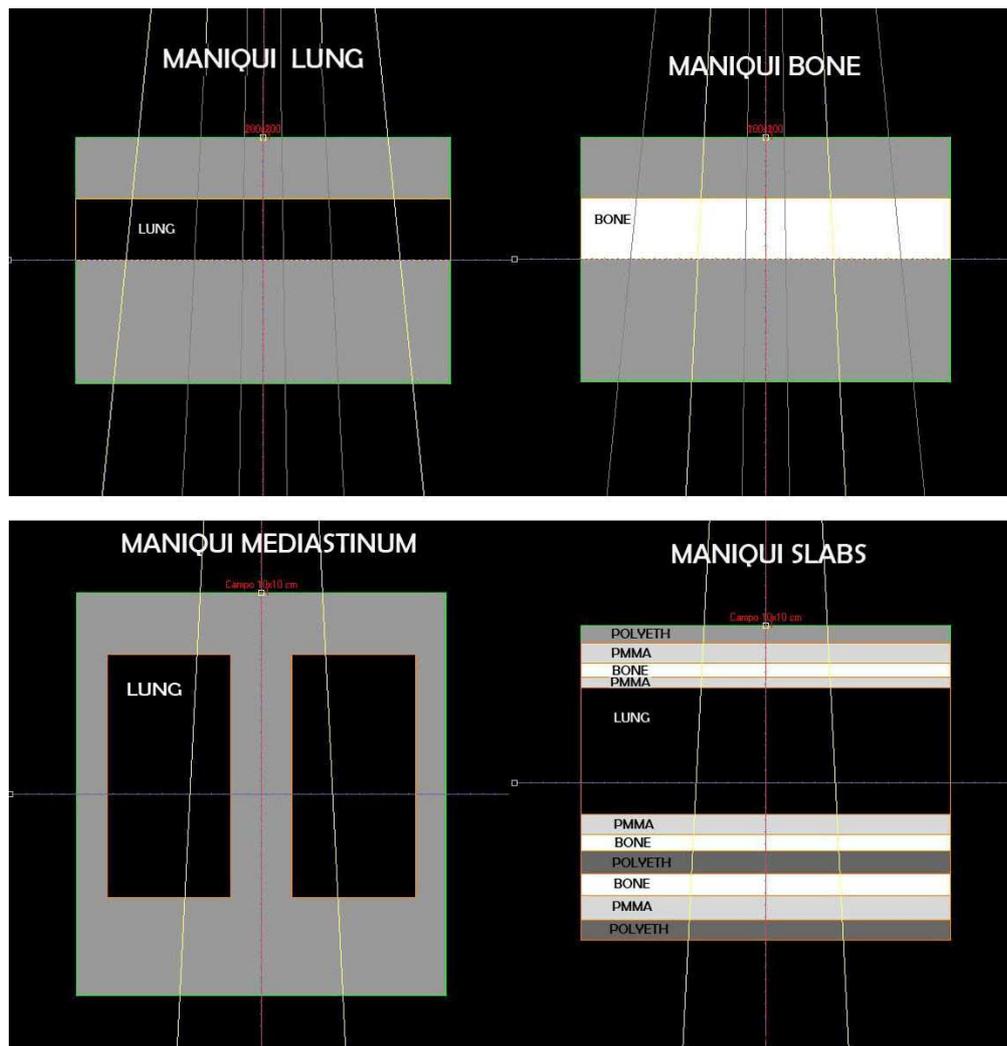


Figura 4.1. Maniqués para la comprobación de la exactitud del cálculo de superposición SCC en presencia de heterogeneidades. LUNG: simula una lámina de pulmón de 5cm de espesor dentro de un maniquí de agua. BONE: Láminas de hueso de 5 cm de espesor dentro de un maniquí de agua. MEDIASTINUM: Maniquí para calcular un campo anterior incidiendo sobre un tórax. SLABS: Maniquí para comprobar el modelado de los cambios a través de muchas láminas heterogéneas.

Previamente al cálculo en los maniqués heterogéneos de la figura 4.1 se simularon espacios de fase para 3 campos, 3 cm x3 cm, 10 cm x 10 cm y 20 cm x 20 cm, y para las dos energías mencionadas (6 y 18 MV), usando para calcularlos otros espacios de fase calculados antes de las mordazas. Estos espacios de fase se propagan con BEAMnrc a través de la disposición de mordazas y MLC propia de cada campo y se proyectan usando DOSXYZnrc contra los maniqués heterogéneos diseñados. Como técnica de reducción de la varianza se empleó la reutilización de historias, hasta conseguir una incertidumbre estadística del 1%

dentro de campo. Los parámetros de transporte seleccionados fueron los establecidos por defecto en el código, siendo las energías de corte para electrones y fotones ECUT=700 keV y PCUT=10 keV respectivamente.

Tablas 4.IIa y 4.IIb. Tamaños, materiales y densidades de los maniqués utilizados en la simulación MC y representados en la figura 4.1.

Tabla 4.IIa

MANIQUI	DIMENSIONES (cm)	MATERIALES DOSXYZNRC	CARACTERÍSTICAS HETEROGENEIDAD
LUNG	30x30x20	H2O700ICRU y LUNG700ICRU	Lámina de LUNG700ICRU de 30x30x5 cm
BONE	30x30x20	H2O700ICRU y ICRPBONE700ICRU	Lámina de ICRPBONE700ICRU de 30x30x5 cm
MEDIASTINUM	30x30x33	H2O700ICRU y LUNG700ICRU	Dos bloques de LUNG700ICRU de 10x30x20cm
SLABS	30x30x25,6	POLYETH700ICRU , PMMA700ICRU, ICRPBONE700, LUNG700ICRU	Espesores de arriba abajo: 1,4/1,6/1,2/0,8/10,4/1,6/1,4/1,8/1,8/2,0 y 1,6 cm de los materiales indicados en la figura 4.1.

Tabla 4.IIb

MATERIAL	DENSIDAD (g/cm ³)	DENSIDAD ELECTRÓNICA RELATIVA AL AGUA
H2O700ICRU	1,0	1,0
LUNG700ICRU	0,26	0,258
ICRPBONE700ICRU	1,85	1,74
POLYETH700ICRU	0,93	0,96
PMMA700ICRU	1,19	1,16

5. Cálculo Monte Carlo de distribuciones de dosis

Entre los diferentes métodos de cálculo existente, el cálculo por MC se considera el que suministra de forma más exacta una distribución de dosis absorbida en un medio irradiado, una vez validadas las unidades de tratamiento y si se reduce la incertidumbre estadística lo suficiente.

Se ha desarrollado una aplicación que permite el cálculo de la dosis absorbida en los pacientes estudiados en el presente trabajo, utilizando un ajuste de los parámetros de la simulación de las unidades de tratamiento que ha sido exhaustivamente verificado.

El código MC que utiliza todo el sistema, al igual que el descrito en el apartado anterior, es BEAM, a su vez basado en EGS4nrc (61). En este caso, al contrario que en 4.2 donde los cálculos se realizaron de manera manual y paso a paso, ha sido necesario generar una utilidad específica automatizada que permita la gestión de pacientes y el manejo de haces de IMRT. Además ha sido necesario crear un “component module” específico para el colimador MLC Optifocus, ya que su geometría no está incluida entre las disponibles en la instalación estándar de BEAMnrc.

El nombre específico de la aplicación desarrollada es **MCVerif** , y realiza el cálculo lanzando un espacio de fase, previamente generado con BEAMnrc, justamente debajo de la cámara monitora a través de las mordazas y el colimador multilámina, generándose un espacio de fase (usando de nuevo BEAMnrc) tras la retícula, listo para ser lanzado sobre el paciente. MCVerif transforma el resultado obtenido con el cálculo de DOSXYZnrc (archivos 3ddose) en un fichero con extensión .d3d que puede ser importado directamente en el sistema de planificación PCRT3D para visualizar la dosis.

Los ficheros 3ddose de DOSXYZnrc contienen la dosis absorbida en cada vóxel del paciente y su incertidumbre, en un formato ASCII fácilmente editable. Podemos ver un esquema general en la figura 5.1.

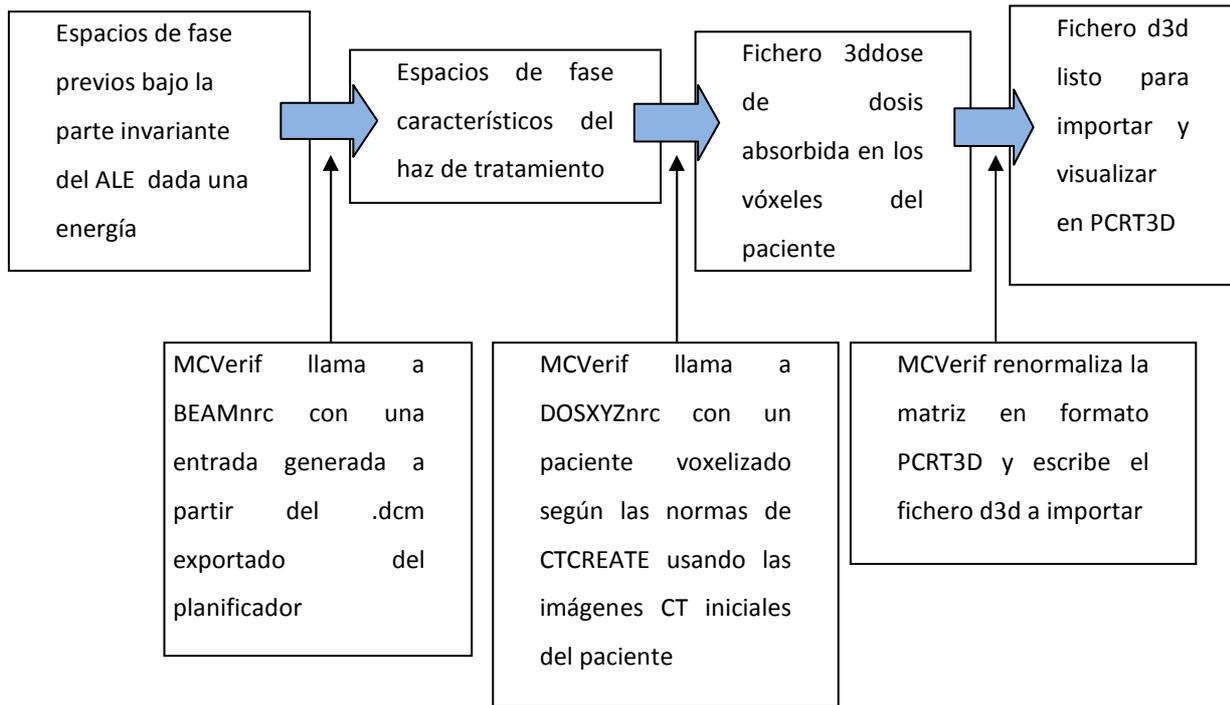


Figura 5.1.- Diagrama de flujo de MCVerif

5.1. Particularidades de adaptación a BEAMnrc y DOSXYZnrc y utilidades de MCVerif

Por motivos de manejo del sistema BEAM, la adaptación a las necesidades de cálculo ha requerido intervención sobre el código de BEAM, de DOSXYZnrc, y de la rutina CTCREATE. Además ha sido necesario modificar y/o agregar los ficheros de espacio de fase intermedios con objeto de tener en cuenta la posible presencia de cuñas dinámicas, campos segmentados en IMRT o bien de obtener una incertidumbre estadística suficientemente baja en los cálculos en un maniquí.

Se han encontrado errores al utilizar ficheros con tamaño superior a 2GB, tamaño bastante habitual bajo las cámaras monitoras. Debido al este problema, el número de partículas máximo en un espacio de fase debajo de las cámaras monitoras está en torno a 70 millones. MCVerif utiliza hasta 8 ficheros de este tipo generados en el proceso de ajuste de la unidad para obtener suficiente incertidumbre estadística. El cálculo responde al esquema de la figura 5.2.

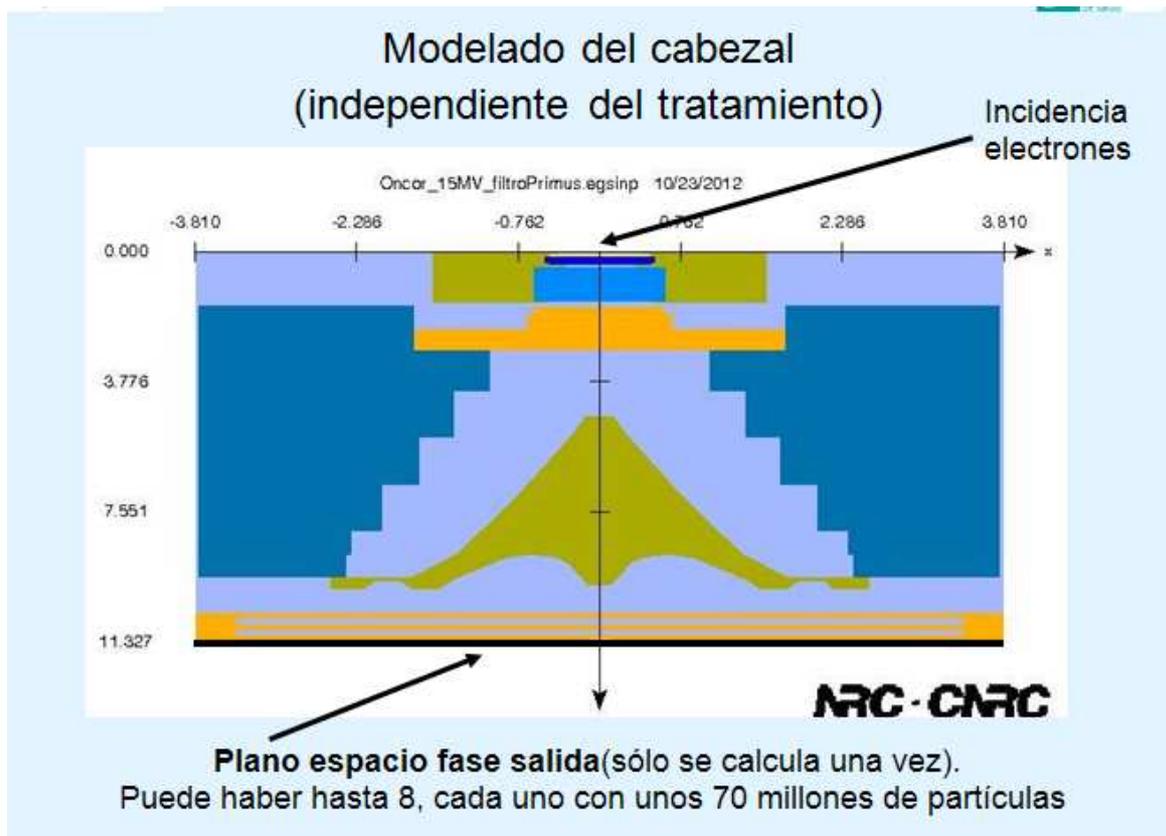


Figura 5.2.- Esquema del cabezal de la unidad de tratamiento y del plano de cálculo del espacio de fase inicial.

Cada uno de estos ficheros se reutiliza con BEAM sobre la geometría de mordazas y MLC (módulo SMLC) para dar lugar a un segundo fichero que ya contiene la conformación del haz sobre el paciente. Este segundo fichero queda con unos pocos millones de partículas, dependiendo del tamaño de campo (4 millones es una referencia). El módulo SMLC ha sido desarrollado en este trabajo para reproducir fielmente las condiciones de transmisión interlámina y de machihembrado (“tongue and groove”) del colimador Optifocus (39). Podemos encontrar una descripción del proceso de ajuste de la unidad de tratamiento en el apartado 5.2.1.

En el caso que nos ocupa, IMRT estática, MCVerif descompone los espacios iniciales bajo las cámaras monitoras en tantos sub-espacios como número de segmentos tiene cada haz de tratamiento, asignándoles tantas partículas como peso relativo tenga el segmento. Cada uno de esos sub-espacios se hace pasar por las mordazas y el MLC. Como resultado, en IMRT se obtienen N sub-espacios en la retícula, siendo N el número de segmentos total del

tratamiento. Cada uno de ellos va a ser calculado independientemente en el paciente generado, tal y como se describe a continuación.

MCVerif llama posteriormente a DOSXYZnrc para realizar el cálculo de dosis sobre las imágenes del paciente. Estas imágenes han sido procesadas previamente para generar un fichero egsphant, usando la utilidad CTCREATE, programa que está incluido en la instalación de BEAM pero que ha sido ampliamente modificado en este trabajo para adaptarlo a nuestra situación. Estas modificaciones han incluido la alteración de la función que asigna densidad y material, para adecuarla a las necesidades de los ficheros .dcm de imagen exportados de PCRT3D y también la aceptación de esas imágenes cuando no son cuadradas (por ejemplo cuando incluyen la mesa de tratamiento del acelerador ONCOR → 517x512 píxeles). El fichero egsphant contiene en formato ASCII las coordenadas de los vóxeles del maniquí, el material asignado a cada uno (aire, pulmón, agua o hueso), y las densidades correspondientes a cada vóxel.

MCVerif también lee el fichero con extensión .dcm (RT-Plan de IMRT), exportado del planificador PCRT3D, para poder configurar convenientemente el módulo SMLC de BEAMnrc y para tener la geometría del espacio de fase con respecto al maniquí egsphant (posición del isocentro, ángulo de brazo, ángulo de colimador, ángulo de mesa). En SMLC es posible realizar el ajuste de la trasmisión interlámina con la selección del parámetro de separación entre láminas, típicamente 0,1 mm.

Antes de calcular la dosis absorbida en un maniquí creado usando la información de sus imágenes TC, a partir de un fichero de espacio de fase, es necesario adaptar la entrada de DOSXYZnrc. El principal problema son las coordenadas de incidencia sobre el paciente. Si se mantienen las coordenadas de isocentro de PCRT3D entonces los ángulos del haz que se usan en la entrada de DOSXYZNR, θ , ϕ y col , han de ser fijados como sigue, partiendo del ángulo de mesa (am), ángulo de brazo (ag) y ángulo de colimador (ac) en radianes del haz exportado desde PCRT3D:

$$\theta = \frac{180}{\pi} \cdot a \cos(\text{sen}(am) \cdot \text{sen}(ag)) \quad \phi = \frac{180}{\pi} \cdot a \tan 2(-\cos(ag), \cos am) \cdot \text{sen}(ag))$$

$$col = \frac{180}{\pi} \cdot a \tan 2(\cos(am), \text{sen}(am) \cdot \cos(ag)) + 180 - ac$$

Un esquema sobre cómo se relacionan los ángulos con el paciente para preparar una entrada de un espacio de fase en DOSXYZnrc puede verse en el manual de DOSXYZnrc (62).

MCVerif fija el número de partículas adecuado para el cálculo teniendo en cuenta las que han atravesado el sistema de colimación. La incertidumbre estadística final obtenida en la dosis calculada en el paciente dependerá del tamaño de vóxel elegido al crear el maniquí y del número de partículas que se reutilicen. Como ejemplo, para un retícula de 0,4 cm x 0,4 cm x 0,5 cm y un campo 10 cm x 10 cm sobre una paciente de mama son necesarias 1200 millones de historias en DOSXYZNRC para obtener una incertidumbre estadística del 2% usando un solo espacio de fase de 2GB. Se calcula la dosis absorbida en cada vóxel del maniquí reutilizando partículas del espacio de fase con un factor que por defecto es 100 y que permite aumentar la eficiencia de la simulación. No se recurre inicialmente a la técnica de reducción de la varianza denominada “photon splitting” (28), aunque dicha técnica puede ser activada si se desea. También se puede seleccionar el número de espacios de fase tras las cámaras monitoras que se van a usar, cuanto más grande se escoge este valor mejor es la estadística del cálculo. Si por ejemplo el procesador tiene 8 núcleos, usar 8 espacios de fase iniciales no implica más tiempo de cálculo ya que cada cálculo se manda a un núcleo distinto. MCVerif tiene en cuenta cuántas partículas hay en los espacios de fase iniciales, por lo que usa ese número de partículas en la simulación a través del sistema de colimación. El número de partículas de los espacios de fase iniciales se cambia editando el código del programa, por lo que si se quiere partir de otros espacios de fase distintos es necesario recompilar el programa.

Finalmente se transforma el fichero con extensión .3ddose que genera DOSXYZnrc en un fichero con extensión .d3d que puede leer PCRT3D para importar la matriz de dosis de cada campo en el planificador. En el caso de IMRT, las matrices generadas por segmento (y por espacio de fase inicial) se suman (respetando el cálculo de incertidumbres) para dar lugar a una matriz única por haz. La dosis en estos ficheros está normalizada al valor que se obtiene en el máximo de un campo 10 cm x 10 cm de referencia, a la profundidad del máximo para la energía seleccionada, de tal manera que, al importar en PCRT3D, la representación de la dosis es la habitual, y el valor de normalización de cada haz conduce directamente a un cálculo válido de unidades de monitor. El hecho de poder importar los resultados de MCVerif

en el TPS es una gran ventaja ya que permite usar todas las herramientas de análisis y exportación de datos del sistema de planificación.

Para comparar la dosis calculada por MC con otro plan de IMRT calculado por el algoritmo SCC, en el plan MC importado en PCRT3D **se cambia la dosis de normalización hasta que las UM se igualen en ambos planes** (rara vez coinciden exactamente). De este modo, la comparación en dosis es pertinente. Las incertidumbres del cálculo están en el fichero con extensión .3ddose y han de ser tenidas muy en cuenta a la hora de evaluar los resultados del cálculo MC. La incertidumbre estimada ($k=1$) para la dosis absorbida calculada por Monte Carlo en los casos planteados es inferior al 2% de la dosis prescrita para el PTV principal en cada caso.

MCVerif incorpora además algunas utilidades que permiten manejar y procesar ficheros con extensión .3ddose y ficheros de espacio de fase para obtener una salida modificada de manera conveniente, es posible desplazar matrices de dosis (contenidas en el tipo .3ddose), renormalizar la matriz de dosis a un valor arbitrario introducido y sumar estas matrices, siguiendo las reglas adecuadas para las incertidumbres. También permite agregar varios ficheros de espacio de fase para obtener uno con más partículas.

5.2. Desarrollo del “component module” SMLC y ajuste de la unidad de tratamiento

En primer lugar es necesario que la unidad de tratamiento esté correctamente ajustada, esto es, que los parámetros de los que depende la simulación con BEAMnrc permitan reproducir medidas experimentales de referencia. Por otro lado, para realizar cálculos MC de campos de IMRT se necesita una descripción suficientemente detallada del MLC, particularmente en lo que se refiere al diseño del machihembrado.

5.2.1. Ajuste de la unidad de tratamiento

Para modelar la unidad de tratamiento a utilizar en la aplicación MCVerif, se simuló un conjunto de campos cuadrados abiertos y se calcularon sus PDP y perfiles de dosis a varias profundidades, todo ello con objeto de ajustar los parámetros que intervienen en el diseño de la fuente de electrones. Para fotones de 6MV y 15MV se usaron módulos de BEAMnrc,

configurados con los datos del fabricante, para definir el blanco, el colimador primario, el filtro aplanador, dos cámaras monitoras, un par de mordazas opuestas y el MLC. Para el MLC se utilizó el módulo VARMLC (ver 5.2.2), ya que en este paso del ajuste, sin campos de IMRT, la estructura de machihembrado del MLC no afecta al resultado. Para la fuente incidente de electrones en el blanco del ALE se eligió un haz monoenergético con una distribución espacial gaussiana y una pequeña dispersión angular del eje a la superficie del blanco. Los parámetros de la fuente son, por tanto, la energía media de los electrones, la anchura a mitad de altura (FWHM) de la distribución espacial y la dispersión angular. Los resultados publicados en la literatura indican que este tipo de fuente es suficiente para caracterizar un haz de 6MV (63,64); la dispersión energética y la excentricidad del punto focal no juegan un papel relevante en su modelado. La dispersión angular sí que es necesaria para reproducir los perfiles de campos grandes (65).

El ajuste se realizó con un método de ensayo y error usando los PDP y perfiles medidos para campos cuadrados durante la puesta en marcha del acelerador. Estas medidas fueron hechas en su día con un diodo no compensado (Scanditronix) hasta un tamaño de campo de 15 cm x 15 cm y con una cámara de ionización de 0,125 cm³ (PTW) para campos más grandes. La incertidumbre estadística de las simulaciones se fijó por debajo del 1% dentro de campo. Un espacio de fase pre-calculado de 560 millones de fotones (tanto para 6 como para 15MV), subdividido en 8 ficheros de 70 millones de partículas cada uno, fue empleado de nuevo como entrada de BEAM por encima de las mordazas y el MLC para calcular la dosis debida a cada campo dentro de un maniquí de agua voxelizado con 3 mm de resolución, a distancia fuente superficie fija y con un tamaño suficiente para cubrir el campo más grande. Para obtener los factores campo Monte Carlo, se utilizaron los valores brutos extraídos a 10 cm de profundidad para cada campo, siendo normalizados al valor del tamaño de campo 10 cm x 10 cm. La incertidumbre correspondiente se propagó convenientemente a través del cociente de dosis absorbida.

5.2.2. Desarrollo de SMLC

Ninguno de los módulos que ofrece BEAMnrc describe de manera apropiada las particularidades del colimador Optifocus usado en los tratamientos de IMRT. Existe un módulo denominado VARMLC (66), que describe una estructura de machihembrado pero

con todas las láminas geoméricamente similares: salientes siempre en el mismo lado e igualmente las ranuras, sin posibilidad de modelar una lámina central diferente con salientes a ambos lados o un par de láminas relacionadas por una reflexión tipo espejo. La razón para ello es que VARMLC se desarrolló para MLC de tipo VARIAN. Se denominará SMLC al nuevo módulo de BEAMnrc desarrollado para simular el colimador Optifocus.

En la figura 5.3 se puede apreciar la estructura de machihembrado de SMLC y de VARMLC.

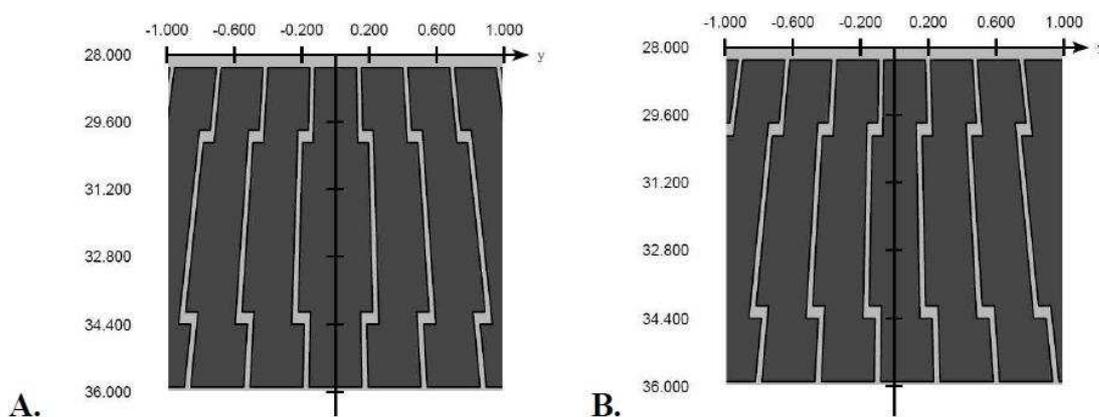


Figura 5.3. Diseños de machihembrado modelados para SMLC (A.) y VARMLC (B.).

El nuevo módulo se validó de dos maneras (39): primero, comprobando directamente sus propiedades geométricas y segundo comparando el resultado de las simulaciones de campos sensibles a la estructura de machihembrado con SMLC y con VARMLC. Después se ajustó el parámetro libre de SMLC, dado que las dimensiones, espesores y materiales del colimador son proporcionadas por el fabricante. Este parámetro es la anchura del hueco de aire que separa láminas adyacentes. Después del proceso que se describe en (39) la anchura quedó fijada en 0,1 mm. En particular, los resultados de una simulación para un campo específico (figura 5.4) pueden apreciarse en la figura 5.5. La simulación MC utilizó vóxeles de 3 mm en las direcciones X y Z y de 1 mm en la dirección Y con objeto de aumentar la resolución en la dirección de interés. Los espacios de fase fueron generados de acuerdo a lo descrito en 5.2.1. En la figura 5.5 se muestran los perfiles de dosis calculados por MC junto con los medidos con el EPID y placas radiocrómicas EBT2, para su comparación. La línea horizontal marca el promedio de valles para un espesor de aire de 0,1 mm, que resultó ser el hueco

ajustado finalmente dado que proporcionaba la mejor coincidencia con el promedio de valles experimental.

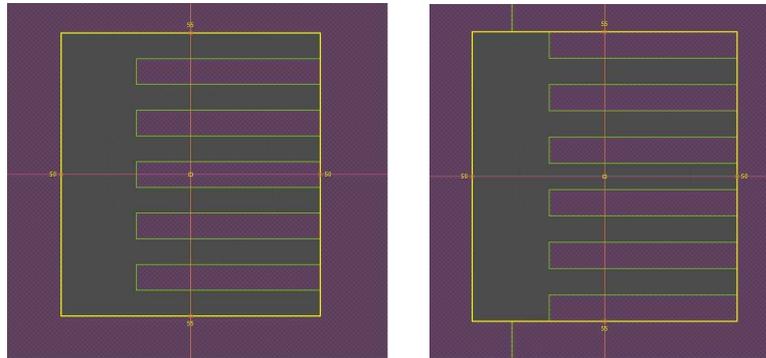


Figura 5.4 Configuración del MLC para comprobar el efecto de machihembrado. El campo de la izquierda y su complementario de la derecha se componen para evaluar el efecto mencionado.

Tras el ajuste del espesor de aire la coincidencia con las medidas experimentales resultó satisfactoria. Como conclusión, se obtuvo que el desarrollo de SMLC permitía simular en detalle el efecto de machihembrado del colimador Optifocus en los tratamientos de IMRT.

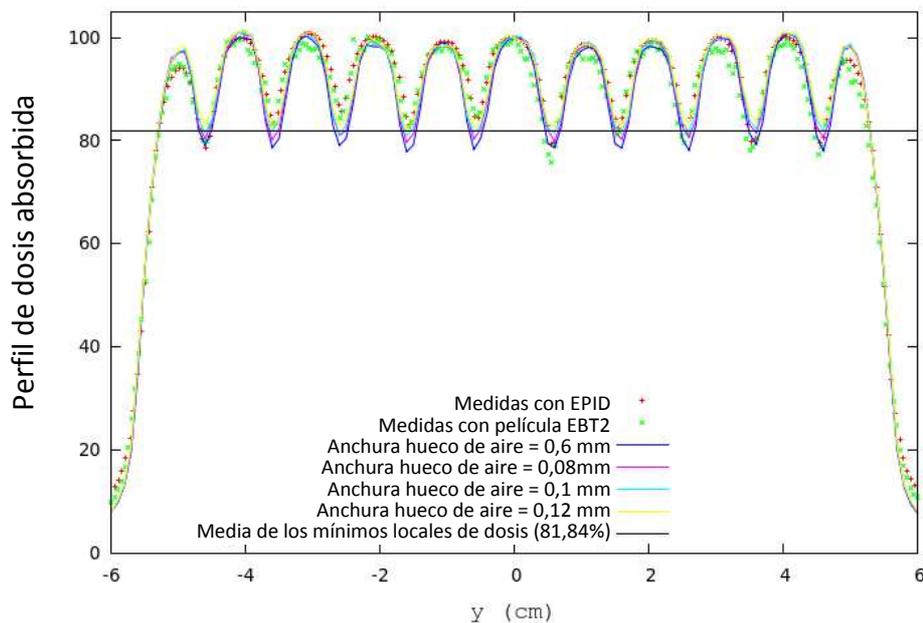


Figura 5.5 Perfiles de dosis calculados por MC para diferentes espesores de aire a lo largo del eje vertical del campo compuesto de la figura 5.4, para la energía de 6MV a una distancia fuente superficie de 98 cm y una profundidad de 2 cm. Se comparan los cálculos con medidas realizadas con EPID (puntos rojos) y con placa radiocrómica (puntos verdes).

6. Métricas a utilizar, sistemas clasificadores y herramientas de discriminación

Como se ha mencionado, en el presente trabajo se van a proponer una serie de métricas que permiten establecer si un plan de tratamiento es aceptable para su tratamiento o no. Es preciso clasificar y comparar los resultados obtenidos para cada métrica y utilizar las herramientas de las que dispongamos para poder hacerlo. En los siguientes puntos se describen las métricas, los sistemas clasificadores y las herramientas que se van a utilizar para compararlos. El proceso se va a estructurar en tres etapas. Primero se evaluará la influencia de las variaciones en dosis por UM de la unidad de tratamiento y la del colimador MLC en la medida de las dosis correspondientes a distintos planes. Después se procederá al análisis ROC de las métricas que se proponen (del inglés “receiver operating characteristic”, o curva característica operativa del receptor) con respecto a las métricas que se consideren como de referencia, estableciéndose los puntos de corte óptimos para la evaluación, y por último, se evaluarán las correlaciones existentes entre métricas si las hubiere utilizando para las métricas continuas los puntos de corte obtenidos anteriormente. Antes de poder utilizarlos se han de establecer una serie de presunciones sobre las que va a funcionar el análisis. En el punto 6.3. se resumirán las métricas a utilizar.

6.1. Punto de partida de la comparación entre HDV y del escalado de histogramas

Las siguientes premisas se suponen válidas para el ámbito del trabajo y constituyen la piedra angular del mismo:

- (a) El algoritmo de cálculo de dosis más exacto del que se dispone es el de Monte Carlo (MC). Se ha modelado particularmente el colimador Optifocus para tener en cuenta la fuga interlámina y el efecto de machihembrado (39). El cálculo se realiza a partir de un espacio de fase bajo las cámaras monitoras y la dosis correspondiente a cada segmento del tratamiento de IMRT es calculada de manera independiente usando el mismo número de UM que calcula el SPT usando el algoritmo SCC. Las otras formas de determinación de la dosis absorbida usadas en este trabajo, mediante los algoritmos CMPC,

CMPM y SCC, utilizan un modelado del espectro efectivo y de la fluencia incidente derivada a partir de datos experimentales. Por tanto MC constituye un modelo de cálculo independiente y validado que se va a considerar como el mejor algoritmo de cálculo del que se dispone.

(b) Las diferencias encontradas entre CMPM y CMPC, una vez validado el modelo en Compass, se deben esencialmente a las diferencias en la dosis por UM de la unidad de tratamiento y al comportamiento del MLC en el día de la medida. Cambia por tanto la forma de la fluencia entre la reconstruida a partir de las medidas y la predicha inicialmente, realizándose posteriormente la determinación de dosis en el paciente usando el mismo algoritmo.

(c) Las diferencias del punto anterior pueden ser utilizadas para corregir los histogramas diferenciales calculados con el algoritmo MC y adaptarlos a los que se obtendrían el mismo día de la medida con Compass. Esto permitiría tener un conjunto de HDV basados en MC pero que incorporan la dependencia con el estado del acelerador. Este conjunto puede ser evaluado para definir si un plan es aceptado o no, por medio de la medida del detrimento que se produce en los índices de calidad del tratamiento, basados en la cobertura de los PTVs con la dosis deseada, el aumento de la dosis media en los órganos en paralelo y de D1 en los órganos en serie. El método escogido para escalar los histogramas diferenciales MC se denominará escalado proporcional, y consiste en evaluar la diferencia en la dosis media del PTV principal ($\Delta\bar{D}_{PTVp}$) entre los algoritmos CMPM y CMPC, y transformar los valores de los puntos D_i de los histogramas diferenciales

obtenidos por MC de acuerdo con la expresión $\Delta D_i = D_i \cdot \frac{\Delta\bar{D}_{PTVp}}{D_{prescrita}}$ donde

$\Delta\bar{D}_{PTVp}$ está expresado en dosis absorbida y ΔD_i sería la variación que se aplica a los puntos del histograma. Este tipo de escalado hace que, para volúmenes con poca dosis, el incremento en la dosis neta sea menor que para PTVs. Para comprobar la validez del escalado de histogramas del algoritmo MC, se evaluará primero si los histogramas obtenidos por el algoritmo CMPC se

pueden transformarse en los de CPM sin más que evaluar la dosis media en el PTV principal y aplicar la expresión. El conjunto de histogramas que se obtienen al escalar los obtenidos por MC por este método se les denominará como “obtenidos por MCesc”.

6.2. Herramientas de evaluación

A continuación se describen las herramientas de evaluación que se van a utilizar para valorar los datos extraídos de los planes de tratamiento.

6.2.1. Gamma global 2D y 3D entre distribuciones de dosis.

La evaluación de la similitud entre matrices 3D, planos y perfiles que comparan datos experimentales y calculados o bien, en general, dos distribuciones de dosis, puede realizarse de acuerdo a un criterio distancia-diferencia de dosis absorbida. El objeto de no comparar simplemente la dosis e incluir la distancia es no penalizar en exceso las diferencias en zonas de alto gradiente de dosis. Si se establecen unas tolerancias para las diferencias en dosis absorbida y distancia (ej: 3% o 3 mm) queda definido un elipsoide en el espacio dosis absorbida-distancia (figura 6.1).

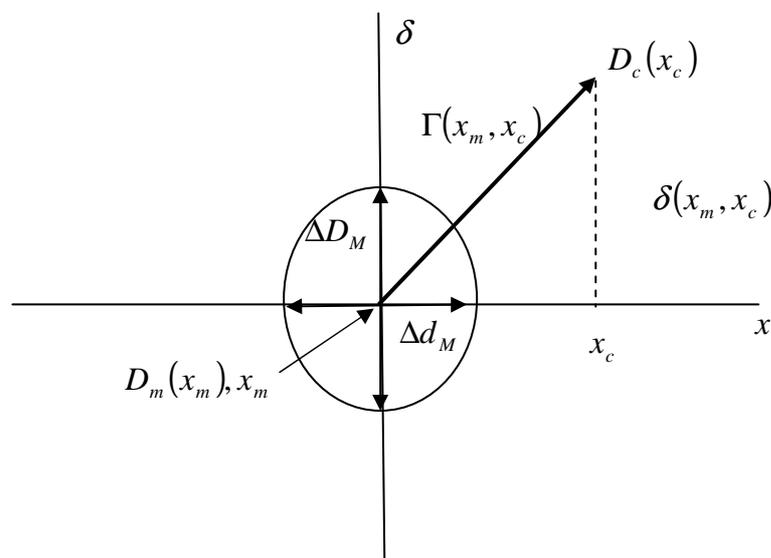


Figura 6.1. Esquema del análisis gamma en 1D. La extensión a distribuciones 2D o 3D es inmediata (20).

Dado un punto de medida dosis absorbida-posición en ese espacio ($D_m(x_m), x_m$) y un punto calculado ($D_c(x_c), x_c$) el vector que los une ha de estar dentro del elipsoide de tolerancia definido por ΔD_M , Δd_M . Si se expresan las diferencias en dosis absorbida y

distancia en unidades de tamaño ΔDM y ΔdM respectivamente, el módulo de dicho vector se denomina gamma, Γ , y ha de ser menor que 1 para que se cumpla el criterio establecido (20). Este concepto ya fue introducido someramente en el punto 1.3.

Sobre el porcentaje de puntos que cumplen que Γ sea menor que 1 se establecen los criterios que sirven para dar por buena una comparación entre dos distribuciones, es decir, para decir que las distribuciones de dosis se consideran semejantes. Cuando se utiliza una utilidad que realice el análisis gamma aplicado a dos matrices que se quieren comparar, es muy importante tener cuidado con dos aspectos fundamentales:

- La resolución de las matrices puede influir en el resultado final que se obtiene, especialmente si la utilidad que realiza la comparación utiliza las resoluciones originales y éstas son parecidas. Si son muy diferentes solo tiene sentido realizar la comparación gamma en los puntos de la matriz con menos resolución.
- El resultado depende mucho de lo que se tome como valor de dosis de normalización al que se asigna el 100%. En el caso de una matriz de dosis absorbida suma de todos los haces en un paciente, parece claro identificar el 100% con la dosis prescrita, pero en el caso de comparar planos de dosis no homogénea e individuales (por haz), el resultado varía mucho según se elija normalizar en el máximo, en la media o en un punto significativo de la distribución.
- La gamma local establece los porcentajes en dosis con respecto al valor local de la dosis absorbida, es decir, compara la dosis medida y calculada y divide por la calculada (considerada de referencia), por lo que siempre resulta más restrictiva que la global, donde los porcentajes se establecen siempre con respecto a un valor representativo de la matriz y en general cercano al máximo (dosis prescrita, zona de dosis homogénea dentro de campo...). Por su naturaleza, la gamma local resalta siempre más las diferencias que se encuentran para valores bajos de dosis.

En el presente trabajo se usa el análisis gamma cuando se comparan planos de dosis medidos con el EPID con los calculados por el SPT. También se comparan con los obtenidos en el SPT los planos de dosis medidos en el interior del maniquí IMRT phantom, en el caso de

la película radiocrómica, o con el detector MatriXX en el interior del maniquí Multicube. Por último, se emplea análisis gamma cuando se analizan distribuciones de dosis con el programa Compass. En los tres primeros casos es una evaluación gamma 2D. Para los planos de dosis provenientes de las medidas realizadas con el EPID, el 100% se define de manera arbitraria en una zona de dosis significativa y homogénea del plano analizado correspondiente. Para las medidas hechas con placas radiocrómicas y para los planos medidos en Multicube se utiliza la dosis prescrita por fracción como dosis del 100% y se realiza un análisis gama2D global. En el caso de los resultados del programa Compass, se realiza un análisis gamma 3D global por órganos, bien en el contorno externo, en los PTVs o en los OARs siendo el 100% de la dosis la dosis prescrita en el PTV principal. Se fija un umbral, por ejemplo 10% de puntos como máximo con $\text{gamma} > 1$, y se analiza que cada volumen lo cumpla. Si alguno de los volúmenes estudiados tiene más puntos fuera del criterio que los fijados en el umbral, el plan se considera un positivo y se rechaza según el criterio. En todos los casos el criterio gamma es 3%-3mm.

6.2.2. Funciones propuestas y variables que evalúan la diferencia entre los HDV obtenidos por dos métodos.

En este trabajo se definen y proponen una serie de nuevas funciones que se construyen a partir de los HDV y de las restricciones empleadas en los tratamientos de IMRT. El objetivo de construir estas funciones es intentar resumir en pocos parámetros las diferencias encontradas entre los HDV extraídos de matrices de dosis de verificación y los HDV obtenidos a partir de las matrices de dosis calculadas con el algoritmo SCC, que son aquellos que se quieren verificar. Se proponen para su estudio cuatro funciones. Las tres primeras que veremos se definen para cada algoritmo de los que se dispone por separado, ya que se calculan usando valores concretos de los puntos del HDV y las restricciones establecidas para los planes que figuran en la tabla 6.I. La última función no usa las restricciones y directamente compara valores de los HDV obtenidos usando dos algoritmos diferentes. Para el PTV principal se evaluarán las diferencias en dosis media (DM), en cobertura a través de la D95 y en los máximos usando D2. Para el PTV secundario se evaluarán las diferencias en dosis media (DM) y en cobertura a través de la D95. Para un OAR dependerá del tipo de órgano; dosis media en parótidas, D1 en medula, tronco, cristalino y nervios ópticos y D20 y

D30 en recto y vejiga. El número total de restricciones evaluadas se denominará N. N es el número de OARs más 3 (restricciones PTV principal) + 2 (en caso de que haya PTV secundario). A continuación se definen las cuatro funciones:

- Función incumplimiento (FI): Definida, para un algoritmo concreto que proporcione una matriz de dosis, como la suma cuadrática de las diferencias entre las dosis de evaluación que figuran en la tabla 6.I, obtenidas a partir del HDV, y las dosis de restricción consideradas según el volumen v, y calculadas solo en caso de que se incumpla la dosis de restricción $D_{v,r}$. El subíndice v puede corresponder por tanto a un PTV o a un OAR y en el caso del PTV puede corresponder a 3 valores (D95, DM y D2) si es un PTV principal y a 2 (D95 y DM) si es un PTV secundario. Esta suma se renormaliza a porcentaje de dosis con respecto a la prescrita para el volumen principal.

$$FI = \frac{100 \cdot \sqrt{\sum_{v=1}^N \delta_{v,r} (D_{v,eval} - D_{v,r})^2}}{D_{prescrita}} \quad (\text{ec 6.1})$$

donde

$$\delta_{v,r}=1 \text{ si } D_{PTV,eval} < D_{PTV,r} \text{ o } D_{OAR,eval} > D_{OAR,r}$$

$$\delta_{v,r}=0 \text{ si } D_{OAR,eval} < D_{OAR,r} \text{ o } D_{PTV,eval} > D_{PTV,r}$$

Las restricciones $D_{v,r}$ que se van a emplear pueden encontrarse en la tabla 6.I. La función está normalizada a la dosis prescrita ($D_{prescrita}$). FI se calcula para cada algoritmo de manera independiente, por ello se utiliza el subíndice eval, que corresponde al método de evaluación, de manera general, y puede ser el que sirve para verificar (verif) o el que se está verificando (SCC).

Habrà por tanto una FI para el sistema verificador FI_{verif} y otra para el verificado FI_{SCC} , de modo que se puede definir la diferencia de ambas como $\Delta FI = FI_{verif} - FI_{SCC}$ y esta diferencia mide, por tanto, el empeoramiento producido en la verificación en términos de incumplimiento de objetivos. Así, nos sirve para evaluar la diferencia entre los HDV obtenidos por dos métodos. En función del algoritmo con el que se evalúe, es decir, del sistema verificador que se emplee, se generarán la variables $\Delta FIMCesc$ y $\Delta FICMPM$.

- Función incumplimiento ponderado (FIP). Igual que la anterior pero aplicando factores de ponderación a cada uno de los sumandos.

$$FIP = \frac{100 \cdot \sqrt{\sum_{v=1}^N w_v \delta_{v,r} (D_{v,eval} - D_{v,r})^2}}{D_{prescrita}} \quad (\text{ec 6.2})$$

donde

$$\delta_{v,r}=1 \text{ si } D_{PTV,eval} < D_{PTV,r} \text{ o } D_{OAR,eval} > D_{OAR,r}$$

$$\delta_{v,r}=0 \text{ si } D_{OAR,eval} < D_{OAR,r} \text{ o } D_{PTV,eval} > D_{PTV,r}$$

siendo w_v el factor de ponderación correspondiente al volumen v . El propósito de esta función es ajustar los valores de w_v para intentar obtener el nivel de discriminación entre planes más eficiente posible para la métrica asociada a la función incumplimiento. De manera análoga a FI se define $\Delta FIP = FIP_{verif} - FIP_{SCC}$. En realidad FI es un caso particular de FIP con todos los pesos iguales a 1, por lo que el análisis de FI se hará conjuntamente con el de FIP. En función del algoritmo con el que se evalúe, que puede ser MCesc o CMPM, se generarán las variables ΔFIP_{MCesc} y ΔFIP_{CMPM} respectivamente.

- Función global de diferencias en dosis por objetivos (FGO): Definida como la suma de las diferencias entre los valores de dosis de las restricciones de la tabla 6.I y las dosis obtenidas al evaluar el HDV para cada una de ellas. Se diferencia de la función incumplimiento en que no es necesario que se incumpla el objetivo para que exista contribución en cada punto de evaluación (nótese que no tiene funciones δ en la definición) y que puede ser positiva o negativa. El signo del sumando cambia según sea la restricción, de dosis mínima o de dosis máxima.

$$FGO = \frac{100 \cdot \sum_{v=1}^N k \cdot (D_{v,r} - D_{v,eval})}{D_{prescrita}} \quad (\text{ec 6.3})$$

donde $k=1$ si v corresponde a un OAR o a D2 del PTV principal y $k=-1$ si v corresponde a DM o D95 de un PTV.

Análogamente a FI y FIP, FGO existe para cada algoritmo por separado. También se renormaliza a porcentaje de dosis con respecto a la prescrita. De nuevo se puede definir $\Delta FGO = FGO_{\text{verif}} - FGO_{\text{SCC}}$ para medir las diferencias entre el sistema verificador y el algoritmo que se verifica SCC. En función del algoritmo con el que se evalúe se generarán las variables $\Delta FGO_{\text{MCesc}}$ y ΔFGO_{CMPM} .

- Función global de diferencias en dosis (FG): Definida como la suma con su signo de las diferencias extraídas de los HDV obtenidos con el algoritmo de verificación (MCesc o CMPM) y el algoritmo verificado (SCC), siempre que esas diferencias supongan un empeoramiento en el sentido establecido por las restricciones de la tabla 6.I. Nótese que esta función compara siempre los resultados correspondientes a dos algoritmos y no utiliza las restricciones impuestas. Las diferencias que hagan que en los puntos de evaluación el algoritmo de verificación proporcione un resultado mejor no se tienen en cuenta.

$$FG = \frac{100 \cdot \sum_{v=1}^N \delta_{\text{verif}, \text{SCC}} (D_{v, \text{verif}} - D_{v, \text{SCC}})}{D_{\text{prescrita}}} \quad (\text{ec 6.4})$$

donde

$$\delta_{\text{verif}, \text{SCC}} = -1 \text{ si } D_{\text{PTV}, \text{verif}} < D_{\text{PTV}, \text{SCC}}$$

$$\delta_{\text{verif}, \text{SCC}} = 1 \text{ si } D_{\text{OAR}, \text{verif}} > D_{\text{OAR}, \text{SCC}}$$

$$\delta_{\text{verif}, \text{SCC}} = 0 \text{ si } D_{\text{OAR}, \text{verif}} < D_{\text{OAR}, \text{SCC}} \text{ o } D_{\text{PTV}, \text{verif}} > D_{\text{PTV}, \text{SCC}}$$

Es importante hacer notar que estas funciones en su definición incorporan el sentido clínico de la verificación, ya que, de una u otra manera, incluyen las restricciones del tratamiento en su definición o bien el hecho de que las diferencias sean aquellas que suponen un empeoramiento de la distribución verificada. Si en el PTV la dosis de verificación que cubre el 95% del volumen es mayor que la de referencia, aunque las dos distribuciones discrepen este hecho no debe ser penalizado, o lo que es lo mismo, diferencias en dosis que vayan en sentido contrario al de las restricciones no se consideran de manera negativa.

Tabla 6.I. Restricciones $D_{v,r}$ que deben cumplirse en los planes de IMRT, a partir del análisis HDV	
ORGANOS DE RIESGO	Restricción
Tronco cerebral	$D1 < 53 \text{ Gy}$
Nervios ópticos y quiasma (hipófisis)	$D1 < 49,25 \text{ Gy}$
Médula espinal	$D1 < 44,5 \text{ Gy}$
Parótidas	$DM < 26 \text{ Gy}$
Cristalinos	$D1 < 8\text{Gy}$
Recto	$D20 < 70 \text{ Gy}$
Vejiga	$D30 < 70 \text{ Gy}$
PTV	Restricción
Restricción 1	$D95 > 95\% \text{ dosis prescrita}$
Restricción 2	$DM > 0,97^* \text{ dosis prescrita}$
Restricción 3	$D2 < 110\% \text{ dosis prescrita}$

Para la evaluación de FI, FIP, FGO, FG y las diferencias asociadas ΔFI , ΔFIP , ΔFGO se ha elaborado una aplicación que se describe en 6.4.

6.2.3. Funciones radiobiológicas

A partir de los HDV diferenciales es posible evaluar la probabilidad de control tumoral (TCP) y la probabilidad de complicaciones en tejido sano (NTCP). Tanto el sistema Compass 3.0, como el algoritmo MC, como el SPT PCRT3D proporcionan matrices de dosis 3D y HDV que pueden usarse para el cálculo de estas probabilidades. Entre los sistemas de verificación y SCC pueden calcularse las diferencias en las probabilidades ΔTCP y $\Delta NTCP$, lo que también supone una métrica de comparación. Estas funciones se han evaluado utilizando parámetros radiobiológicos estándar tomados de la bibliografía y siguiendo el formalismo que se expone en (19). En dicho formalismo se utiliza un modelo fenomenológico de dos parámetros para determinar la TCP basado en una expresión tipo logit:

$$TCP = \prod_{i=0}^N \left[\frac{1}{1 + \left(\frac{D_{50}}{D_i} \right)^{4\gamma_{50}}} \right]^{v_i} \quad (\text{ec 6.5})$$

En este modelo se asume un tumor homogéneo con una distribución de dosis no homogénea. v_i es el volumen fraccional que recibe un nivel de dosis D_i , D_{50} es la dosis que

da como resultado una TCP del 50% y γ_{50} es la pendiente normalizada de la curva de TCP en D_{50} . Los valores que se han tomado se pueden encontrar en la tabla 6.II.

Para el cálculo de la NTCP se ha utilizado el modelo de Lyman-Kutcher-Burman (LKB) (67), el cual es ampliamente aceptado. Se ha seguido la siguiente formulación:

$$NTCP(gEUD) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{u^2}{2}} du \quad (\text{ec 6.6})$$

donde

$$t = \frac{gEUD - TD_{50}}{m \cdot TD_{50}} \quad (\text{ec 6.7})$$

$$gEUD = \left(\sum_i v_i D_i^{\frac{1}{n}} \right)^n \quad (\text{ec 6.8})$$

Por tanto se hace depender la probabilidad de control tumoral de la dosis equivalente uniforme generalizada, gEUD, propuesta por Niemierko (68). TD_{50} es la dosis que produciría un 50% de complicaciones si el órgano fuera irradiado de manera uniforme, m es un parámetro específico de un tejido, inversamente proporcional a la pendiente de la curva dosis-respuesta, que representa la probabilidad de complicación en el OAR en función de la dosis, y n es el parámetro de efecto volumen, que indica la dependencia de la probabilidad de complicación con el volumen irradiado y que coincide con el parámetro “a” descrito en (68). Para tejido sano n es positivo y menor o igual que 1. A la luz de la ecuación 6.8, es claro que cuando $n = 1$, gEUD es la dosis media en el órgano. La expresión 6.6 para la NTCP puede reformularse en términos de la función error para la que, a su vez, se ha utilizado un desarrollo polinómico que permite calcularla fácilmente.

gEUD puede ser calculada también para el PTV, y se pueden encontrar valores estándar para el parámetro a de la gEUD en (69).

Los índices radiobiológicos han mostrado ya su utilidad como métrica de verificación de tratamientos y también como indicadores objetivos de la posible robustez de un plan con respecto a diferentes tipos de error, ya que el impacto radiobiológico de un cambio en la dosis que recibe un órgano depende de la diferencia entre dicha dosis y aquellas que tienen influencia en las curvas de dosis-respuesta biológica para un órgano concreto (19).

Los valores a utilizar para calcular los índices se encuentran en la tabla 6.II.

Tabla 6.II Parámetros utilizados en las ecuaciones 6.5 a 6.8				
ÓRGANO	D50/TD50	γ_{50}	n (1/a)	m
PTV cabeza y cuello	5044	1,73	-0,1	
PTV próstata	7050	2,66	-0,1	
PTV cerebral	4800	2	-0,1	
PTV pulmón	5100	1,6	-0,1	
RECTO	7200		0,06	0,15
VEJIGA	6200		0,13	0,11
CABEZA FEMORAL	6500		0,25	0,12
MEDULA	6650		0,05	0,175
TRONCO	7000		0,05	0,18
PAROTIDA	2840		1	0,18
CRISTALINO	1250		0,3	0,27
NERVIO OPTICO	6500		0,25	0,14
CORAZON	4800		0,35	0,1
RIÑON	4000		0,7	0,1
QUIASMA	6500		0,25	0,14
INTESTINO DELGADO	5500		0,15	0,16

Un concepto interesante y que será de utilidad para la evaluación de planes es el de robustez, descrito en (19), pero no propuesto para tal fin por los autores. Los índices radiobiológicos son, evidentemente, sensibles a un plan que cumple las restricciones con escaso margen. Un cambio moderado en el valor de la dosis de verificación con respecto a la dosis verificada para los valores de la tabla 6.I, extraídos de los HDV correspondientes, puede conducir a saltos grandes en la NTCP. En este sentido parece adecuado incorporar un **índice de robustez (IR)** al plan a la hora de estudiar su aceptabilidad para su tratamiento y también para comparar con los índices de robustez que proporcionan los diferentes

algoritmos de verificación utilizados. La manera de construir el índice se propone este trabajo, y sería la siguiente:

- En primer lugar es necesario adaptar las curvas dosis-respuesta a las restricciones impuestas desplazándolas hasta situarlas en una zona de valores de dosis próxima a los valores de restricción. La razón, para ello, es que en general la TCP y la NTCP no son sensibles a cambios de dosis en el entorno de las restricciones de la tabla 6.I (u otras que se propusieran) si los parámetros de D_{50} y de TD_{50} no se corresponden con las dosis de restricción para PTV u OAR respectivamente. Un ejemplo sencillo es pensar que ocurre con la TCP de un paciente cuyo plan de IMRT corresponde tan sólo a una fase del tratamiento. Por ejemplo, la TCP de un tratamiento de próstata de 56 Gy (ver tabla 2.I), es menor del 10% y se encuentra en zonas de la curva dosis respuesta poco sensibles a cambios en la dosis. Para incrementar esta sensibilidad se puede cambiar, de manera artificial, el valor de D_{50} en el cálculo de la TCP, de tal modo que la dosis prescrita se sitúe en una zona de probabilidad de control tumoral en torno al 80 o 90%. Del mismo modo la curva dosis-respuesta en médula tiene una NTCP muy baja si la D1 de médula está por debajo de 45 Gy, como ocurre en todos los planes calculados con SCC; en este caso conviene modificar TD_{50} para que el gradiente de la curva dosis respuesta caiga en la zona de interés. El resultado son unas curvas dosis-respuesta modificadas artificialmente que conducen a unos índices que denominaremos modificados: TCP_{MOD} y $NTCP_{MOD}$, contruidos a partir de $D_{50,MOD}$ y $TD_{50,MOD}$. Los índices radiobiológicos modificados pierden, evidentemente, la significación radiobiológica, y responden a una modificación matemática que permite que se puedan utilizar como posibles predictores de la viabilidad de una planificación.
- Se calculan las pendientes de las curvas de TCP_{MOD} y $NTCP_{MOD}$ en función de gEUD utilizando el modelo del apartado 6.4. Usando las ecuaciones de la 6.5 a la 6.8 se puede evaluar la derivada parcial en el caso de la TCP para cada PTV. En el caso del TCP, es necesario calcular primero la gEUD siguiendo la ecuación 6.8 para luego derivar la ecuación 6.5 en el caso en el que el volumen fraccional es 1 pues todos los vóxeles del PTV tienen una dosis igual a gEUD.

$$\frac{\partial TCP_{MOD,i}}{\partial gEUD_i} = \frac{4 \cdot D_{50,MOD} \cdot \left(\frac{D_{50,MOD}}{gEUD_i} \right)^{4 \cdot \gamma_{50} - 1} \cdot \gamma_{50}}{\left(1 + \left(\frac{D_{50,MOD}}{gEUD_i} \right)^{4 \cdot \gamma_{50}} \right)^2 \cdot gEUD_i^2} \quad (\text{ec 6.9})$$

así como la derivada parcial de NTCP para cada uno de los órganos de riesgo

$$\frac{\partial NTCP_{MOD,i}}{\partial gEUD_i} = \frac{0.398942 \cdot e^{-\frac{(gEUD_i - TD_{50,MOD})^2}{2 \cdot m^2 \cdot TD_{50,MOD}^2}}}{m \cdot TD_{50,MOD}} \quad (\text{ec 6.10})$$

Las pendientes máximas de las curvas se dan cuando $gEUD = D_{50,MOD}$ para el PTV y $gEUD = TD_{50,MOD}$ para los órganos de riesgo, siendo sus valores

$$\frac{\partial TCP_{i,MOD}}{\partial gEUD_i} \Big|_{\max} = \frac{\gamma_{50}}{D_{50,MOD}} \quad \text{y} \quad \frac{\partial NTCP_{i,MOD}}{\partial gEUD_i} \Big|_{\max} = \frac{0.398942}{m \cdot TD_{50,MOD}}$$

respectivamente

- Los valores obtenidos de $gEUD$ para cada plan y órgano se usan para evaluar el gradiente de las curvas de TCP y NTCP en el punto que se esté considerando y su proporción con respecto al gradiente máximo posible, multiplicándose las componentes para cada volumen i , siendo N el número total de órganos incluido el PTV. Se define así el **índice de robustez (IR)**, como:

$$IR = \prod_{i=0}^N \left(1 - \frac{\frac{\partial XCP_{i,MOD}}{\partial gEUD_i}}{\frac{\partial XCP_{i,MOD}}{\partial gEUD_i} \Big|_{\max}} \right) \quad (\text{ec 6.11})$$

Nótese que en la ecuación 6.11 se ha dejado la primera inicial del índice radiobiológico para sustituir por "T" o por "NT" en función de que el volumen considerado sea un PTV o un OAR.

El índice de robustez se hace 0 en cuanto uno de los volúmenes tiene una $gEUD$ que se sitúa en el gradiente máximo de la curva de TCP o de NTCP, por lo que otra posibilidad es ponderar la contribución máxima de cada volumen en función de lo que importe su

inestabilidad en la verificación. Por ejemplo, es evidente que si los cristalinicos están en un punto crítico esto importa menos que para la médula o para el PTV, por lo que la expresión podría sustituirse por una del tipo:

$$IR = \prod_{i=0}^N 1 - w_i \frac{\frac{\partial XCP_{i,MOD}}{\partial gEUD_i}}{\frac{\partial XCP_{i,MOD}}{\partial gEUD_i}_{max}} \quad (\text{ec 6.12})$$

Siendo w_i un valor entre 0 y 1 definido para cada volumen.

Nótese también que con respecto al índice de robustez, un plan que rebase ampliamente las restricciones en un OAR también puede ser un plan robusto. Por ejemplo, si una de las dos parótidas en un caso de cabeza y cuello se da por perdida y, como consecuencia, la dosis media que recibe está lejos (muy por encima) de la restricción, IR no se verá afectado aunque la evaluación de dosis en ella sea muy diferente para las dos distribuciones que se estén comparando.

6.2.4. Correlación entre sistemas clasificadores

Se van a introducir dos tipos de métrica (ver apartado 6.3). El primer tipo es aquel que proporciona una variable continua. Se tratará de evaluar la correlación entre este tipo de variables continuas, por medio del coeficiente de correlación de Pearson y su significación.

Por otro lado podemos definir métricas categóricas, correspondientes a variables binarias, que proporcionan directamente una colección de positivos y negativos con tantos elementos como número de planes estudiados. Con cada variable binaria se construye una serie cuyos valores pueden coincidir o no con los que proporcionan otras. Por convención, se identifica el positivo con el valor 0 en la variable binaria y con el plan rechazado según esa variable. De manera análoga se identifica el negativo con el valor 1 y con el plan aceptado. Entre cada pareja de variables binarias se pueden establecer tablas de contingencia sencillas. Si una de las variables que se compara se considera de referencia, simplemente por conveniencia, la tabla se puede caracterizar tal como se muestra en la tabla 6.III.

Tabla 6.III. Tabla de contingencia genérica				
Tabla de contingencia		VARIABLE REFERENCIA		Total
		Positivo 0	Negativo 1	0
VARIABLE CONTRASTADA	Positivo 0	VP	FP	VP+FP
	Negativo 1	FN	VN	FN+VN
Total		POSITIVOS	NEGATIVOS	TOTAL CASOS

En la tabla 6.III se establece la existencia de verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN). Tal como se indica, al considerar una variable como de referencia, pueden establecerse sus resultados como “verdaderos” y por tanto la variable que se compara con ella generará “falsos” si no coincide en todos los casos. La comparación se hace por tanto relativa a la variable considerada como de referencia y esa elección es en principio, arbitraria y a conveniencia, ya que en el caso que nos ocupa **no hay ningún clasificador que se considere perfecto**.

La prueba a realizar sobre las tablas de contingencia será la Chi-cuadrado y el valor de la significación asintótica tiene que ser bajo para descartar que los resultados de dos métricas sean independientes. Esto es sencillo de deducir, ya que la prueba Chi-cuadrado tiene como hipótesis nula que la distribución de la variable de referencia sobre la variable contrastada no es independiente, es decir, que ambas están relacionadas. Dos métricas estarán muy relacionadas cuando el número de FP y FN sea pequeño en comparación con sus respectivos VN y VP, es decir, las frecuencias observadas han de diferir notablemente de las esperadas, que corresponden a un reparto uniforme en función del número de positivos y negativos reales impuesto por la variable de referencia. Por tanto esperamos un valor de significación inferior a 0,05 y una Chi-cuadrado lo más alta posible para poder identificar los resultados de dos métricas como relacionados.

Una forma de construir una variable binaria a partir de una continua, es buscando valores que superan o no un umbral prefijado. En particular, a través del análisis ROC presentado en 6.2.5, se pueden obtener valores de corte que permitan construir dicha variable binaria separando los planes entre aquellos que tienen un valor por encima de la variable y aquellos que lo tienen por debajo. Una vez establecido el valor de corte, denominado también punto

de corte (PC), estas variables continuas también pueden participar en el establecimiento de tablas de contingencia.

6.2.5. Análisis ROC de los sistemas clasificadores.

Las curvas ROC y su posterior análisis son utilizadas siempre que se desea establecer cómo de apropiado es un sistema clasificador, caracterizado por una variable continua (obtenida del análisis de los datos), para distinguir entre los objetos que clasifica, en nuestro caso planes aceptables y planes inaceptables.

Una curva ROC es un gráfico de sensibilidad (S) frente a 1-especificidad (E), siendo definidas ambas como

$$S=VP/(VP+FN) \quad E=VN/(VN+FP)$$

Ambas se suelen expresar en porcentaje; la sensibilidad es tanto mejor para un sistema clasificador cuanto menos falsos negativos proporciona, y la especificidad es tanto mejor cuanto menos falsos positivos detecta.

Las curvas ROC se construyen variando los valores umbrales utilizados para clasificar los objetos (en este caso los planes de tratamiento) entre aceptables y no aceptables y representando los puntos en el espacio 2D, empleando para el eje de abscisas 1-Especificidad y la Sensibilidad para el eje de ordenadas. Para definir cómo de bueno es un sistema clasificador, o lo que se denomina capacidad de discriminación, se proporciona el área bajo la curva ROC, denominada como AUC (del inglés "area under the curve"). AUC se mueve entre 1 y 0,5. Si AUC = 1 el sistema clasificador es igual de bueno que el sistema de referencia mientras que si es 0,5 el sistema actúa como un clasificador aleatorio.

A modo de guía para interpretar las curvas ROC se han establecido los siguientes intervalos para los valores de AUC:

[0,5-0,6): Test malo [0,6-0,75): Test regular

[0,75-0,9): Test Bueno [0,9-0,97): Test muy bueno [0,97-1): Test excelente

AUC puede interpretarse como la probabilidad de que un positivo escogido aleatoriamente de la muestra tenga un valor más alto de la variable continua que se está usando como clasificador que un negativo escogido del mismo modo (asumiendo que un

positivo da valores más altos de la variable que un negativo, ya que también puede ser al revés). En el caso del presente trabajo, esto significa que si para un sistema clasificador de planes sometido a análisis ROC, AUC es 0,8, existe un 80% de probabilidad de que la clasificación de un plan basada en ese sistema coincida con la del sistema de referencia.

Además, es posible encontrar aquellos valores de corte que proporcionan el criterio óptimo para una mejor clasificación, por lo que el análisis ROC es muy útil para establecer el nivel de acción de un sistema verificador. El valor de corte geoméricamente más adecuado es aquel donde se encuentra el punto sobre la curva ROC más alejado de la diagonal que marca que la variable seleccionada actúa como un clasificador aleatorio. Encontrar el punto es sencillo, basta con girar virtualmente la curva ROC hasta que la diagonal se convierta en horizontal y ver qué punto es el que tiene más altura. Desde el punto de vista matemático y geométrico es correcto, pero esta aproximación ignora el coste relativo de un falso positivo y un falso negativo, así como la prevalencia. Por ejemplo si la prevalencia de un plan inaceptable es baja, tener un punto de corte muy sensible puede conducir a un número excesivo de falsos positivos, lo que supone un gasto inútil de tiempo y recursos. De manera opuesta si el coste relacionado con aceptar un plan que es inaceptable es alto, un punto de corte sensible podría ser preferible a uno de alta especificidad. Este asunto se puede analizar por medio del uso de puntos de corte optimizados por coste; los puntos de coste óptimos se calcularían por medio de una función que calcula el coste de los falsos negativos y la prevalencia de planes inaceptables. Los valores de corte son, en realidad, dependientes de la situación particular de cada centro y pueden incluir factores como los riesgos de irradiar un paciente con un plan inaceptable, ponderados con el tiempo extra que se necesita para corregir planes inaceptables. Este método de asignación de costes se antoja complicado y subjetivo por lo que una estimación razonable puede corresponder a variar el coste relativo de un falso positivo frente a un falso negativo hasta llegar al punto de corte óptimo para que, por ejemplo, en el caso de la verificación con cámara de ionización se sitúe en el 3% tradicional. Esto permite trabajar hacia atrás desde un punto de corte deseado y derivar el peso del coste que conduciría a un punto de corte con cámara óptimo en diferencia de dosis. La sistemática es aplicable al resto de métricas.

1-S y 1-E expresan la probabilidad de tener un falso negativo y un falso positivo respectivamente en el valor de corte examinado. Por tanto podemos usar un peso relativo k

que descompense la suma de esas probabilidades, lo que dependiendo del valor de k minimiza la función de coste en otro punto distinto del que geoméricamente proporciona la curva ROC.

$$\text{Prevalencia de planes inaceptables} = 100 \cdot (\text{VP} + \text{FN}) / (\text{VN} + \text{FP}) \quad (\text{ec. 6.13})$$

$$\text{Coste} = k \cdot (1 - S) + (1 - E) \quad (\text{ec. 6.14})$$

Como se ha dicho más arriba, el uso de curvas ROC lleva implícito que se conoce la clasificación correcta de negativos y positivos. No es el caso del presente trabajo, ya que no existe una referencia absoluta que clasifique los planes sin lugar a dudas. Eso no resta utilidad a la herramienta, ya que, por conveniencia, se va a utilizar la métrica de diferencias en dosis y la métrica de diferencias en dosis basada en objetivos, descritas en 6.3, como el sistema clasificador perfecto, en unos casos basada en CPM y en otros basada en MCesc. Frente a él se van a evaluar el resto de métricas. No obstante, de manera general, la métrica de referencia puede ser cualquiera de las variables binarias definidas. En ese caso el análisis ROC está midiendo cómo de bien predice la variable contrastada los resultados de la utilizada como referencia.

También es importante aclarar que aunque se indica que la entrada del análisis ROC es una variable continua, es posible también usar un conjunto de variables binarias caracterizadas cada una de ellas por un umbral de aceptación. Aunque el conjunto es finito, cada tabla de contingencia, una por variable, permite generar un punto en el espacio ROC. La colección completa de puntos permite dibujar una curva ROC cuya área y punto de corte pueden ser calculados.

Para la realización del análisis ROC de las variables implicadas se dispone de un programa de análisis estadístico (SPSS para Windows versión 15.0). En el caso de las variables continuas (FI, FIP, FGO y FG), SPSS dibuja la curva automáticamente al ser contrastadas con respecto a una variable binaria de estado (cualquiera de las variables cuyo nombre empieza por "PASA"). En el caso de querer contrastar un conjunto de variables binarias caracterizada por distintos umbrales de aceptación (medidas con cámara, detrimento en las funciones radiobiológicas, función gamma por órganos), las curvas ROC son dibujadas usando una hoja de cálculo, y SPSS se utiliza para calcular la sensibilidad y especificidad en cada punto a través de la tabla de contingencia específica de cada umbral.

6.3. Resumen de métricas de comparación y variables binarias

En este apartado se presenta un resumen de las métricas utilizadas, es decir, de las distintas variables usadas para poder establecer si un plan de tratamiento calculado con el SPT se considera que es aceptado o no. El resultado de la evaluación de un plan con una determinada métrica será de tipo binario, el plan se acepta o se rechaza, como queda dicho en el apartado anterior. Las variables binarias se asignan de modo que 1 significa que el plan se acepta (para la métrica en cuestión) y 0 significa que el plan se rechaza.

Las métricas que se van a comparar para los planes pertenecientes a los pacientes escogidos con respecto al cálculo realizado por el SPT (SCC), son las siguientes:

1) Verificación convencional

Dosis en un punto medida con cámara de ionización y comparación gamma de distribuciones planares de dosis medidos con EPID, placa radiocrómica o matriz de cámaras de ionización en el interior del maniquí Multicube: Se acepta el plan si $|\Delta D| < 3\%$ para la dosis en la cámara de ionización y si el % de puntos con gamma < 1 es menor de 10% para la media de todos los planos y, además, si la media más la desviación estándar no supera el 11%. A partir de estos criterios de aceptación se construye la variable binaria PASA_CONVENCIONAL. También se puede usar exclusivamente la cámara de ionización como elemento discriminador. Además, variando el umbral de aceptación se construye un conjunto de variables binarias PASA_CAMARAX, con el criterio de aceptación $|\Delta D| < X\%$.

2) Métrica de diferencias en dosis

Se evalúan las diferencias de valores resultantes de los puntos del HDV que nos sirven para fijar las restricciones (ver tabla 6.I) obtenidos con el método de verificación y el usado como referencia. Se acepta el plan si el empeoramiento, en caso de producirse, en los valores de los HDV de verificación, expresados en la tabla 6.I es menor del 3% de la dosis prescrita. Por empeoramiento se entiende dosis de verificación más baja en PTVs para D95 y DM, más alta para D2, y en OARs dosis más alta de cualquier parámetro usado como restricción. Si el PTV es secundario no se tiene en cuenta la restricción en D2.

El método de verificación puede ser MCesc o CPM. A partir de esta métrica se construyen las variables binarias PASA_MCesc y PASA_CPM. Esta métrica es también válida para los resultados obtenidos a partir de los histogramas con Delta4, siendo en ese caso la variable asociada PASA_DELTA. Excepcionalmente también se usa esta métrica entre CPM y CMPC, para explorar la validez del escalado proporcional en el apartado 7.3.3, y también usando MC sin escalar para valorar los efectos de las diferencias en el cálculo exclusivamente en el apartado 7.4.2 (se usa PASA_MC).

La elección, a priori no evidente, de la dosis prescrita como valor sobre el que se hace el 3%, no sólo para PTV, sino también para los OAR, se fundamenta en que el porcentaje de desviación local para los OAR siempre es mayor que para el PTV dada una desviación fija de dosis entre la restricción que se considere para algoritmo verificador y verificado. Un ejemplo extremo son los cristalinos. Se busca para ellos que D1 no supere 8Gy en tratamiento de 66Gy. Si el valor de D1 es 7Gy para SCC y 8Gy para CPM, por ejemplo, la diferencia en porcentaje local es del 14%, mientras que con respecto a la dosis prescrita es un 1,5%. Si la métrica de diferencias en dosis se deja con un valor del 3% local, el número de positivos aumentaría mucho, por lo que, o bien se expande la tolerancia en función de lo baja que sea la dosis evaluada, o bien se usa un porcentaje de dosis global, dejando el 3% usado para el PTV.

3) Métrica de diferencias basada en objetivos

La métrica anterior pudiera producir un exceso de planes rechazados, ya que no tiene en cuenta si se está sobrepasando la restricción establecida para el OAR o el PTV, es decir, puede producirse en la verificación un empeoramiento del 5% global en el tronco, por ejemplo, pero D1 puede continuar estando muy por debajo de la restricción establecida (53Gy), por lo que, aunque la dosis de verificación discrepa notablemente de la planificada, las consecuencias son insignificantes para el paciente, y por tanto rechazar un plan basándose en la métrica anterior podría suponer un gasto innecesario de recursos. Por ello la métrica propuesta en este apartado tiene estos hechos en cuenta y da por buenas diferencias que no incumplan los objetivos. Así, se construyen las variables binarias PASA_CPMOBJ,

PASA_MCescOBJ y PASA_DELTAOBJ, que, por definición, deben contener los mismos negativos que PASA_CMPM, PASA_MCesc y PASA_DELTA o más, pero nunca menos.

4) Métrica basada en la función gamma en los volúmenes de interés

Exclusivamente para el algoritmo CMPM se evalúa la función gamma global, en dichos volúmenes (PTVs, OARs y Externo) y se penaliza si en alguno de ellos se supera el 10% de puntos con $\gamma > 1$, siendo 3%-3mm el criterio gama aplicado. Para calcular la función gamma global se normaliza la dosis en cada elemento de volumen respecto a la dosis prescrita, a la que se le asigna el 100%. A partir de esta métrica se construye la variable binaria PASA_GAMMA. El umbral del 10% puede cambiarse para investigar cuál puede ser más apropiado, construyéndose de este modo un conjunto de variables binarias PASA_GAMMAX, con X variable. Se usarán los siguientes valores de X: 1%, 5%, 10%, 15%, 25%, 40% y 100%.

5) Métrica basada en diferencias de la función incumplimiento ponderado, ΔFIP

La función incumplimiento ponderado se expresa en dosis, y por tanto puede normalizarse como % de dosis con respecto a la dosis prescrita. Se intentará determinar cuál es un buen punto de corte discriminatorio para el valor de esta función, al ser una variable continua. Para ello se emplea el análisis ROC con respecto a PASA_CMPM, PASA_CMPMOBJ, PASA_MCesc y PASA_MCescOBJ. Una vez determinado el valor de corte, puede definirse una variable binaria asociada a la función incumplimiento, PASA_ΔFIPX, siendo X el valor de corte procedente del análisis. Dado que FIP puede calcularse para los algoritmos CMPM o MCesc, el nombre del algoritmo al que se refiere aparecerá al final del nombre de la variable. Con respecto a FI, cabe recordar que es un caso particular de FIP.

6) Métrica basada en las diferencias en la función global basada en objetivos, ΔFGO

Variable continua sobre la que se espera, también, obtener un valor de corte. Para ello se emplea el análisis ROC con respecto a PASA_CMPM, PASA_CMPMOBJ, PASA_MCesc y PASA_MCescOBJ. Puede construirse la variable binaria PASA_ΔFGOX a partir del valor de corte, indicándose también a qué algoritmo se refiere.

7) Métrica basada en la función global de diferencias en dosis, FG

Variable de la misma índole que las dos anteriores. Para ello se emplea el análisis ROC con respecto a PASA_CMPM, PASA_CMPMOBJ, PASA_MCesc y PASA_MCescOBJ. Puede construirse la variable binaria PASA_FGX a partir del valor de corte, indicándose también a qué algoritmo se refiere.

8) Métrica basada en diferencias de TCP y de NTCP

Las TCP y NTCP se calculan a partir de los HDV diferenciales que proporcionan los algoritmos CMPM, MCesc y SCC. Se acepta el plan si el empeoramiento de CMPM o MCesc con respecto a SCC, en caso de producirse, en los valores de TCP es inferior al 5% y los valores de NTCP no se incrementan más del 5% para los órganos considerados. A partir de esta métrica se puede construir la variable binaria PASA_TPCMPM5, extensible a otros umbrales que se consideren, o, de manera general PASA_TPCMPMX, como conjunto de variables binarias asociado. Si el algoritmo utilizado para evaluar los índices radiobiológicos es MCesc la variable binaria asociada se denomina PASA_TPCMescX.

9) Índice de robustez

El índice de robustez se calcula para SCC y CMPM por separado. Se puede medir la diferencia en dicho índice entre el plan de verificación (CMPM) y el plan de referencia (SCC), siendo las variables asociadas a esa diferencias ΔIR y PASA_IR (continua y binaria), añadiendo el algoritmo utilizado en el nombre de la variable.

10) Métrica basada en la complejidad del plan

Se ha encontrado que un buen parámetro para cuantificar la complejidad de un plan es, entre otros, el denominado ALPO (promedio de la apertura de pares de láminas, del inglés "Average Leaf-Pair Opening") (70,71), definido para IMRT como:

$$ALPO = \frac{\sum_{s,j} |x_R - x_L|_{s,j} \Delta UM_s}{\sum_{s,j=paresabiertos} \Delta UM_s} \quad (\text{ec 6.15})$$

siendo s el subíndice para los segmentos que constituyen un campo de IMRT estática y j el subíndice para las láminas que forman cada segmento. Las posiciones de cada

una de las láminas del par son x_R y x_L . Por tanto se suma el tamaño de las láminas que estén abiertas y se pondera por las unidades de monitor de cada segmento. La variable ALPO puede ser introducida en el análisis ROC como variable continua sobre la que determinar un punto de corte de discriminación. Otras variables que podrían estar relacionadas con la complejidad del plan son el número de segmentos NSEG, las UM totales (UMTOTALES) y las UM por segmento (UMSEG) y la variable compuesta $UM*SEG/DSES$. Para evaluar ALPO en cada uno de los planes estudiados se ha elaborado una rutina que lo calcula a partir del archivo RTplan (DICOM-RT) correspondiente al plan, de acuerdo con la ecuación 6.15. Hay un valor de ALPO por haz, por lo que al promedio para todos los haces se le va a denominar ALPOTOTAL.

En la tabla 6.IV se presenta un resumen de las diferentes métricas a evaluar.

MÉTRICA	MODO DE ANALISIS	VARIABLES ASOCIADAS	Se acepta plan si
Verificación convencional CÁMARA + EPID o Placa radiocrómica o MatriXX +Multicube	En Maniquí gamma 2D campo por campo gamma 2D global gamma 2D global	PASA_CAMARAX PASA_CONVENCIONAL	$\Delta D < X\%$ en CI $\Delta D < 3\%$ en CI $(\% \Gamma > 1)_{media} < 10\%$ y $(\% \Gamma > 1) + (\% \Gamma > 1)_{desvest} < 11\%$
Diferencias en dosis	Con MCesc y CMPM (válida también para DELTA4)	PASA_MCesc PASA_CMPM PASA_DELTA	Pérdida de D95 en PTV $< 3\%$ en los PTV principales Pérdida de DM en PTV $< 3\%$ en los PTV principales Aumento de D2 en PTV principal $< 3\%$ Aumento en los DX de OAR de tabla 6.I $< 3\%$ global
Diferencias basadas en objetivos	Con MCesc y CMPM (válida también para DELTA4)	PASA_MCescOBJ PASA_CMPMOBJ PASA_DELTAOBJ	Igual que la anterior, pero aunque no cumpla se acepta el plan si no se incumplen las restricciones de la tabla 6.I.
Basada en la función gamma en los volúmenes de interés	Con CMPM en PTVs, OARs y Externo	PASA_GAMMAX	$(\% \Gamma > 1 = X)$ variando umbral X entre 1 y 100%

Diferencias de la función incumplimiento ponderado	Con MCesc y CPM Se determina punto de corte PC y su correspondiente umbral de aceptación U	ΔFIP_{MCesc} ΔFIP_{CPM}	Se construye la variable binaria PASA_ΔFIPU donde U determina el umbral de aceptación ΔFI es un caso particular de ΔFIP
Diferencias en la función global basada en objetivos	Con MCesc y CPM Se determina punto de corte PC y su correspondiente umbral de aceptación U	$\Delta FGOM_{MCesc}$ $\Delta FGOC_{CPM}$	Se construye la variable binaria PASA_ΔFGOU donde U determina el umbral de aceptación
Función global de diferencias en dosis	Con MCesc y CPM Se determina punto de corte PC Y su correspondiente umbral de aceptación U	FGMCesc FGCPM	Se construye la variable binaria PASA_FGU donde U determina el umbral de aceptación
Diferencias de TCP y de NTCP	MCesc+ ΔTCP ΔNTCP CPM+ ΔTCP ΔNTCP	PASA_TCP _{MCesc} X PASA_TCP _{CPM} X	ΔTCP > -X y ΔNTCP < X para todos los volúmenes, con X variable donde X determina el umbral de aceptación
Índice de robustez	Del conjunto de HDV de SCC frente a las restricciones y diferencia $\Delta IR = IR_{CPM} - IR_{SCC}$ Se determina punto de corte PC y su correspondiente umbral de aceptación U	IR_SCC IR_CPM PASA_IR	Se construye la variable binaria PASA_IRU donde U determina el umbral de aceptación
Complejidad del plan	Media de la apertura de pares de láminas Se determina punto de corte PC y su correspondiente umbral de aceptación U	ALPO para cada haz ALPOTOTAL, promedio para todos los haces	Se construye la variable binaria PASAALPOU donde U determina el umbral de aceptación

6.4. Procesado de los HDV

Para el cálculo de todas las variables continuas y binarias de la tabla 6.IV es necesario el procesado de los HDV generados por Compass, por MC y por el sistema de planificación (SCC). Se dispone de cuatro matrices 3D de dosis por plan, correspondientes a las distribuciones de dosis calculadas con CMPC, CMPM, MC y SCC. Las correspondientes a MC y SCC ya están en el sistema de planificación, tal como quedó dicho en los apartados 4 y 5. Las matrices que provienen de CMPC y CMPM se han importado en formato RTDose dentro de PCRT3D. De este modo es posible generar cuatro conjuntos de HDV. Dado que para el cálculo de índices radiobiológicos es necesario disponer de los histogramas en formato diferencial y que la exportación de datos de los histogramas desde Compass no se podía automatizar, la opción de generar todos los histogramas en el SPT era la única posible para poder procesar los resultados. Se generaron 232 ficheros (58 planes, 4 matrices por plan), con los HDV diferenciales. En primera instancia, una aplicación, desarrollada en Visual C++, realiza varias tareas sobre ellos:

- Genera el conjunto MCesc procesando los HDV diferenciales para realizar un escalado de acuerdo con la descripción del apartado 6.1.
- Empareja conjuntos de histogramas para evaluar las métricas entre ellos. Se generaron las siguientes parejas: MC-SCC, CMPM-CMPC, CMPM-SCC, MCesc-SCC y CMPC-SCC.
- Cada pareja es procesada para calcular el histograma integral y extraer los valores de la tabla 6.I en cada volumen de interés. Identificando el tipo de volumen de manera automática se extrae, además, para los histogramas diferenciales la gEUD y, posteriormente, el TCP o NTCP (y también los valores modificados de acuerdo con lo expuesto en 6.4). También se calcula el índice de robustez de cada plan teniendo en cuenta la ecuación 6.11. El resultado es un conjunto de 58 ficheros por cada uno de los emparejamientos del párrafo anterior.

Los ficheros son importados en hojas Excel (Microsoft Office), para el agrupamiento de los valores de interés en tablas únicas para cada pareja de algoritmos. Una vez agrupados los datos de todos los pacientes en una hoja, se aplica un código de análisis en Visual Basic que calcula FI, FIP, FGO y sus correspondientes diferencias entre algoritmos, FG, y todas las

variables binarias descritas en la tabla 6.IV. El formato en columnas propicia la importación directa de datos en SPSS para el análisis ROC de variables continuas y las correlaciones entre variables. Los análisis ROC que corresponden a variables binarias, para las que se va cambiando el umbral de aceptación, se han realizado de forma manual, introduciendo el resultado de la variable binaria para cada umbral en una hoja Excel, generando las tablas de contingencia en SPSS y calculando sensibilidad y especificidad para cada umbral, integrando el área bajo la curva una vez analizados todos los puntos.

La variable ALPO y las variables de la verificación convencional (resultados de la verificación con cámara, gamma 2D por planos, etc...) han sido introducidos a mano en las tablas.

7. Resultados

7.1. Verificación del algoritmo de superposición de cono colapsado

El proceso de verificación del algoritmo SCC está descrito en (40). Los resultados principales, como se menciona en el apartado 4.2, corresponden a la comparación entre los calculados con el algoritmo SCC y con el MC, en una serie de maniqués descritos en la figura 4.1 y en la tabla 4.II, para dos energías de fotones, 6 y 18 MV, y tres tamaños de campo diferentes. Previamente se comprobó que los PDP y los perfiles de dosis en un maniqué de agua, calculados con MC usando DOSXYZnrc, coincidían con los experimentales, por lo que se dieron los seis espacios de fase por validados en agua dentro de la incertidumbre asociada al cálculo. La comparación entre los resultados experimentales y los calculados, de los PDP y los perfiles a 5, 10 y 20 cm de profundidad, condujo a un índice gamma < 1 para más del 95% de los datos estudiados. El criterio gamma considerado fue 2% de dosis relativa a la profundidad del máximo en el eje y 2 mm de distancia.

En cuanto a los maniqués heterogéneos es necesario mencionar que la dosis viene expresada en las figuras que siguen como dosis absorbida en el medio, dado que es la salida natural de los algoritmos MC y SCC. En las figuras 7.1 a la 7.4, aparece también un algoritmo denominado preciso (PR), que fue previo al desarrollo del algoritmo SCC en el sistema de planificación. PR está basado en un método muy similar al de Clarkson (40), siguiendo una separación del cálculo en dosis primaria y dispersa, ambas obtenidas a partir de datos experimentales. En el caso del algoritmo PR, la dosis en el interior de heterogeneidades fue escalada de acuerdo con (73), utilizándose un factor de corrección uniforme por densidad electrónica y por espectro dado un material. El objeto de incluir el algoritmo PR en la comparación es apreciar la mejora que supone el uso de algoritmos tipo SCC. Los cálculos se encuentran normalizados para todos los algoritmos a la profundidad del máximo en el eje de cada campo.

En la figura 7.1 podemos observar el resultado en el maniqué LUNG, donde se superponen porcentajes de dosis en profundidad calculados con MC, SCC y PR. Para los tres campos y las dos energías utilizadas, SCC presenta un comportamiento más cercano a MC que PR, siendo

prácticamente igual para los campos 10 cm x 10 cm y 20 cm x 20 cm y existiendo más discrepancias en el caso del campo pequeño 3 cm x 3 cm, donde se encuentra algún punto que supera el 7% de diferencia con respecto a MC, especialmente en el caso de alta energía. Las razones para este desacuerdo son consistentes con la bibliografía (74) y se fundamentan en la pérdida de equilibrio electrónico para campos pequeños y alta energía, que sólo puede ser modelada correctamente a partir de cálculos MC, ya que los núcleos de dispersión en los que se fundamenta SCC están calculados en un maniquí de agua en condiciones de equilibrio y son directamente adaptados a las condiciones del paciente utilizando la densidad másica y electrónica que se va encontrando a lo largo del eje del cono colapsado.

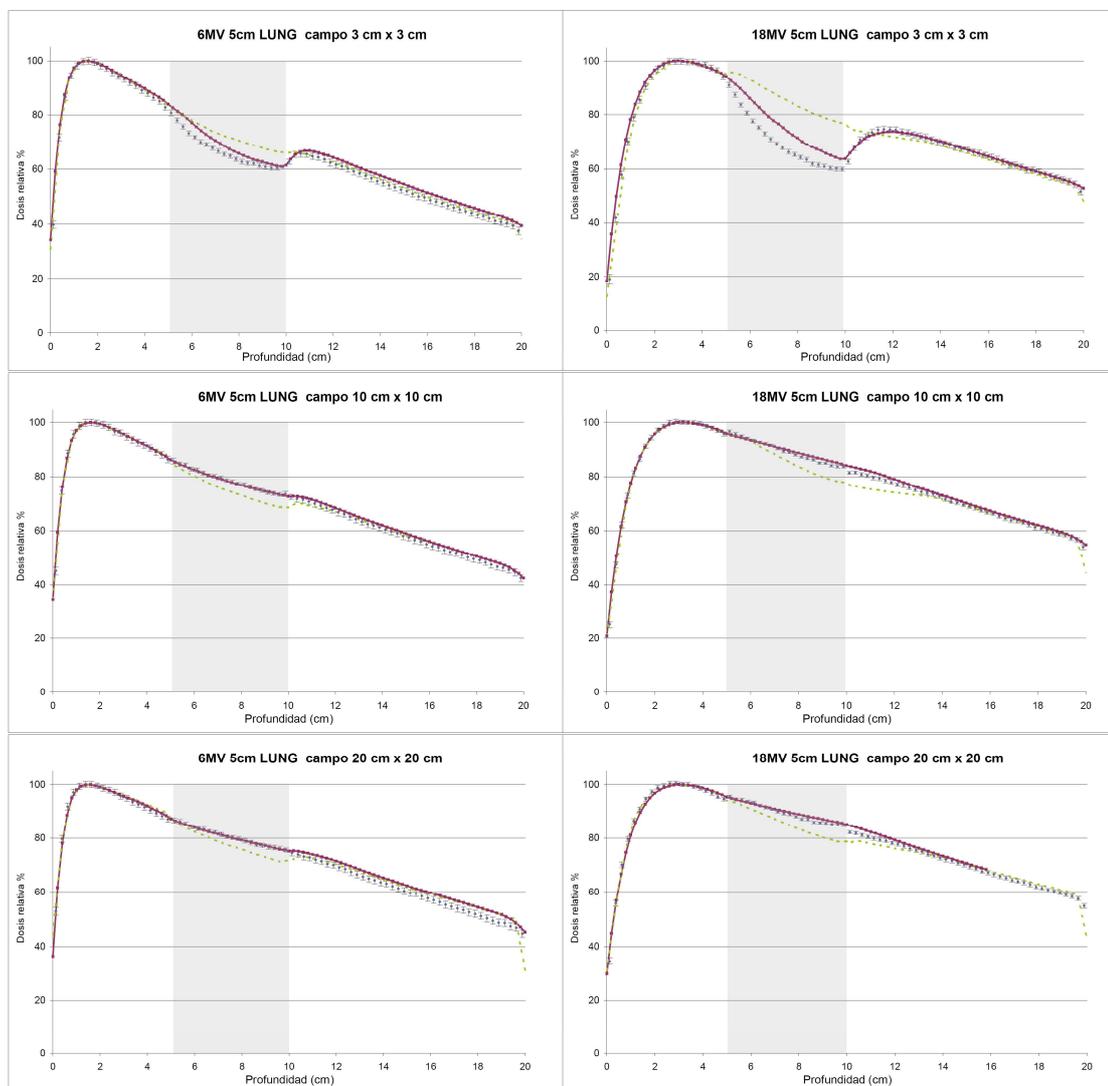


Figura. 7.1. Resultados en el maniquí LUNG para 6 MV (columna izquierda) y 18 MV (columna derecha) y diferentes tamaños de campo. La zona sombreada corresponde a la heterogeneidad. Se superponen las curvas de los tres algoritmos utilizados de acuerdo con la notación: **MC** **SCC** **PR**

En la figura 7.2 se observa lo mismo para el maniquí BONE. Para este maniquí el acuerdo es mejor que para el maniquí LUNG para todos los campos y las dos energías. Si se observan diferencias entre MC y SCC se localizan en las interfases, donde puede haber algún punto que supere el 3% de diferencia en dosis relativa al máximo. En particular, para un campo pequeño, el acuerdo continúa siendo muy bueno. Para 6 MV el algoritmo PR no da resultados muy diferentes, siendo su desacuerdo bastante mayor para 18 MV.

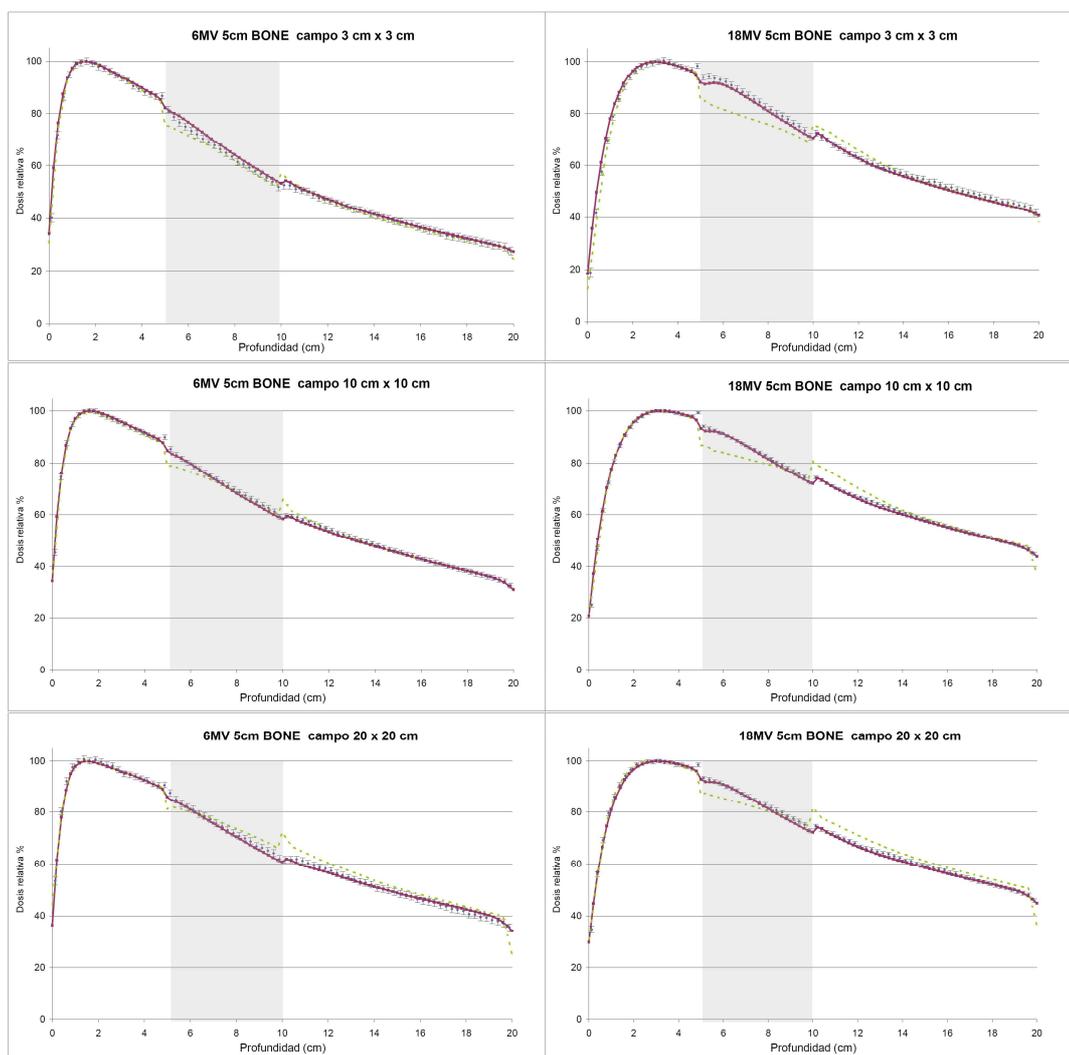


Figura 7.2. Resultados en el maniquí BONE para 6MV (columna izquierda) y 18MV (columna derecha) y diferentes campos. La zona sombreada corresponde a la zona de la heterogeneidad. Se superponen las curvas de los tres algoritmos utilizados de acuerdo con la notación: **MC** **SCC** **PR**

La figura 7.3 muestra los resultados para la geometría MADIASTINUM para 6 y 18 MV. Se presenta la comparación de tres perfiles a tres profundidades diferentes (10, 15 y 20 cm) para los tres algoritmos estudiados. Los tres perfiles cruzan la heterogeneidad de baja densidad a diferentes alturas. Los calculados con MC y SCC muestran acuerdo en un entorno de un 3% de la dosis relativa al máximo para más del 95% de los puntos estudiados. Aun estando dentro de valores aceptables, dentro de la heterogeneidad, SCC sobreestima la corrección con respecto a MC, quizá debido a que en el eje central el valor calculado con MC también es inferior al obtenido con SCC. En el caso de esta geometría el algoritmo PR muestra una corrección en profundidad insuficiente para 6 MV pero bastante acertada en el caso de 18 MV, si bien se pone de manifiesto que la penumbra del haz es más corta de lo que debería dentro de la heterogeneidad de baja densidad.

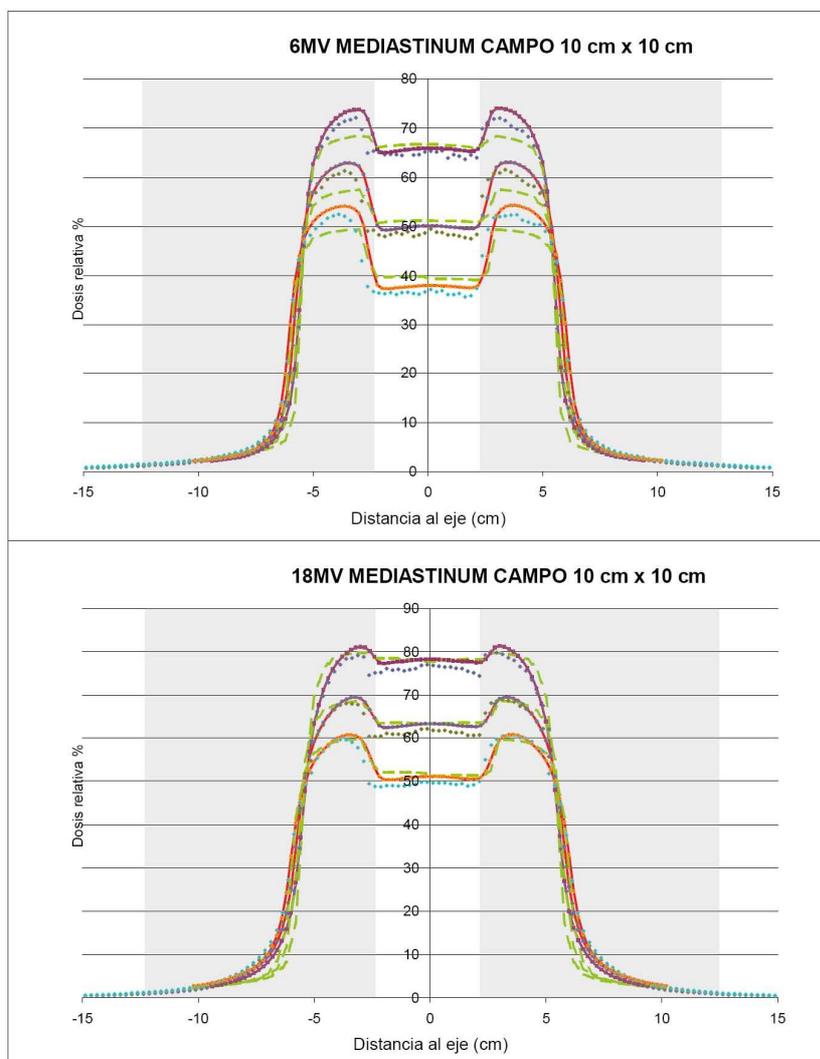


Fig. 7.3. Perfiles en el maniquí MADIASTINUM para 6 MV (arriba) y 18 MV (abajo) a 3 profundidades distintas, del perfil más alto al más bajo 10, 15 y 20 cm. La zona sombreada corresponde a la zona de la heterogeneidad LUNG. Se superponen las curvas de los tres algoritmos utilizados de acuerdo con la notación:

● MC
■ SCC
▲ PR

Por razones de claridad en la imagen se han omitido las barras de incertidumbre del algoritmo MC.

En la figura 7.4 se presentan los resultados para la geometría SLABS. En este caso se presentan, además, dos perfiles para 6 y 18 MV dentro de la heterogeneidad de pulmón. En ellos se puede apreciar que la penumbra del haz se modela adecuadamente comparando MC y SCC, dado que uno de los efectos conocidos de las heterogeneidades de baja densidad, debido al mayor recorrido de los electrones secundarios, es el ensanchamiento de las penumbras. SCC se muestra capaz de modelar este efecto no contemplado en el algoritmo PR.

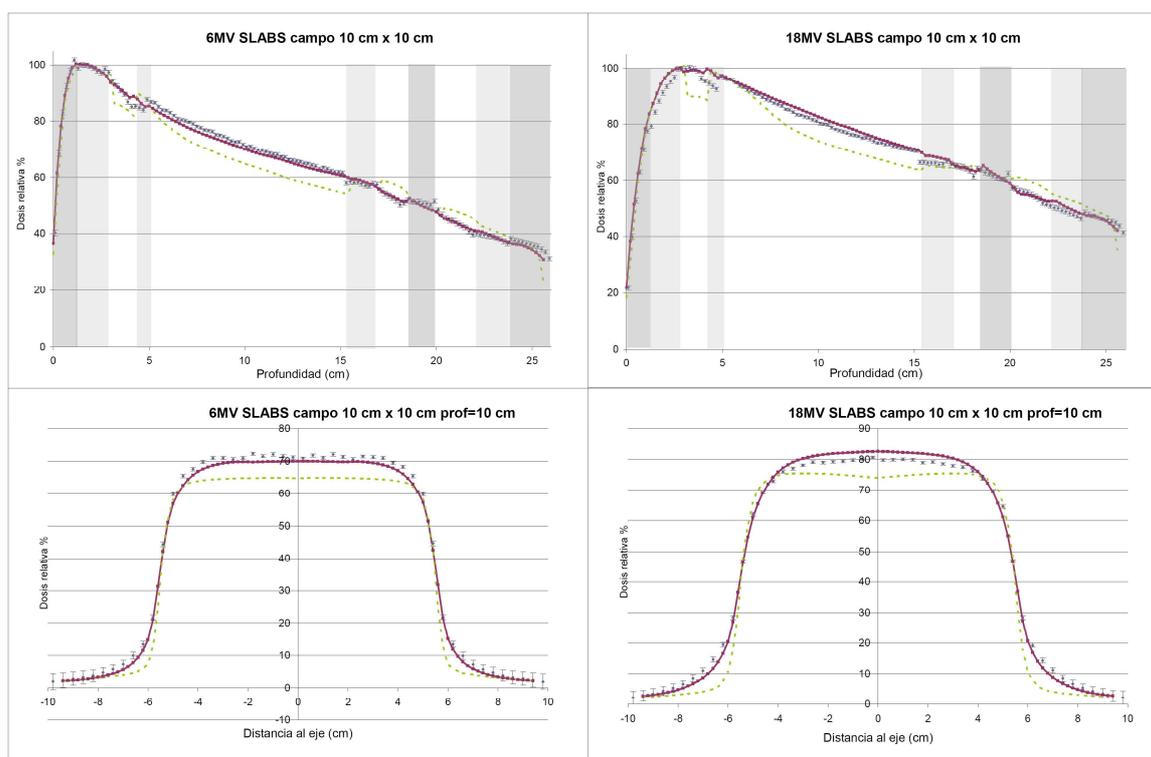


Fig. 7.4. Porcentajes de dosis y perfiles a 10cm de profundidad (dentro de pulmón) en el maniquí SLABS para 6MV (izquierda) y 18MV (derecha). Se superponen las curvas de los tres algoritmos utilizados de acuerdo con la notación: **• MC** **— SCC** **..... PR**

Los resultados obtenidos en la comparación entre SCC y MC son comparables a los de trabajos previos (74–76) y permiten utilizar el algoritmo SCC con similares características a otros utilizados para tratamiento.

7.2. Ajuste de la unidad de tratamiento en el cálculo por Monte Carlo

El ajuste de los parámetros que definen la fuente de la unidad de tratamiento dio como resultado una energía media de los electrones incidentes de 6,522 MeV para la unidad de 6MV. La FWHM para la distribución gaussiana del haz circular de electrones quedó fijada en 1 mm y el ángulo medio de dispersión en 1 grado (sirve para ajustar los perfiles de campo más ancho).

En la figura 7.5 se muestran los PDP para campos 3 cm × 3 cm, 5 cm × 5 cm, 10 cm × 10 cm, 20 cm × 20 cm, y 30 cm × 30 cm, junto con los correspondientes resultados de la simulación Monte Carlo. Las incertidumbres de la simulación son menores del 1%, aunque no se muestran por motivos de claridad en las curvas.

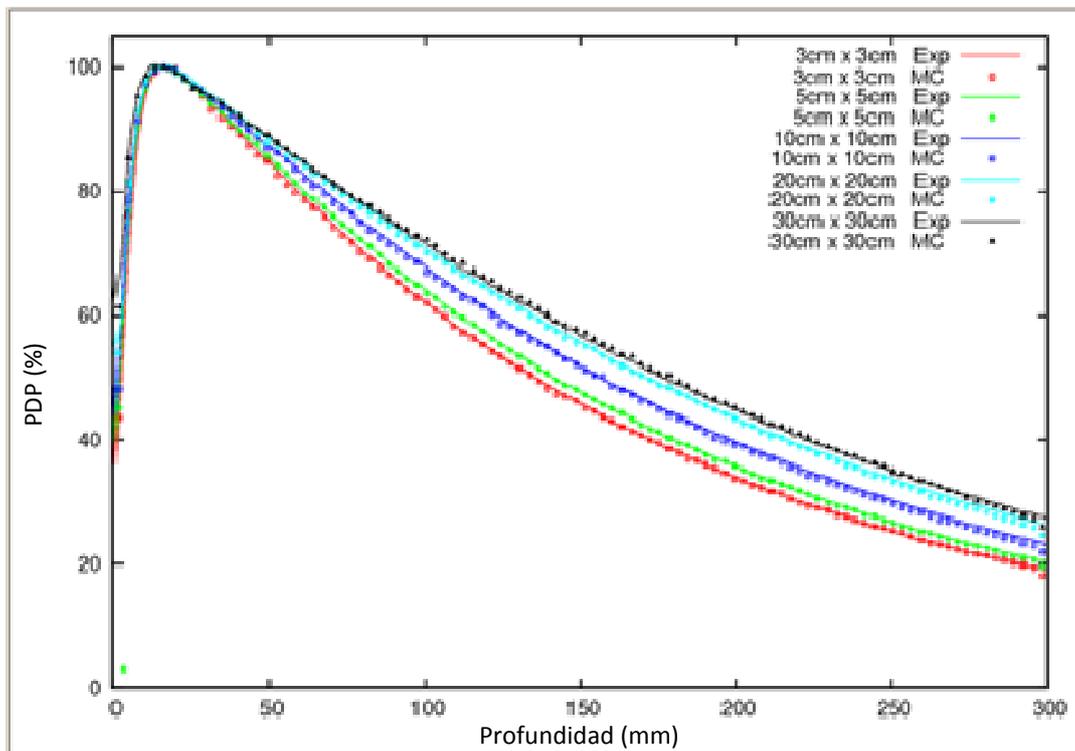


Figura 7.5.- PDPs para diferentes campos obtenidos experimentalmente (Exp) y por medio de simulaciones Monte Carlo (MC).

En la figuras 7.6, 7.7 y 7.8 se muestran los perfiles de dosis medidos y simulados para los mismos campos a la profundidad de 1,5 cm, 10 cm y 20 cm. Existe buen acuerdo entre las simulaciones y los resultados experimentales. En términos de comparación gamma, el 100% de los puntos mostrados cumplen el criterio 1%/1mm tomando como dosis de normalización

la que se deposita en el máximo de cada campo. Las incertidumbres en la simulación MC están en torno al 1% dentro de campo.

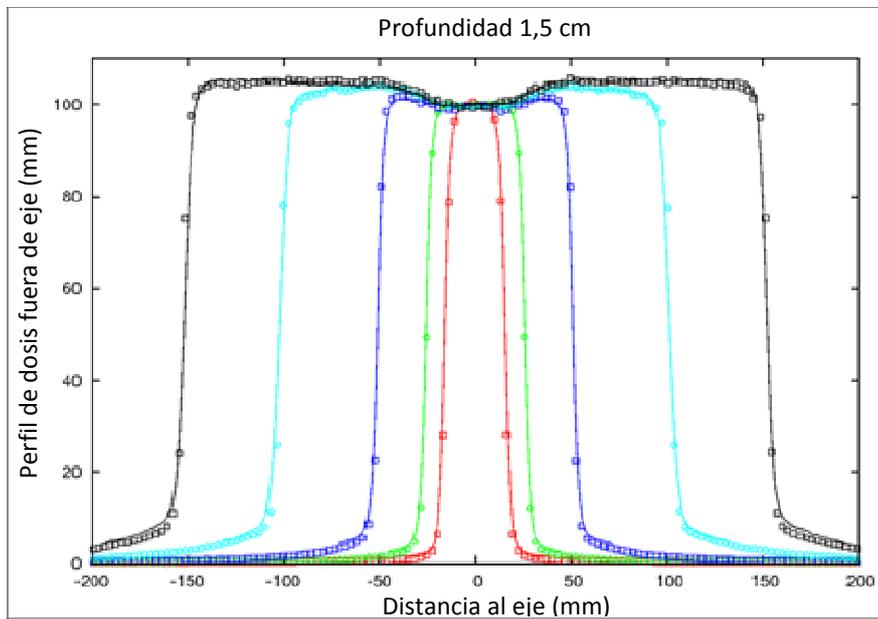


Figura 7.6. Perfiles transversales para diferentes tamaños de campo a 1.5 cm de profundidad y 100 cm de distancia fuente-superficie. Las líneas sólidas corresponden a medidas realizadas con diodo para campos pequeños y cámara de ionización para campos grandes y los puntos a las simulaciones MC.

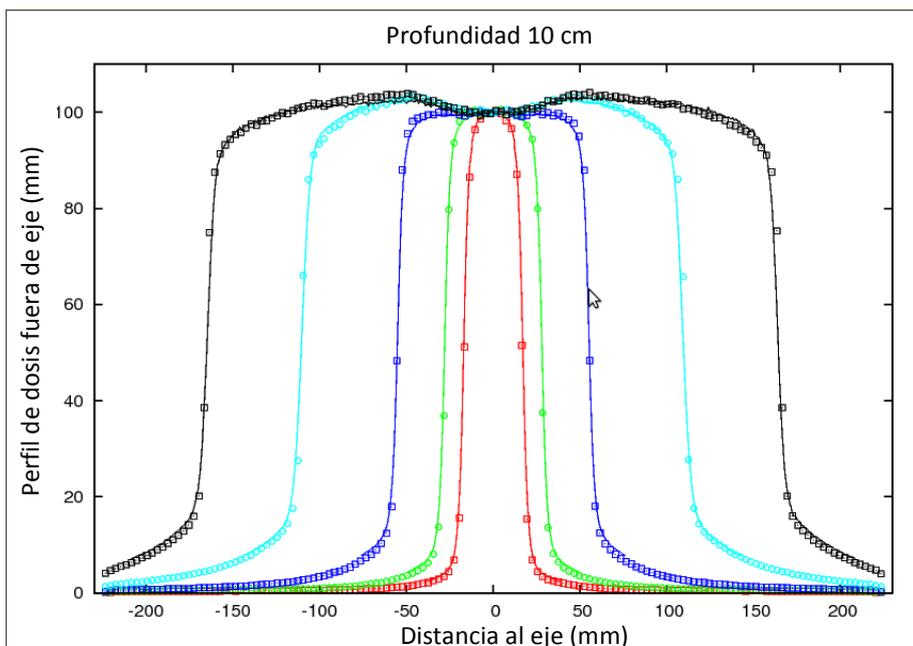


Figura 7.7. Igual que la figura 7.6 pero a 10 cm de profundidad.

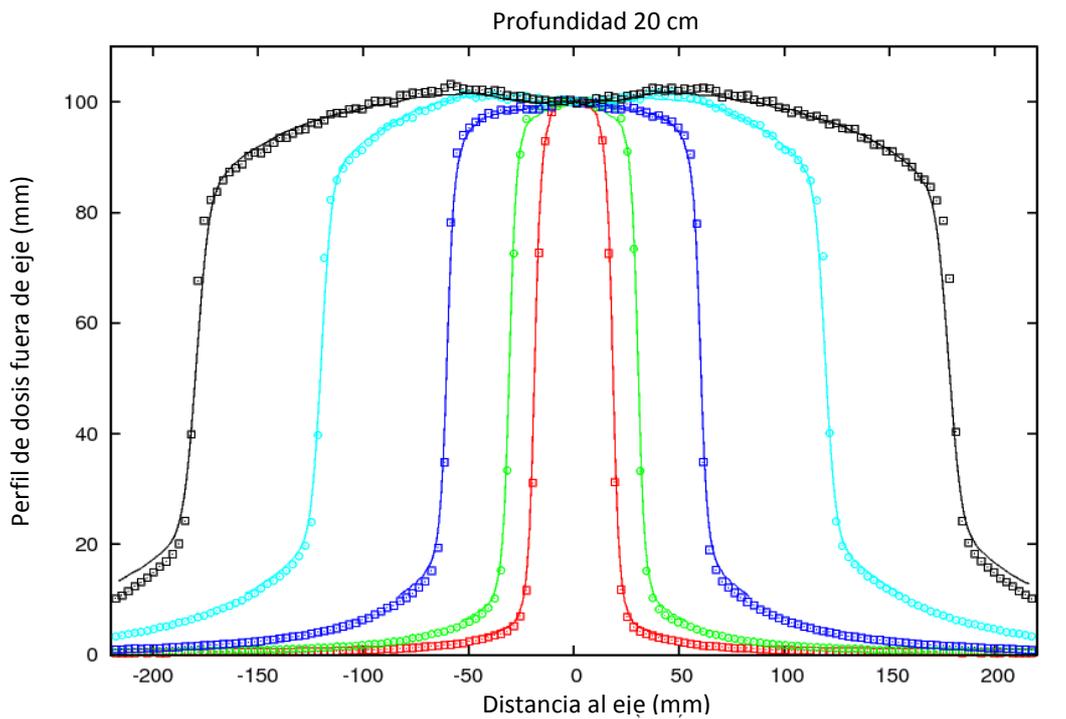


Figura 7.8. Mismo que la figura 7.6 pero a 20 cm de profundidad.

En la figura 7.9 se presentan los factores campo calculados por Monte Carlo y comparados con los medidos. Se observa que coinciden dentro de las incertidumbres presentes en el proceso.

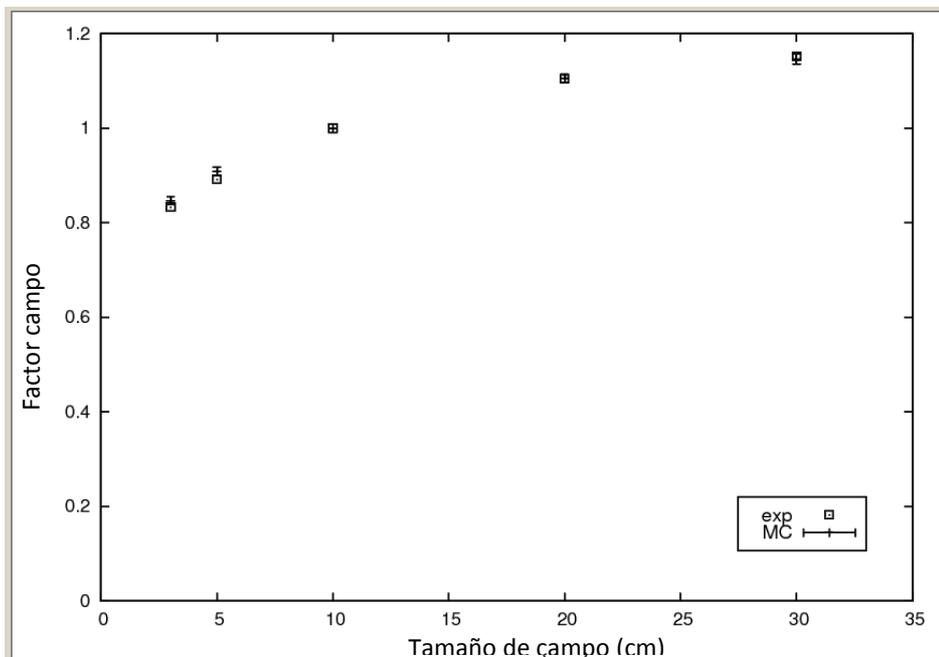


Figura 7.9. Factores campos medidos (exp) y calculados por MC.

7.3. Influencia del estado del acelerador en el resultado final

En este apartado se resumirá la influencia de la dosis por UM proporcionada por el acelerador y del estado del MLC en las medidas de IMRT realizadas. Ambas incertidumbres están medidas y pueden ser expresadas en términos de variación en los histogramas dosis-volumen.

7.3.1. Incertidumbre en la dosis por UM del acelerador

Las medidas con cámara de ionización en la verificación de IMRT están corregidas por la variación en la dosis por UM del acelerador y, por tanto, son relativas al valor de cGy/UM introducido en el SPT. En cuanto a las medidas con el sistema MatriXX+Compass, la calibración geométrica y en dosis del detector se realiza mensualmente por recomendación del fabricante. La reproducibilidad semanal y mensual se midió durante la puesta en marcha del acelerador Oncor. Como resultado se obtuvo que la incertidumbre asociada a la variación mensual del acelerador es del $\pm 1,2\%$ ($k=2$). Esa estabilidad se ha comprobado a lo largo de la vida útil del acelerador.

7.3.2. Incertidumbre en el posicionamiento del MLC

Los resultados obtenidos en el test de bandas adyacentes muestran la alta sensibilidad del test. En la figura 7.10 se presentan los resultados para la variación del parámetro PPBC en 7 medidas consecutivas, en función de los desplazamientos provocados en el banco de láminas X1. El promedio para un desplazamiento de +2mm es 23% y para -1mm es -18%. El parámetro PPBC, que caracteriza el estado de las láminas del MLC, se ha normalizado a 0 en la situación de desplazamiento 0 mm para que se aprecie el efecto exclusivamente debido al movimiento forzado de las láminas.

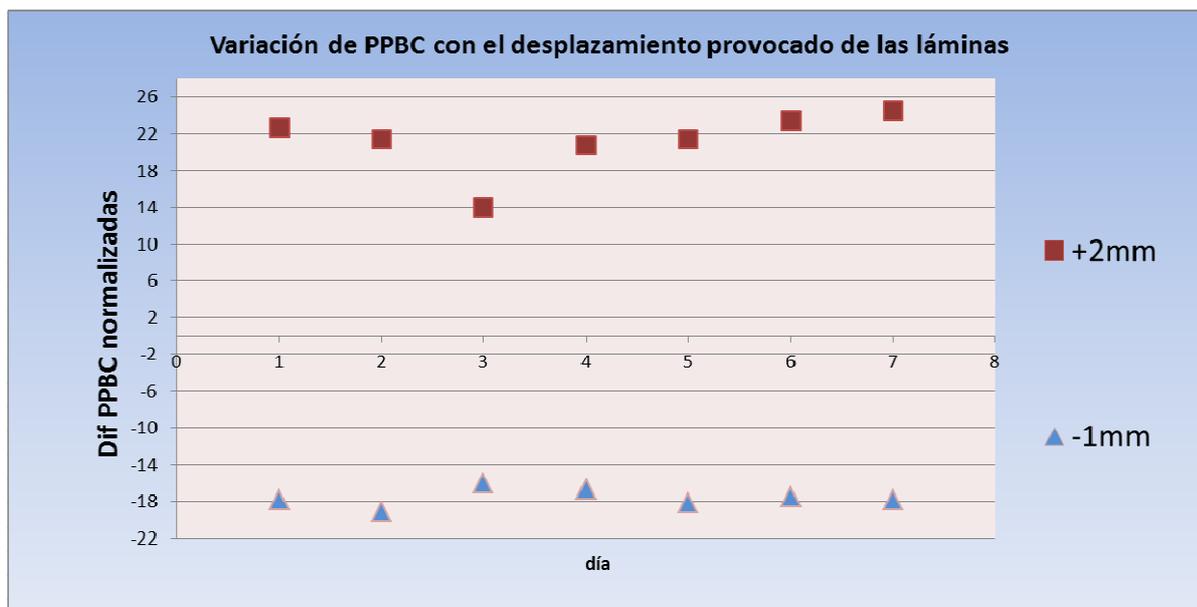


Figura 7.10. Medidas consecutivas de PPBC con un desplazamiento provocado en el banco X1 de 2mm (puntos rojos) y de -1mm (puntos verdes). PPBC está normalizado a 0 en la situación de desplazamiento 0 mm. El promedio para +2mm es 23% y para -1mm es -18%.

Del análisis del test se pueden extraer otros resultados adicionales:

- La prueba muestra una alta reproducibilidad en tres irradiaciones sucesivas. En la siguiente tabla (7.I), se presenta el valor del parámetro PPBC obtenido en medidas consecutivas, tanto antes como después de realizar una autocalibración del MLC.

Tabla 7.I. Reproducibilidad de PPBC. La autocalibración vuelve a colocar el parámetro PPBC en el valor que tenía cuando se calibró el MLC.

Valores de PPBC en test sucesivos	PPBC antes de autocalibración	PPBC después de autocalibración
Medida 1	91,1	104,5
Medida 2	90,5	104,2
Medida 3	90,4	103,9
Media	90,7	104,2
Desv estándar	0,4	0,3

- Los resultados son repetibles para cada situación del brazo (ver figura 7.11)

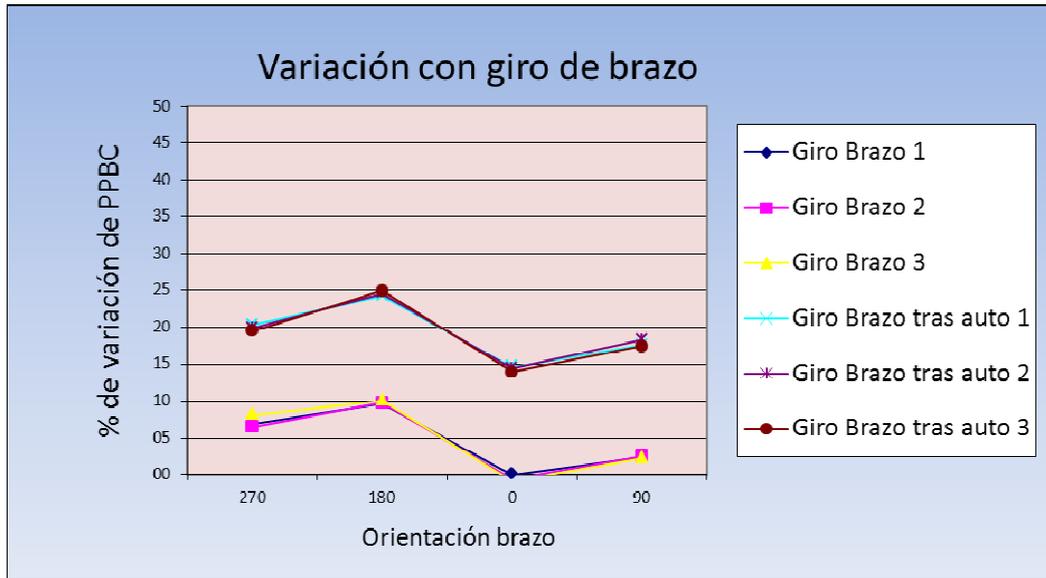


Figura 7.11. Variación del parámetro PPBC con el giro de brazo, antes y después de realizar una autocalibración. PPBC se mueve en valores relativos con respecto a la situación inicial, 0 equivale a 90% de PPBC.

En el haz de 180° se encuentra un máximo en el parámetro PPBC. Se observa que a 180 grados PPBC es mayor que 0, en torno a un 15%. La posible causa para este incremento es que la gravedad sitúa las láminas unas décimas de milímetro hacia afuera en eje perpendicular al haz, y que ese movimiento no sigue la focalización del colimador. El cambio en el valor de PPBC con el giro de brazo indica que los desplazamientos que se producen en las láminas son inferiores a 1 mm de acuerdo con los resultados de la figura 7.10.

7.3.2.1 Influencia de la incertidumbre en el posicionamiento del MLC en el HDV correspondiente al PTV:

Se muestran los resultados obtenidos para un plan de tratamiento IMRT seleccionado. En cada uno de los segmentos que constituyen los haces de tratamiento, se ha movido en el SPT el banco de láminas X1 en un caso +2mm y en el otro -1mm. Se han comparado los HDV obtenidos en cada caso, y, en particular, la dosis media en el PTV principal. Se observa un aumento del 2% en el primer caso y una disminución del 0,7% en el segundo (figura 7.12).

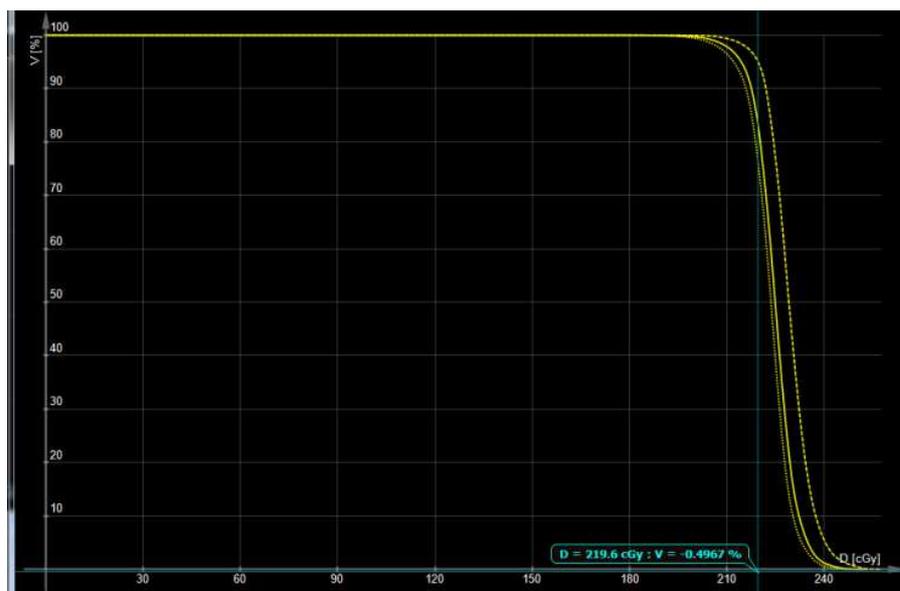


Figura 7.12. Histogramas obtenidos en el SPT del PTV principal con la situación provocada de desplazamiento de láminas. La línea de puntos gruesos corresponde a +2mm y la de puntos finos a -1mm. La dosis media en el PTV cambia un +2% en el primer caso y un -0,7% en el segundo.

Por tanto, un desplazamiento global de las láminas produce, a su vez, un desplazamiento en el HDV del PTV principal del plan.

7.3.2.2 Medidas en casos reales:

Se ha demostrado que el parámetro con el que se puede caracterizar el estado del MLC frente a la variación de la dosis media medida en los PTVs principales del plan es el % promedio de picos (PPBC), procedente del test de bandas contiguas. La comparativa de los histogramas de cinco de los pacientes estudiados, medidos en diferentes días y con diferentes resultados en el parámetro PPBC controlado diariamente, muestra diferencias en dosis compatibles con las variaciones en el posicionamiento del MLC encontradas en el test precedente en el que se provocaban desviaciones sistemáticas de las láminas. En la figura 7.13, se presentan la variación de la diferencia de dosis media en los PTVs obtenida a partir de las medidas realizadas con el sistema Compass en función del estado del colimador, caracterizado por el parámetro PPBC, para los cinco pacientes mencionados.

Como se deduce del párrafo anterior, en este caso las variaciones en el MLC no se han provocado sino que corresponden a diferentes situaciones del mismo a lo largo del tiempo,

variando PPBC a lo largo del tiempo en un rango de aproximadamente el 20%. Dada la forma particular de los perfiles del acelerador y las características del test, un valor medio para el parámetro es el 103%, que correspondería al promedio de dosis relativa al centro que se deriva de un solapamiento perfecto de todas las bandas adyacentes para configurar el campo rectangular 20 cm x 28 cm al que equivale la suma de las bandas. Nótese que en la figura 7.13 al valor 103% no le corresponde una desviación del 0%, dado que en la verificación con CMPM influyen más factores. Se observa en dicha figura que la tendencia de la variación es creciente y queda bien ajustada por una recta en todos los casos, si bien la pendiente es diferente entre pacientes.

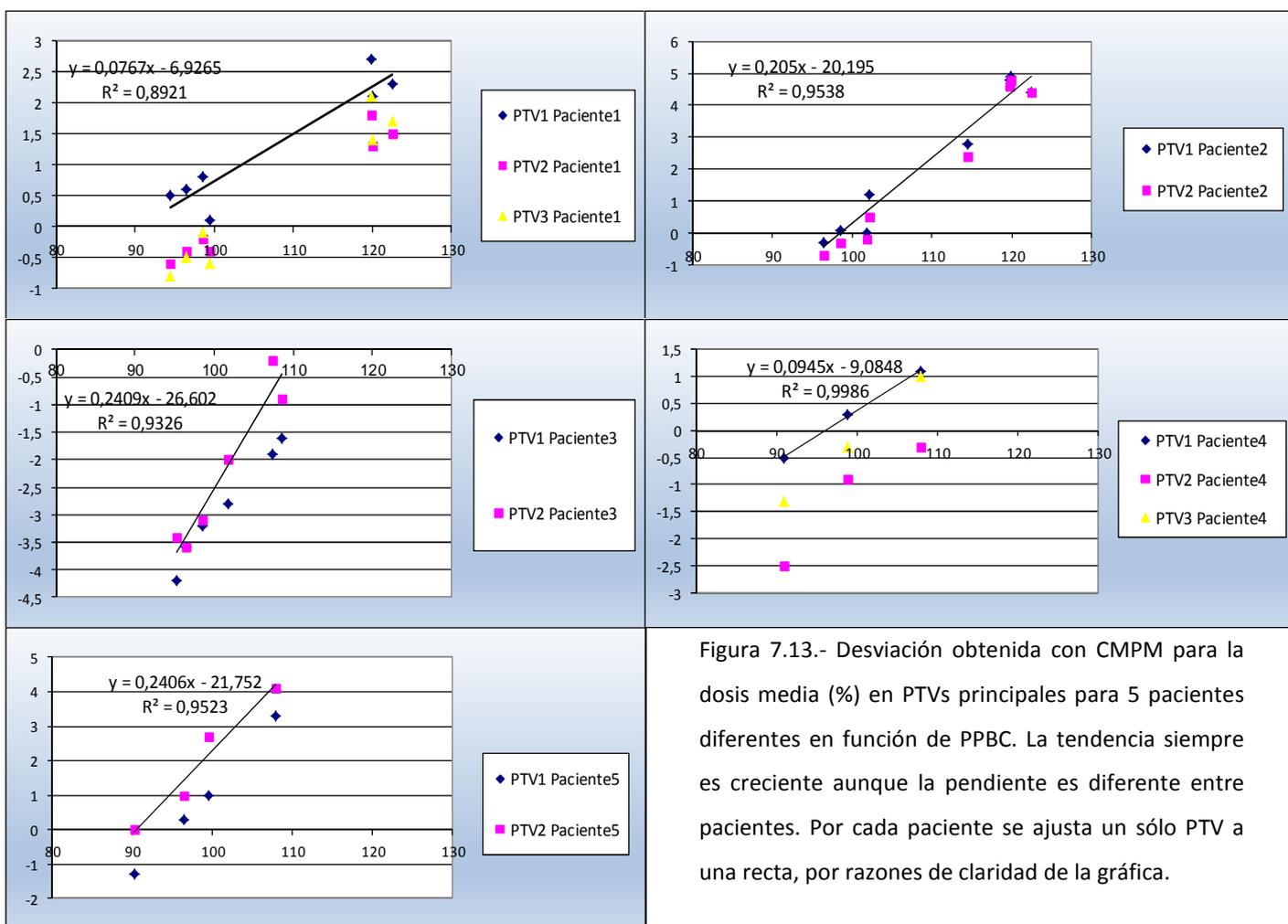
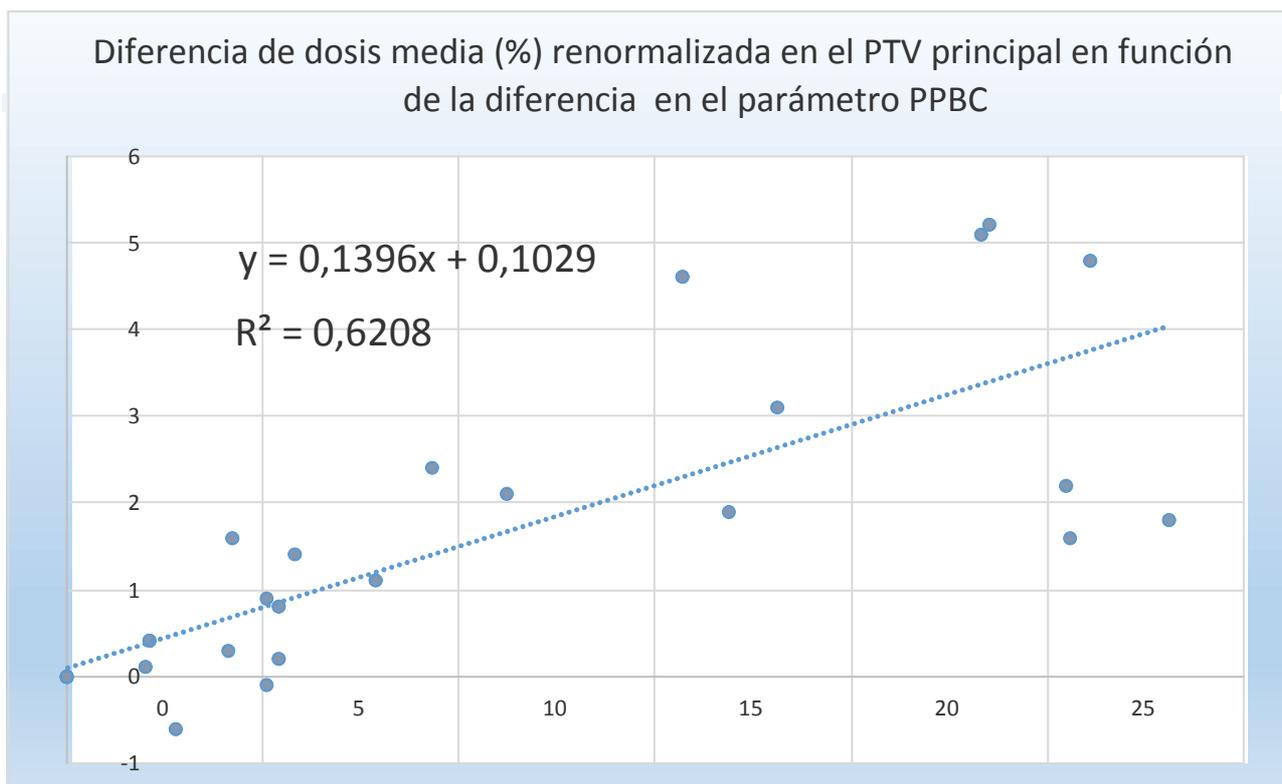


Figura 7.13.- Desviación obtenida con CMPM para la dosis media (%) en PTVs principales para 5 pacientes diferentes en función de PPBC. La tendencia siempre es creciente aunque la pendiente es diferente entre pacientes. Por cada paciente se ajusta un sólo PTV a una recta, por razones de claridad de la gráfica.

Por tanto, para obtener una estimación de la incertidumbre asociada a la calibración del MLC no podemos usar los datos crudos de diferencia ya que la diferencia en dosis media en el PTV varía dado un mismo resultado del test bandas para pacientes distintos. Es posible

normalizar todas las variaciones tanto en promedio de picos como en diferencia de dosis al valor más bajo para cada paciente. De este modo el primer punto de cada una de las gráficas de la figura 7.13 se convierte en el origen de la nueva una nueva gráfica y podemos construir una sola gráfica para todos los pacientes. Para ella es posible ajustar una recta de cuya pendiente poder extraer el valor promedio $\Delta\%$ dosis/ $\Delta\%$ PPBC. Dicha gráfica se presenta en la figura 7.14. El coeficiente de determinación del ajuste es 0,6208 y la pendiente 0,14.



30

Figura 7.14.- Diferencia de dosis media renormalizada en el PTV principal en función del incremento del PPBC para todos los pacientes estudiados. La pendiente de la recta es 0,14 y el error típico de 0,021. PPBC está renormalizado al valor más bajo para cada paciente.

Lo que nos permite inferir que, si el resultado del test bandas contiguas está en el entorno del $103\pm 10\%$, la dosis media en los PTVs principales de un paciente variará en promedio un $\pm 1,4\%$ debido a los efectos de incertidumbre en el posicionamiento de las láminas para el colimador Optifocus.

En los pacientes estudiados en este trabajo, el estado del colimador durante el verificación con el sistema Compass presentó un parámetro PPBC con la distribución de valores que podemos ver en la figura 7.15. El 85% de los valores de la variable se

encontraron en el intervalo 103 ± 10 %, y el 95% de los valores en el intervalo [90%,116%] lo que dejó la incertidumbre estimada en la dosis que recibe el PTV principal debida al estado del MLC ($k=2$) en torno al ± 2 %, de acuerdo con la figura 7.14.

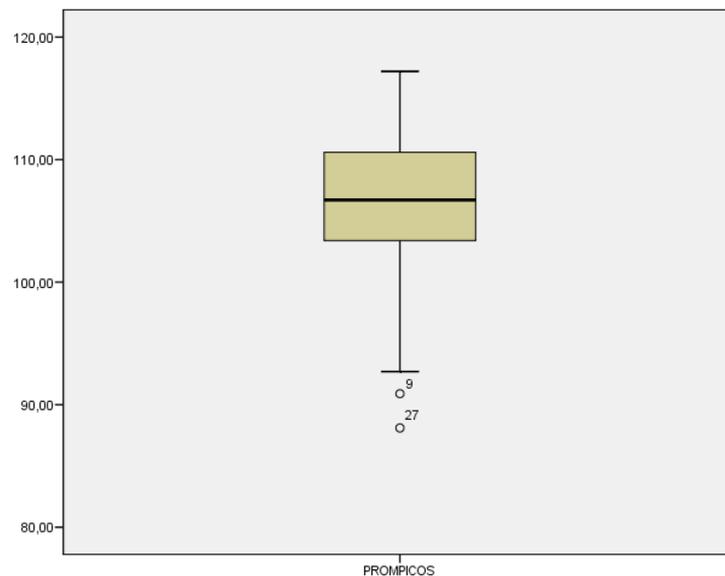


Figura 7.15.- Diagrama de cajas para el parámetro PPBC de los 58 pacientes del trabajo cuando se midieron con Compass. La media es 105,9% y la desviación estándar es 6,4%.

Cabe preguntarse a la vista de la influencia del estado del colimador en el resultado de la verificación del tratamiento, si, además, es el único factor que determina las diferencias que se encuentran entre la medida (CMPM) y la predicción por parte de CMPC. En este sentido podemos comparar las diferencias que se producen en las desviaciones en dosis media del PTV principal con respecto a SCC, al obtenerse con CMPM y con CMPC. Dado que ambos utilizan el mismo algoritmo para calcular en el paciente, las diferencias entre ellos proceden de las diferencias entre la respuesta predicha del detector y la respuesta medida. Si el factor principal de influencia fuera el estado del MLC, existiría una correlación fuerte para el PTV principal entre las diferencias CMPC-SCC y las diferencias CMPM-SCC, si corregimos CMPM previamente teniendo en cuenta el parámetro PPBC que había el día de la medida. La figura 7.14 nos permite calcular cuál es el factor de corrección para CMPM en lo que respecta al PTV principal. En la figura 7.16 se han representado las diferencias en dosis media del PTV principal entre CMPM corregido por PPBC y SCC, frente a las mismas diferencias entre CMPC y SCC, para los 58 planes estudiados. La correlación existe y el valor del coeficiente de

determinación indica que el 60% de la variación es explicable a través del parámetro PPBC, lo que indica que es el principal parámetro que influye en el resultado de la medida frente al cálculo previsto por el algoritmo CMPC para la dosis media del PTV principal. La significación bilateral de la correlación es 0. Otro factor que lógicamente influye en la diferencia de dosis media en el PTV principal es la dosis/UM del ALE el día de la medida, como se ha visto en 7.3.1. En el siguiente apartado se profundiza más en las diferencias que existen entre CPM y CMPC.

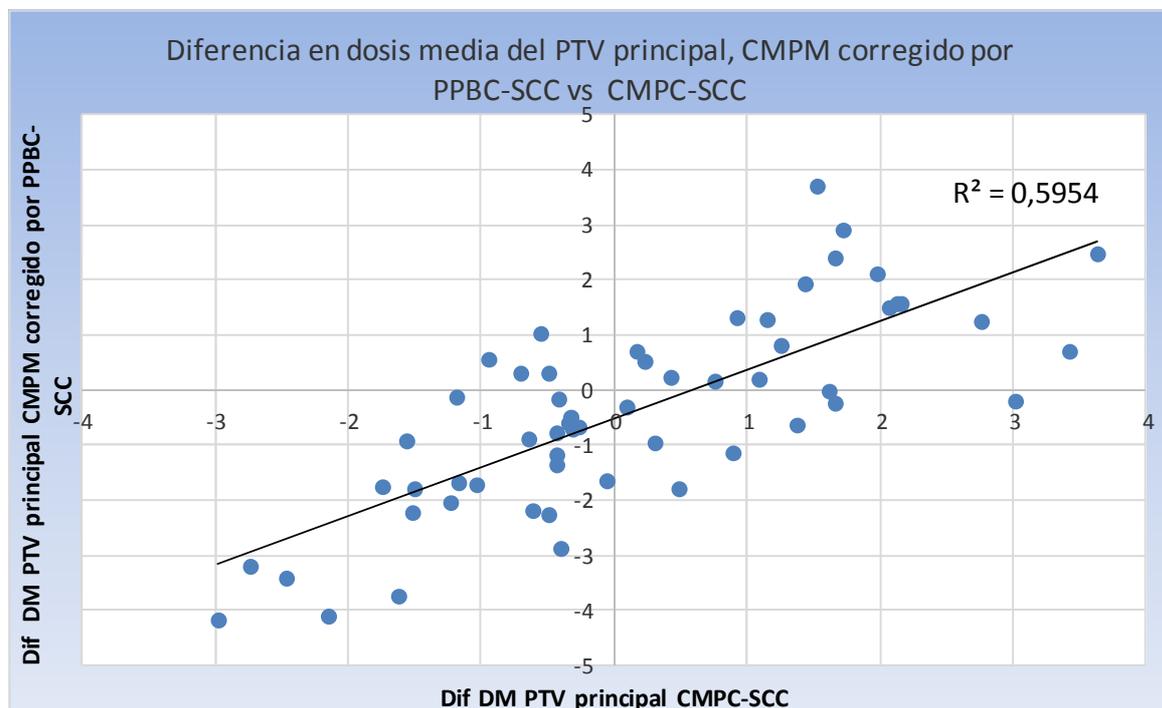


Figura 7.16.- Correlación entre las diferencias en la dosis media del PTV principal para el algoritmo CPM corregido por PPBC con respecto a SCC frente a al algoritmo CMPC también con respecto a SCC.

7.3.3. Evaluación de la validez del escalado proporcional entre CMPC y CPM

Tal como se ha dicho en el apartado 6.1, las diferencias entre la fluencia predicha y la fluencia reconstruida se deben al estado del acelerador en el día de la medida, en particular del MLC, ya que en IMRT estática no influyen otros parámetros dinámicos, como pudiera ser la tasa o la velocidad de las láminas y del brazo. También existe una diferencia residual entre

ambas fluencias que se debe al modelo de reconstrucción. Estas diferencias han sido evaluadas por Godart (7) y pueden tener influencia en el resultado final debido al tamaño del núcleo de reconstrucción que se usa al determinar la fluencia reconstruida (ver ecuación 3.4). Este artefacto puede crear inexactitudes cerca de los bordes de campo de segmentos individuales y puede conducir a que la dosis máxima en zonas calientes sea sobreestimada. Este efecto podría estar también presente en el análisis de los histogramas.

Se ha procedido al escalado de los histogramas obtenidos con CMPC para ver si, con una operación sencilla, son equiparables a los obtenidos con CPM. La operación sencilla definida en 6.1 es el escalado proporcional a partir de los datos de dosis media en el PTV principal y la dosis prescrita. La primera métrica con la que se han comparado los resultados es con la métrica de diferencias entre el algoritmo CPM y el algoritmo CMPC pero, en este caso, se ha penalizado igualmente tanto la pérdida como el aumento de dosis en los puntos de evaluación usados para fijar las restricciones de la tabla 6.I., ya que se trataba de ver cómo de diferentes eran los histogramas, no si el plan era aceptable o no. Así, a partir de los histogramas se han extraído las dosis en los puntos de evaluación y se ha optado en este trabajo por considerar que los histogramas eran similares si la diferencia entre los valores obtenidos con CPM y CMPC eran inferiores al 3% de la dosis prescrita en el PTV principal. Como resultado se obtuvo que 11 de los 58 planes no pasaron la prueba de comparación, cuando se esperaba que todos la pasaran si se quería dar por válida la posibilidad de escalar los histogramas. El problema está principalmente en los OARs. Sólo 2 de los 11 planes no pasan la prueba para el PTV, es decir, los histogramas son muy similares en cobertura, dosis media y dosis máxima. Los OARs sin embargo hacen que los 11 planes no pasen la prueba, por lo que se deduce que es en el posicionamiento particular de las láminas en el borde de cada segmento donde está la diferencia que no puede ser tenida en cuenta con un escalado proporcional. Dicho de otro modo, la contribución principal a las diferencias en los PTV entre CMPC y CPM viene dada por el estado del MLC a través del parámetro PPBC pero esto no es extrapolable a los OARs. Esto se puede apreciar en la figura 7.17, donde se observa la dispersión de los valores de las diferencias para cada órgano. Tras el escalado, tronco, quiasma, vías ópticas y cristalinos conservan diferencias por encima del 3% en algunos casos, por lo que dichas diferencias no pueden resolverse con una operación sencilla y dependen del comportamiento particular de determinadas láminas.

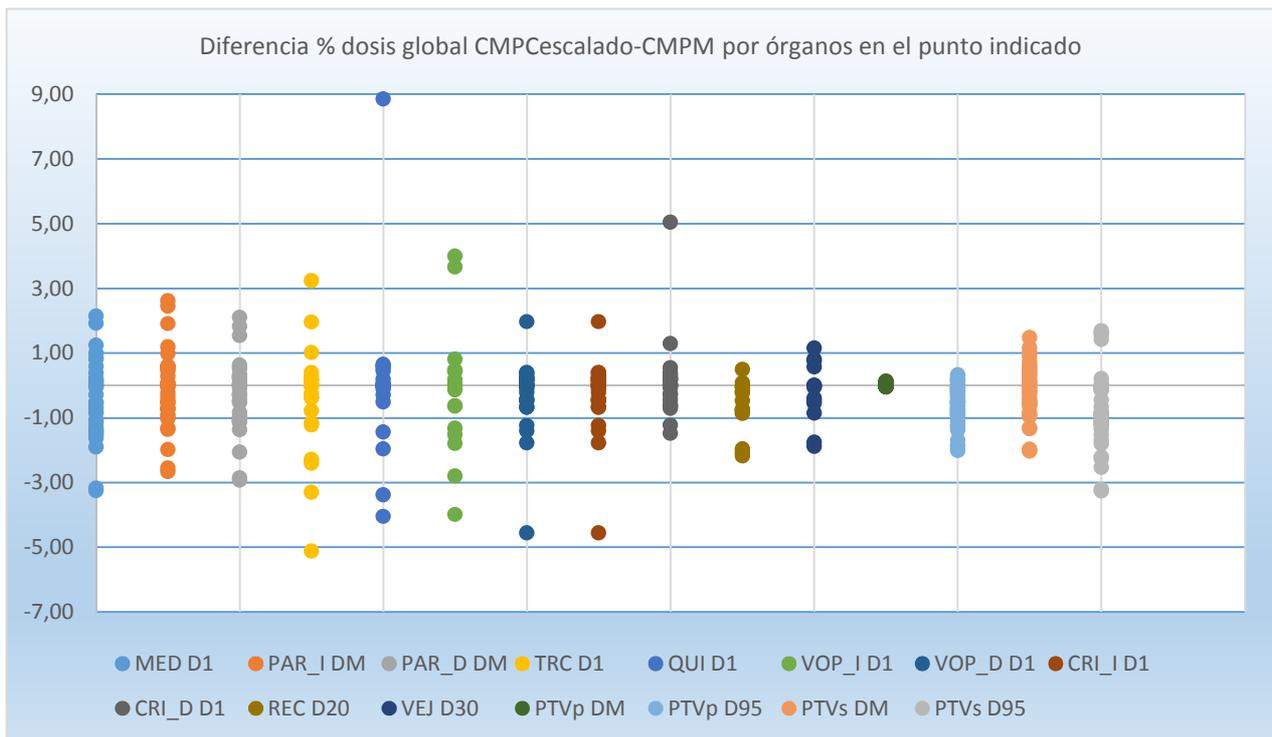


Figura 7.17.- Dispersión de datos por órgano y valor comparado entre CMPC escalado y CMPM. Si todos los valores estuvieran entre 3% y -3% CMPC escalado y CMPM podrían considerarse equivalentes desde el punto de vista de la métrica de diferencias en dosis.

Este resultado, obtenido comparando resultados para CMPM y CMPC, limita la validez del escalado que se puede realizar sobre los histogramas obtenidos con MC. Si los valores de la figura 7.17 estuvieran en un intervalo más estrecho, especialmente para los OARs, se podría concluir que una operación sencilla del tipo escalado podría proporcionar un resultado MCesc que pudiera ser considerado de referencia.

Aunque no sea posible establecer un sistema clasificador de referencia absoluto, la utilidad de MCesc está clara, por un lado proporciona otro algoritmo basado en un cálculo redundante más exacto que CMPC, y por otro lado, incorpora las posibles desviaciones debidas a la medida en cuanto al estado global del MLC y a la variación de la dosis por UM del ALE. Por tanto, en los análisis de los dos siguientes apartados se van a usar tanto MCesc como CMPM, como sistema clasificador de referencia para poder extraer conclusiones acerca de las métricas propuestas.

7.4. Análisis ROC de las métricas evaluadas

Los resultados en este apartado pretenden establecer una clasificación de las diferentes métricas empleadas con respecto a un sistema clasificador de referencia, escogido de manera alternativa entre aquellas que provienen de CMPM o MCesc según el algoritmo sobre el que se aplican. El análisis ROC (18) permite, además, establecer valores de corte para las métricas que se analizan. A través de las curvas ROC se pretende evaluar cuál de las métricas definidas es mejor predictor de la clasificación realizada con la métrica de diferencias, bien sea evaluada a partir de los HDV obtenidos de CMPM con respecto a los HDV obtenidos de SCC, o bien sean los HDV obtenidos de MCesc los que se comparan con los provenientes de SCC. El primer paso para proceder al análisis ROC es construir las variables binarias descritas en 6.3.: PASA_CMPM, PASA_CMPMOBJ, PASA_MCesc, PASA_MCescOBJ y PASA_MC para ser usadas como variables binarias clasificadoras de referencia.

7.4.1. Análisis de las funciones FI, FIP, FGO y FG

El comportamiento de estas funciones para su uso como sistemas clasificadores se va a estudiar comparando con las variables mencionadas en el párrafo anterior.

Inicialmente, se buscan los factores de ponderación que funcionan mejor con la función de incumplimiento ponderado (FIP), determinándose que conjunto conduce a la mejor predicción. En la siguiente tabla (7.II) se presentan los valores seleccionados como pesos para la FIP.

	Conjunto1	Conjunto2	Conjunto3	Conjunto4	Conjunto5	Conjunto6
wptv1d95	1	1	0,001	1000	0,5	1
wptv1d2	1	1	0,001	1000	0,5	1
wtrcd1	0,04	0,4	1	0,1	2	1
wmedd1	0,075	0,75	1	0,1	2	1
wrd20	0,05	0,5	1	0,1	2	1
wrd30	0,05	0,5	1	0,1	2	1

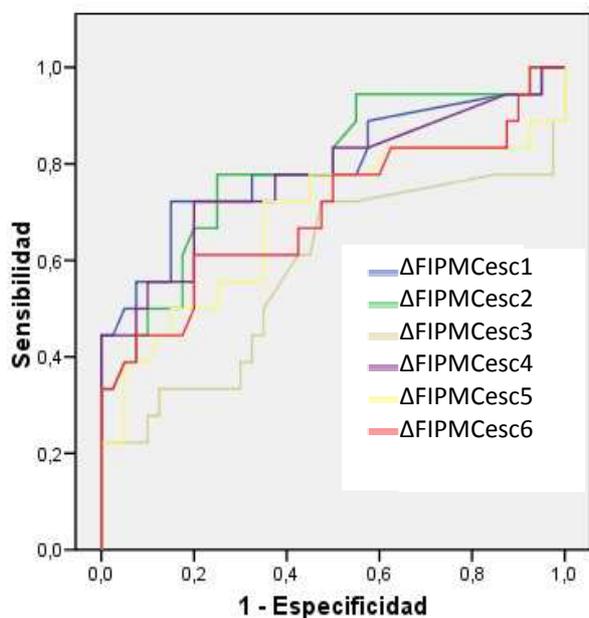
wpardm	0,05	0,5	1	0,1	2	1
wcrd1	0,025	0,25	1	0,1	2	1
wvopd1	0,04	0,4	1	0,1	2	1
wquid1	0,04	0,4	1	0,1	2	1

Cada conjunto ha sido escogido para representar una combinación diferente de ponderaciones, oscilando desde tener en cuenta sólo los PTVs hasta sólo los OARs para determinar el valor de la FIP. Para investigar qué conjunto de valores proporciona la mejor predicción, podemos trazar las curvas ROC correspondientes a la predicción que realiza la función FIP de los resultados de la métrica de diferencias en dosis, tanto usando MCesc como usando CPM. Las variables binarias asociadas son PASA_CPM y PASA_MCesc. Se usan como clasificadoras las variables continuas ΔFIP_{MCesc} y ΔFIP_{CPM} y los resultados se presentan en la figura 7.18. En ella se muestra, además de las curvas ROC, una tabla con la AUC que se obtiene para cada conjunto de pesos considerado.

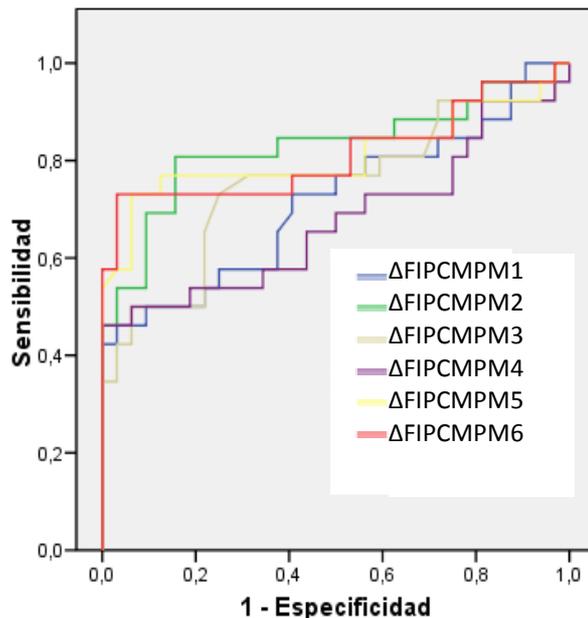
Por tanto, y a la vista de los resultados para MCesc y CPM, el segundo conjunto de pesos parece el más adecuado para ser usado en la función incumplimiento ponderado como predictor ya que proporciona el área bajo la curva más grande tanto si se usa como referencia MCesc como si se usa CPM. Nótese que la función incumplimiento es un caso particular de la función incumplimiento ponderado con todos los pesos =1.

Una vez seleccionado el mejor conjunto de pesos para la ponderación podemos comparar cómo se comportan todas las funciones definidas como predictoras en 6.2.2, FI, FIP, FGO y FG, construyendo curvas ROC donde las cuatro variables continuas son usadas como predictoras de las siguientes variables binarias: PASA_CPM, PASA_MCesc, PASA_CPMOBJ y PASA_MCescOBJ, distribuciones binarias obtenidas a partir de la métricas de diferencias en dosis y la métrica de diferencias en dosis basada en objetivos.

Curva ROC PASA_MCesc



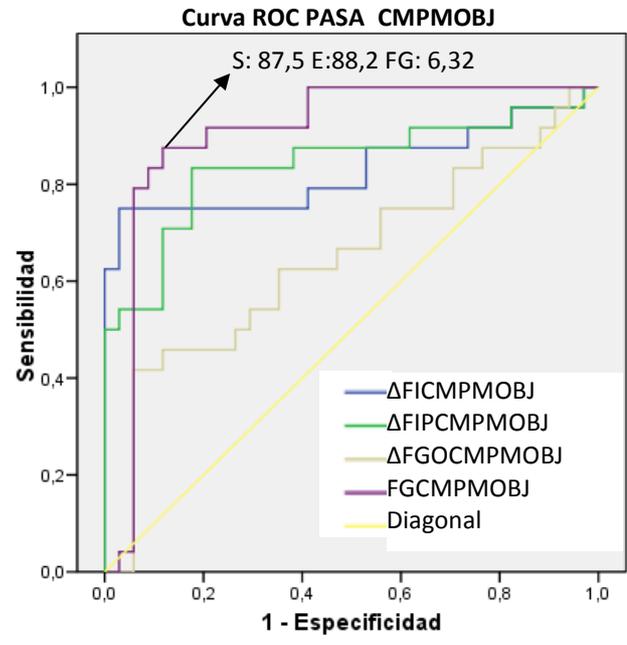
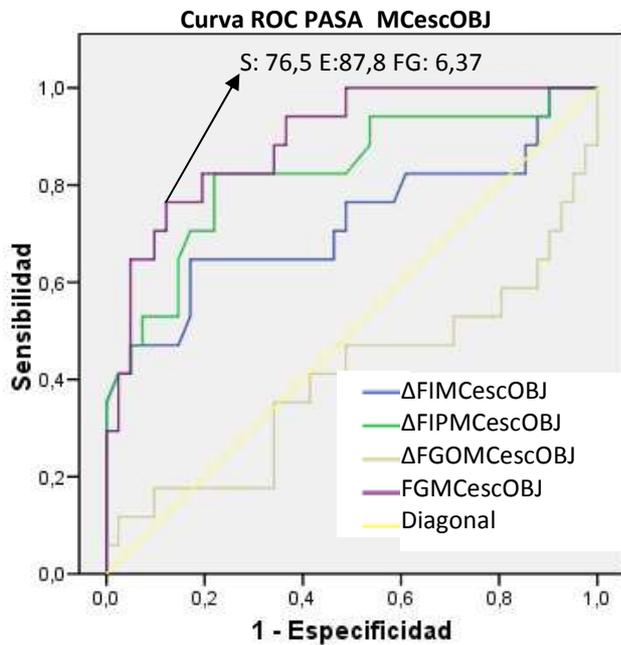
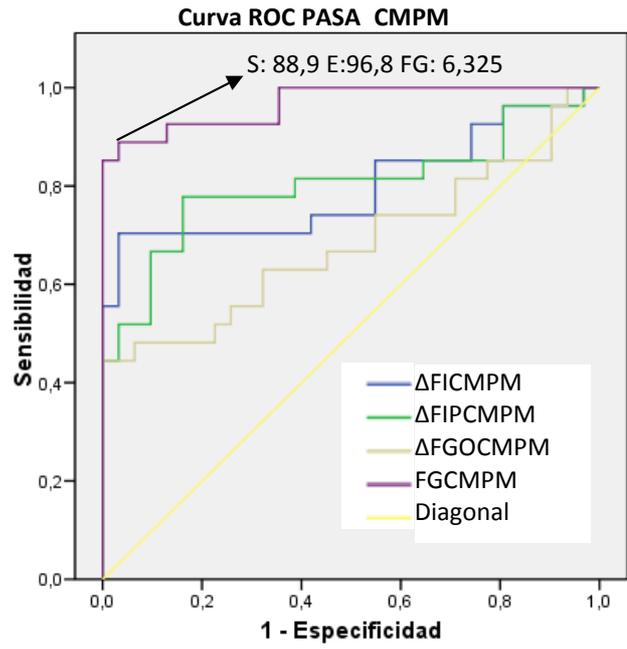
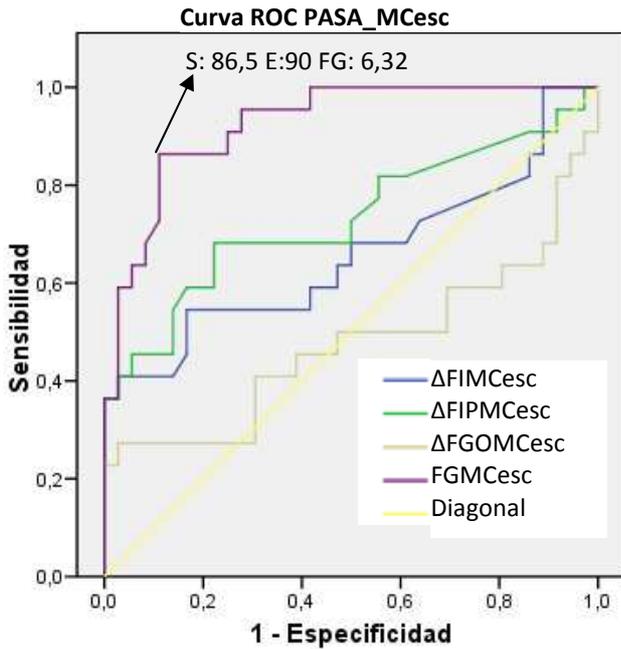
Curva ROC PASA_CMPM



Variables resultado de contraste	AUC MCesc	Variables resultado de contraste	AUC CMPM
ΔFIPMCesc1	,794	ΔFIPCMPM1	,716
ΔFIPMCesc2	,792	ΔFIPCMPM2	,828
ΔFIPMCesc3	,581	ΔFIPCMPM3	,751
ΔFIPMCesc4	,775	ΔFIPCMPM4	,677
ΔFIPMCesc5	,684	ΔFIPCMPM5	,811
ΔFIPMCesc6	,699	ΔFIPCMPM6	,813

Figura 7.18.- Curvas ROC para diferentes selecciones de los pesos de la función incumplimiento ponderado, numerado cada conjunto del 1 al 6.

El objeto es evaluar su poder predictivo, calculando AUC para cada curva y su punto de corte (figura 7.19). Las curvas ROC de la figura 7.19 determinan que la mejor función predictora para todas las variables binarias es la función global de diferencias de dosis (FG), ya que AUC tiene valor mayor que para las otras funciones con respecto a las cuatro variables binarias cuyo resultado se trata de predecir. El valor de corte matemático es parecido en los cuatro casos y está en torno a 6,35, estando expresado como un valor de FG en la figura. La capacidad predictiva de FG es buena y en algunos casos muy buena.



Variables resultado de contraste	AUC PASA_MCesc	AUC PASA_CMPM	AUC PASA_MCescOBJ	AUC PASA_CMPMOBJ
ΔFI	,657	,798	,721	,83
ΔFIP	,732	,802	,825	,841
ΔFGO	,475	,682	,400	,647
FG	,925	,968	,891	,903

Figura 7.19.- Curvas ROC de predicción basadas en FI, FIP, FGO y FG para las variables binarias PASA_MCesc (izq arriba), PASA_CMPM (dch arriba), PASA_MCescOBJ (izquierda abajo) y PASA_CMPMOBJ (derecha abajo). Los resultados indican que el mejor predictor es la función global de diferencias en dosis FG, para un valor de corte (geométrico) muy parecido, en torno a 6,35. Se indica la sensibilidad y la especificidad en el punto de corte. Los resultados de área bajo la curva (AUC) clasifican FG como un buen predictor.

separado para cada conjunto de histogramas, y no se limitan a comparar las diferencias entre referencia y verificación.

En particular, analizando lo que sucede en los diferentes planes de tratamiento seleccionados para este trabajo, se ha observado que la función global de objetivos (FGO) demuestra, incluso, una tendencia inversa a la esperada, es decir, da un valor menor de la diferencia entre referencia y verificación en algunos planes que son inaceptables desde el punto de vista de las métricas de diferencias en dosis. Esto ocurre cuando la curva ROC está por debajo de la diagonal en la figura 7.19. La razón para ello es que FGO contabiliza todas las diferencias con respecto a la restricción sobrepasen o no un umbral determinado.

Con respecto al punto de corte de 6,35 para FG, y tal como se ha mencionado en el apartado 6.2.5, la función de coste está minimizada en este punto si se considera que el peso por aceptar un falso negativo es el mismo que el de rechazar un falso positivo. Evidentemente, si el centro que analiza los datos quiere ser conservador, sería necesario bajar el punto de corte. Por ejemplo, si suponemos una función de coste con la forma

$$\text{Coste} = k \cdot (1 - S) + (1 - E) \quad (\text{ec 7.1.})$$

Estaríamos indicando que si k es distinto de 1 se penaliza diferente el hecho de tratar a un paciente con un plan inaceptable que el de tener que replantear, modificar y/o volver a medir un plan que era aceptable. En los casos estudiados es necesario incrementar k desde 1 hasta 2,5 para cambiar el valor de corte, siendo en ese caso $FG=4.1$, válido para las cuatro variables binarias presentadas. Estos valores se obtienen fácilmente ya que el análisis ROC proporciona las tablas de sensibilidad frente a especificidad en cada punto de la curva, por lo que cada punto tiene una función de coste y buscar el mínimo resulta sencillo.

7.4.2. Análisis de la complejidad de planes

Con respecto a la complejidad de los planes, caracterizada por la variable ALPO, por el número de segmentos NSEG, por las UM totales, por las UM por segmento y por la variable compuesta $UM \cdot SEG / DSES$, podemos utilizar de nuevo la métrica de diferencias en dosis, en este caso con las variables asociadas PASA_CMPM y PASA_MC, según el método de referencia considerado. Usaremos dichas variables binarias para contrastar si tienen poder predictivo. El hecho de utilizar PASA_MC permite contrastar SCC con MC desde el punto de

vista del algoritmo y sin incluir medidas. Es necesario indicar que 11 de 58 planes son positivos para PASA_MC y 26 de 58 lo son para PASA_CMPM, lo que supone una prevalencia muy alta de positivos en ambos casos. Los resultados para cada una de las variables se muestran a través de las tablas de AUC:

Variables resultado de contraste PASA_MC	AUC	Variables resultado de contraste PASA_CMPM	AUC
UM*SEG/DSES	,695	UM*SEG/DSES	,679
ALPOTOTAL	,657	ALPOTOTAL	,522
NSEG	,566	NSEG	,555
UMTOTALES	,773	UMTOTALES	,708
UMSEG	,756	UMSEG	,701

En ellas se observa que ALPO y el número de segmentos son las variables menos relacionadas con los resultados obtenidos, encontrándose una relación moderada entre las otras tres variables y el hecho de que los planes resulten aceptables o no con la métrica de diferencias. El valor de 0,522 para AUC en el caso de ALPOTOTAL en relación a PASA_CMPM indica que es muy mal predictor de los resultados obtenidos a partir de las medidas con Compass.

7.4.3. Análisis de los índices radiobiológicos

Para analizar el valor predictivo de los índices radiobiológicos con objeto de realizar una clasificación adecuada de los planes de IMRT considerados, usaremos una métrica que evalúa las diferencias de TCP y NTCP con sus variables asociadas PASA_TCPCMPMX y PASA_TCPMCescX, descritas en 6.3., donde se acepta el plan si $\Delta TCP > -X\%$ y $\Delta NTCP < X\%$ para todos los volúmenes evaluados en el plan de manera simultánea. El análisis ROC realizado usando este conjunto de variables binarias que se construye variando el umbral X, conduce a curvas de bajo poder predictivo. Los valores que se seleccionaron para el umbral fueron ΔTCP ó $\Delta NTCP = 1,2,3,5,7,9,12$ y 15. Este análisis ROC tomó como sistema clasificador perfecto PASA_MCescOBJ y PASA_CMPMOBJ. En las figuras 7.20 y 7.21 se muestran las curvas obtenidas.

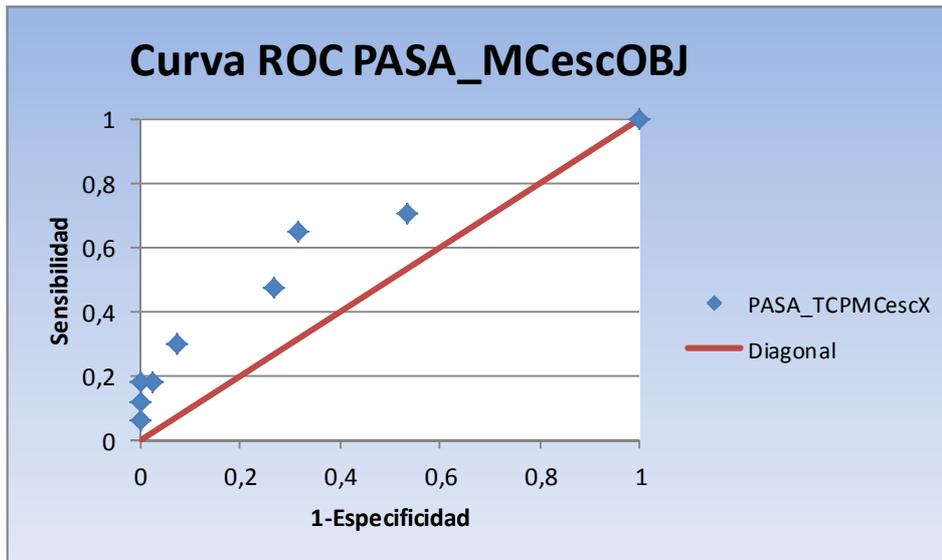


Figura 7.20.- Curvas ROC de predicción para la variable PASA_TCPMCescX. El sistema clasificador perfecto considerado es PASA_MCescOBJ.. Para esta curva AUC=0,66 PC=2

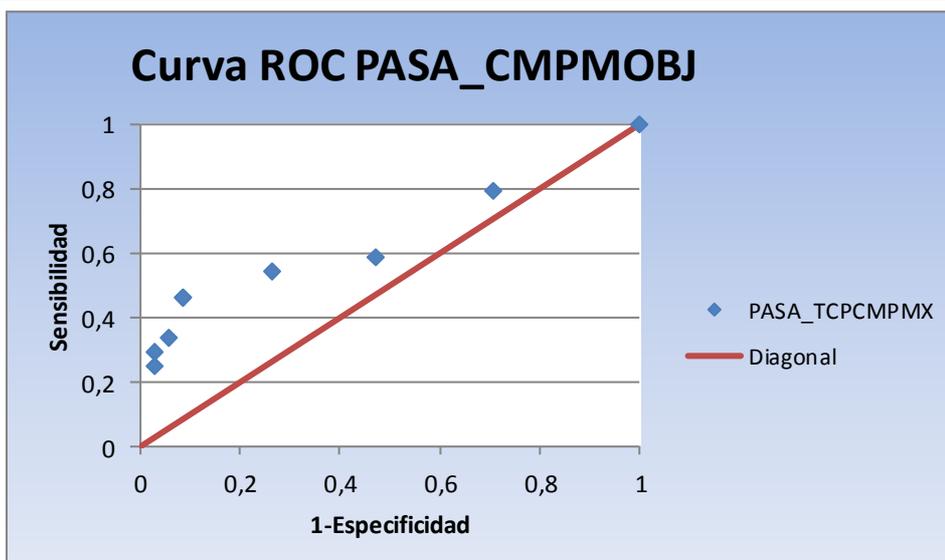


Figura 7.21.- Curvas ROC de predicción para la variable PASA_TPCMPMX. El sistema clasificador perfecto considerado es PASA_CMPMOBJ. Para esta curva AUC=0,65 PC=5

El resultado es pobre e indica que las variaciones de TCP y NTCP tienen poco poder predictivo en cuanto a las métricas definidas. Esto es perfectamente esperable por la definición de TCP y NTCP. Si bien en (19) se demuestra que hay clara correlación entre las variaciones de los histogramas y las variaciones entre los índices radiobiológicos, cuando intentamos utilizar los índices como predictores las cosas no son tan sencillas. La razón está

en que restricciones del plan y puntos críticos de las curvas de TCP y NTCP no coinciden. Esta problemática se ha mencionado en 6.2.3 y se amplía aquí a través de un ejemplo.

Si tomamos uno de los pacientes de cabeza y cuello que tiene un plan rechazado por CPM de debido a dosis alta en médula podemos observar que la curva de NTCP que se va a emplear es la de la figura 7.22.

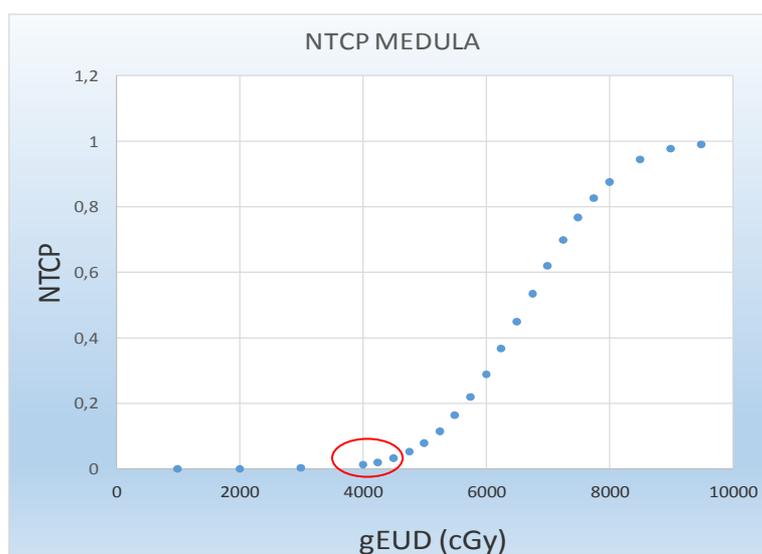


Figura 7.22.- Curva NTCP para la médula en función de la gEUD. En rojo la zona en la que se mueve el ejemplo. Valores de NTCP en tanto por 1.

Esto nos da la siguiente tabla de valores

	SCC	CMPM	SCC	CMPM	SCC	CMPM
	gEUD	gEUD	D1	D1	NTCP	NTCP
Dosis (cGy)	3750	4107	4339	4754	0,006	0,014

Con lo que un incremento de D1 que, claramente, sobrepasa los límites establecidos en la tabla 6.I ($D1 < 44,5$ Gy), no incrementa NTCP más del 1%, quedando el plan aceptado desde el punto de vista radiobiológico, ya que este era el único punto de la comparación de histogramas que ofrecía problemas para este plan. Este hecho puede apreciarse también en la literatura. En concreto en (19) se presentan 5 casos de pacientes de cabeza y cuello y 5 de próstata. Al situar los valores obtenidos para los OARs sobre las curvas dosis-respuesta se aprecia que en el caso de médula y la laringe las NTCP quedan por lo general en una zona de la curva poco sensible a variaciones en la gEUD.

Para aumentar la sensibilidad del test radiobiológico podemos adaptar los valores de la tabla 6.I siguiendo nuestro interés de obtener resultados similares a los de PASA_CMPMOBJ. Para ello, como se ha indicado en el apartado 6.2.3, para los OARs sustituimos los valores de TD_{50} por los valores de las restricciones de la tabla 6.I. Por ejemplo, en el caso de la médula el valor de TD_{50} pasa de 66,5 Gy a 44,5 Gy, que corresponde a identificar $TD_{50}=D1$. En el caso de los PTV tenemos que tener en cuenta que no todos los planes tienen una dosis total que coincida con una TCP alta. Podemos estar hablando de planes que cubren solo una parte del tratamiento o que no tienen intención curativa. Esto quiere decir que puede ser más sensato a la hora de caracterizar la curva de TCP usar un D_{50} que sea un porcentaje de la dosis prescrita para el plan concreto. Se propone el 80%, con objeto de situar la curva dosis-respuesta en la zona de interés. Se busca que los valores de TCP para todos los planes estudiados estén en torno al 80 o 90%. El índice TCP modificado resultante será más sensible a variaciones de la dosis en el PTV, y genera un nuevo conjunto de variables binarias PASA_TCPMODCMPMX, que evalúan las diferencias en TCPMOD y NTCPMOD de acuerdo con el umbral X. En resumen, la modificación de D_{50} y TD_{50} situará los gradientes de las curvas de TCP y NTCP en la zona de interés para el análisis de las restricciones. Si volvemos a calcular todos los índices TCP y NTCP modificados de esta manera artificial, el sistema clasificador mejora su predicción con respecto a la métrica de diferencia de dosis basada en objetivos, tal como se aprecia en la figura 7.23. Se obtiene un $AUC = 0,79$ y un $PC=2$. Esta AUC indica que persiste aún diferente criterio de clasificación entre ambos sistemas. En particular, en el punto de corte ($PC=2$) la sensibilidad es 95,8% pero la especificidad es 58,2% indicando que en ese punto hay muchos planes que no pasan PASA_TCPMODCMPM2 pero están validados según PASA_CMPMOBJ.

En la figura 7.24 se presenta la correlación existente entre los parámetros de cobertura del PTV principal, D_{95} y D_{99} , y los valores de ΔTCP modificados para el algoritmo CPM. Los coeficientes de determinación son 0,88 para D_{95} y 0,67 para D_{99} . Utilizando ΔTCP los coeficientes serían 0,63 y 0,43 respectivamente, lo que evidencia la mejoría que se consigue con la modificación.

Sin embargo, sigue habiendo discrepancias que hacen que ambos sistemas de clasificación no sean iguales. El motivo principal es que las pérdida de cobertura que generan positivos en PASA_CMPMOBJ no siempre lo hacen en PASA_TCPMODCMPMX. Por ejemplo,

una pérdida de cobertura que supera el 3% en D95 o el 5% en el D99 no supone una pérdida de TCPMOD tan grande, situándose en niveles en torno al 1-3%.

Por otro lado, en el punto de corte, que queda situado en el 2%, la especificidad es baja lo que indica que existe un número alto de falsos positivos con respecto a la métrica de diferencias en dosis basada en objetivos. La razón para ello es que ahora en los OARs las diferencias en dosis se ven amplificadas al ser convertidas en Δ TCPMOD.

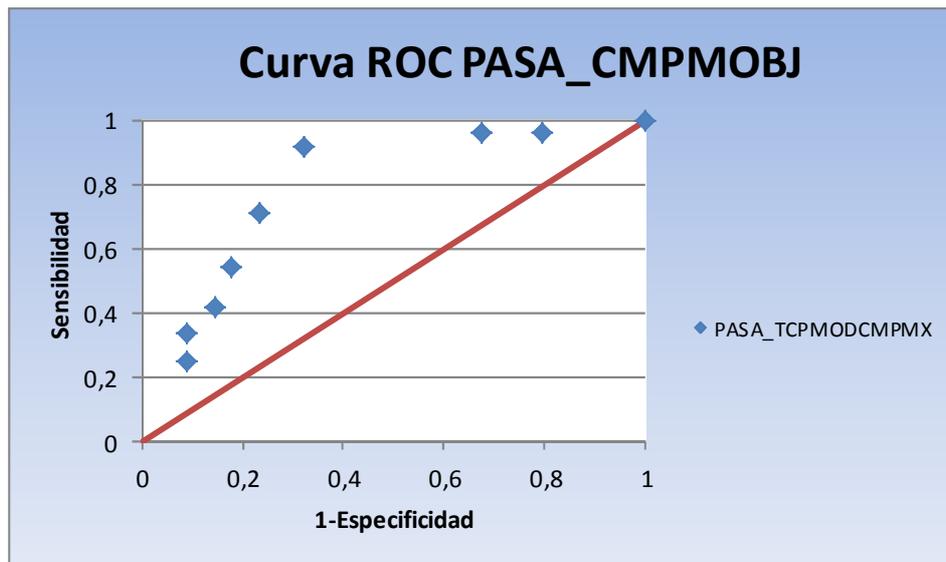


Figura 7.23.- Curva ROC para la variable PASA_TCPMODCMPMX con respecto al clasificador PASA_CMPMOBJ. El D_{50} para el cálculo de TCP se ha adaptado para cada plan. AUC = 0,79. PC=2.

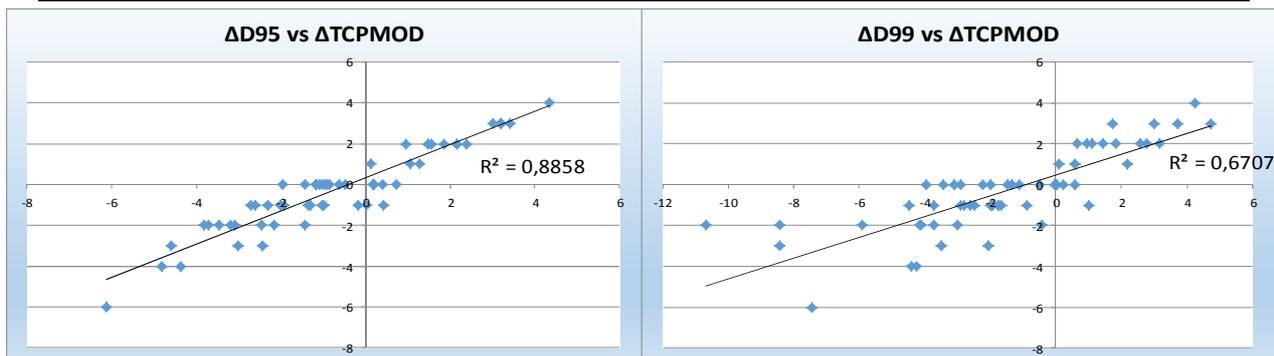


Figura 7.24.- Correlación entre cambios en D95 y D99 del PTV principal con respecto al incremento de TCP modificado. El D_{50} para el cálculo de TCPMOD se ha adaptado para cada plan, situándose en el 80% de la dosis total. Con ello la correlación mejora con respecto al índice TCP habitual.

Al modificar la curva de NTCP, la hemos colocado en un punto en el cual cambios pequeños en la dosis son tenidos en cuenta si se encuentran en las inmediaciones de las dosis de restricción. Podemos verlo a través de un ejemplo: si construimos la variable binaria

PASA_TCPMODCMPM5 (1 si $\Delta TCP > -5\%$ y $\Delta NTCP < 5\%$ y 0 si $\Delta TCP > -5\%$ y $\Delta NTCP < 5\%$) a partir de los valores obtenidos para los incrementos de TCP y NTCP modificados, podemos elaborar su tabla de contingencia (ver 6.2.4) con respecto a PASA_CMPMOBJ (tabla 7.III).

Tabla 7.III Tabla de contingencia PASA_TCPMODCMPM5 * PASA_CMPMOBJ				
		PASA_CMPMOBJ		Total
		0	1	0
PASA_TCPMODCMPM5	0	14	7	21
	1	10	27	37
Total		24	34	58

Este tipo de tablas de contingencia, descritas en 6.2.4, son las que subyacen en todo el análisis presentado, existiendo una por cada punto de las curvas ROC. En este caso vemos que, con respecto a PASA_CMPMOBJ, hay 7 falsos positivos y 10 falsos negativos. Los 7 falsos positivos corresponden a planes poco robustos en los que NTCP aumenta por encima del 5% pese a que los criterios de que CMPMOBJ no se incumplen. De los siete, 4 corresponden a cristalino y 3 corresponden a parótidas. En cuanto a los 10 falsos negativos, 7 son por pérdidas de cobertura fuera de tolerancia en PASA_CMPMOBJ pero que no suponen gran pérdida de TCP. Hay 2 que corresponden a que D2 del PTV sobrepasa el incremento del 3%. Que aumente D2 lleva consigo que la dosis máxima en el PTV crezca y, en consecuencia, aumente el TCP. Esto favorece que sea aceptado el plan según PASA_TCPMODCMPM5 mientras que, según PASA_CMPMOBJ, debe ser rechazado. Por último, hay un plan donde el cristalino ya estaba muy por encima de su dosis de tolerancia y así se aceptó, pero que, además, discrepa entre verificación y referencia más del 3%. Obviamente, no tiene consecuencias clínicas pero CMPMOBJ no lo da por válido. A la luz de este ejemplo podemos comprobar que, aunque coinciden muchas veces, en algunas ocasiones las métricas basadas en índices radiobiológicos reflejan aspectos que no contemplan las basadas en diferencia de dosis y viceversa.

Los positivos de la métrica PASA_TCPMODCMPM5 del ejemplo anterior pueden ser relacionados con el concepto de robustez del plan, introducido en este trabajo (ver apartado 6.2.3). Un análisis basado en IR revela que muchas de las causas que conducen a planes inaceptables según PASA_CMPMOBJ no tienen el mismo efecto en la robustez de los planes.

Un índice de robustez pequeño o cercano a cero para un plan calculado con SCC no tiene por qué conducir a un plan no aceptado según PASA_CMPM o PASA_CMPMOBJ, pero sí ha de tener mayor probabilidad de conducir a ello, ya que se encuentra más cerca de las restricciones establecidas. Además, IR no será sensible a la pérdida de cobertura o al aumento del valor máximo de dosis dentro del PTV, de manera análoga a PASA_TCPMODX. Se profundizará más en este asunto en el apartado 7.5.4.

7.4.4. Análisis del índice gamma 3D

En lo referente a la métrica gamma, aunque el criterio para aceptar o rechazar un plan de tratamiento se ha fijado en 3%/3mm para el contorno externo, PTVs y OARs, se puede variar el % de puntos que se permite que incumpla el criterio y construir con ello también una curva ROC tomando como clasificador perfecto PASA_CMPM y PASA_CMPMOBJ. La gamma 3D solo ha sido evaluada para el algoritmo CMPM en todos los volúmenes de interés. En la figura 7.25 se representa la curva ROC elaborada usando diferentes umbrales de aceptación para el criterio gamma: 1,5,7.5,15,25 y 40 %. Le corresponde un AUC=0,88, y un punto de corte óptimo que fija un umbral en el porcentaje de puntos que se permite que tengan gamma >1 en el 10%, coincidente además para ambas variables, siendo la sensibilidad del 77% y la especificidad del 90% para PASA_CMPM y 79% y 84% para PASA_CMPMOBJ. Este resultado es muy bueno para la métrica gamma global por órganos y está por encima de lo esperado, ya que dicha métrica no distingue entre diferencias en dosis “a peor”, es decir, menos dosis en PTV y más dosis en el OAR, lo que debería estar proporcionando una cantidad apreciable de falsos positivos con respecto a PASA_CMPM. Lo que ocurre en realidad es que cuando $\Delta D95$ y ΔDM son mayores del 3%, hecho que suele conducir a un positivo según PASA_GAMMA10, pero que es negativo según PASA_CMPM en un principio, algún OAR tiene un ΔDX superior al 3% y, por tanto, el resultado de PASA_CMPM también resulta positivo.

A la vista de este resultado, incluir en el informe de verificación de IMRT los resultados de gamma global en cada uno de los órganos parece adecuado, ya que, además, la función gamma suele estar incluida en los sistemas de verificación. Tal como se describe en 6.2, todos los volúmenes son analizados simultáneamente para ver si tienen un porcentaje de puntos fuera de tolerancia por encima del umbral. De los 20 planes positivos según

PASA_GAMMA10, sólo 1 tiene un porcentaje de puntos fuera de tolerancia por encima del 10% en el contorno externo, por lo que la matriz total no es tan sensible con respecto a PASA_CMPM como el análisis volumen a volumen.

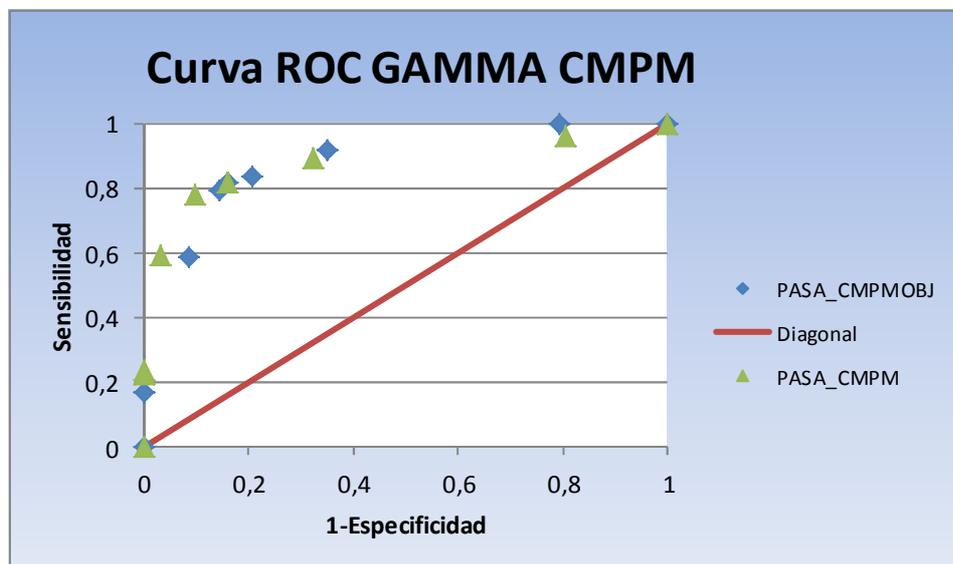


Figura 7.25.- Curva ROC para la gamma global en órganos, determinada a partir de CMPM, tanto con respecto a PASA_CMPMOBJ como PASA_CMPM. Se utilizan umbrales del 1,5,7.5,15,25 y 40%. AUC = 0,88 y PC=10% en ambos casos.

7.4.5. Capacidad predictiva de las medidas con cámara

En la figura 7.26 se representa la curva ROC para la diferencia en valor absoluto de dosis medida con la cámara pin-point en un punto significativo del maniquí de verificación, tanto con respecto a PASA_CMPMOBJ como PASA_CMPM. Se utilizan umbrales del 0.5,1,1.5,2,3,4,5,6 y 7%, construyéndose el conjunto de variables binarias PASA_CAMARAX. Su uso para predecir la métrica de diferencias sea o no basada en objetivos no difiere del de un clasificador aleatorio, lo que en un principio es un resultado sorprendente si tenemos en cuenta que la medida con cámara en un punto es el uno de los elementos principales de la verificación convencional de planes. Las razones para ello son variadas:

- Diferencias en el estado del colimador entre el día de la medida con cámara y el día de la medida con otros dispositivos. Se ha visto que la incertidumbre asociada al estado del colimador para la dosis en un punto dentro del PTV principal está en torno al $\pm 2\%$. No se conoce el valor del parámetro PPBC, los días en los que se verificó el tratamiento midiendo en un punto de un maniquí con cámara, para un

70% de los planes estudiados en el trabajo, por lo que no es posible corregir dichas medidas por PPBC.

- Diferencias en el maniquí. Las medidas con cámara se realizan en el interior de un maniquí (cúbico o con hemcilindros acoplados para simular pelvis, ver figura 6.1), mientras que Compass calcula en una geometría diferente (TC del paciente).
- Diferencias locales de dosis en un punto compatibles con un histograma con similares valores de dosis media y cobertura. La posición particular de una lámina o un segmento sobre el punto de medida puede dotar de inestabilidad a la medida de dosis en ese punto, pese a ser escogido siempre en una zona de bajo gradiente de la dosis global.

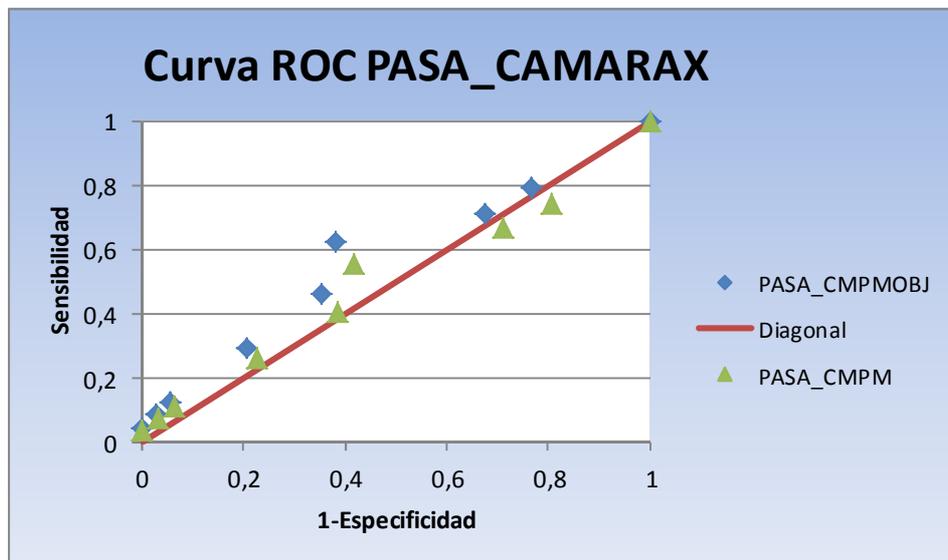


Figura 7.26.- Curva ROC para la diferencia en valor absoluto de dosis medida con la cámara pin-point tanto con respecto a PASA_CMPMOBJ como PASA_CMPM. Se utilizan umbrales del 0.5,1,1.5,2,3,4,5,6 y 7%. AUC < 0.58 y PC=2%.

Para tratar de aislar los efectos mencionados se ha realizado un estudio controlando el valor de PPBC y de las variaciones de dosis/UM del acelerador y calculando con COMPASS en el mismo maniquí en el que, simultáneamente, se está midiendo con cámara. Los resultados se presentan en 7.5.3.

7.5. Correlación entre métricas de aceptación

En este apartado se expresan las correlaciones encontradas entre los diferentes sistemas de clasificación y se discuten las diferencias encontradas.

En la tabla 7.IV podemos observar el resultado de todas las posibles pruebas de Chi-cuadrado derivadas de las diversas tablas de contingencia que se pueden construir entre variables con resultado POSITIVO (0) o NEGATIVO (1) en la verificación. Las variables consideradas en la tabla son aquellas que se consideran más relevantes tras el análisis ROC realizado en el apartado anterior. En la tabla de doble entrada se consignan las significaciones exactas bilaterales de las correspondientes pruebas así como el valor de Chi-cuadrado obtenido entre paréntesis. Se han marcado en negro las variables que se han encontrado dependientes, con significaciones por debajo de 0,05 y en rojo aquellas que resultan independientes con significación por encima de 0,05. En esta tabla se condensa mucha información acerca de las distribuciones de positivos y negativos obtenidos. Todas las variables que provienen de las medidas realizadas con el sistema COMPASS conducen a distribuciones dependientes de positivos y negativos, ya que el valor de la significación está por debajo de 0,05. PASA_CMPM OBJ, PASA_GAMMA10, PASA_TCPMODCMPM2, PASA_TCP5CMPM y PASA_FG6,35CMPM están fuertemente relacionadas, y además, se encuentra que las distribuciones de PASA_MCescOBJ y PASA_FG6,35MCesc también lo están. Si analizamos la regresión entre las variables continuas FGCMPM y FGMCesc obtenemos la figura 7.27, donde se muestran los planes que discrepan en la tabla de contingencia para el valor de corte de 6,35 (ver apartado 7.4.1).

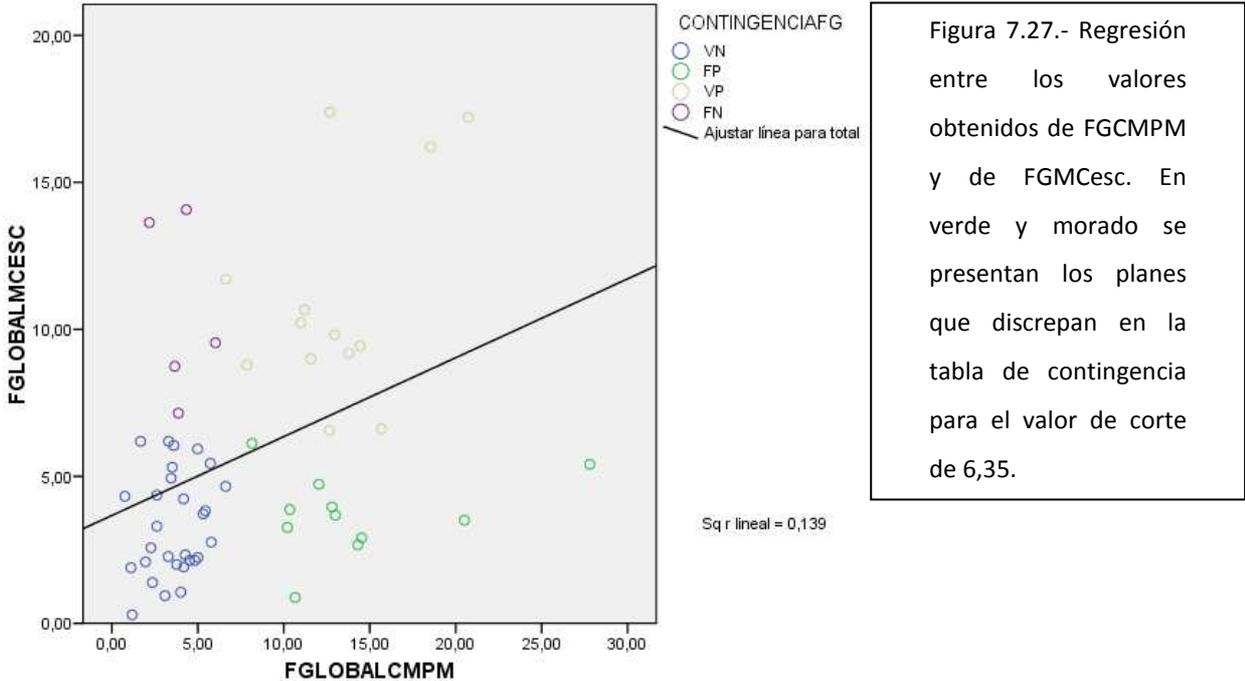


Figura 7.27.- Regresión entre los valores obtenidos de FGCMPM y de FGMCesc. En verde y morado se presentan los planes que discrepan en la tabla de contingencia para el valor de corte de 6,35.

Tabla 7.IV Tabla cruzada con la significación del test Chi-cuadrado y valor de Chi-cuadrado entre paréntesis. Se destacan en rojo las significaciones por encima de 0,05 que emparejan variables que se consideran independientes, y por tanto reflejan que la clasificación de planes que realizan no está relacionada.

	PASA_CMPMOBJ	PASA_MCescOBJ	PASA_GAMMA10%	PASA_DELTA	PASA_DELTAOBJ	PASA_CONVENCIONAL	PASA_TPCMPMOD2	PASA_TCP5 CPM	PASA_FG6,35MCesc	PASA_FG6,35CPM
PASA_CMPMOBJ	0	0,02 (5,395)	0,0 (24,1)	0,77(0,089)	0,018 (5,573)	0,014 (6,086)	0,0 (18,196)	0,001 (10,524)	0,009 (6,880)	0,0(32,905)
PASA_MCescOBJ		0	0,082 (3,017)	0,035(4,43)	0,001 (11,062)	0,346 (0,888)	0,196 (1,673)	0,051 (3,813)	0 (23,196)	0,032 (4,576)
PASA_GAMMA10%			0	0,88 (0,023)	0,396 (0,719)	0,157 (2,002)	0,0 (13,771)	0 (14,954)	0,009 (6,880)	0,0 (32,905)
PASA_DELTA				0	0,0 (14,4)	0,686 (0,164)	0,009 (6,806)	0,039 (4,251)	0,097(2,756)	0,296 (1,092)
PASA_DELTAOBJ					0	0,201 (1,636)	0,134 (2,25)	0,257 (1,286)	0,134 (2,25)	0,098(2,731)
PASA_CONVENCIONAL						0	0,118 (2,438)	0,210(1,569)	0,049(3,88)	0,182 (1,782)
PASA_TPCMPMOD2							0	0,001 (10,474)	0,008 (7,117)	0,0 (24,939)
PASA_TCP5 CPM								0	0,002(9,533)	0,0 (18,628)
PASA_FG6,35MCesc									0	0,003 (9,024)
PASA_FG6,35CPM										0

En ella se aprecia que la regresión entre ambas variables no da un resultado satisfactorio, si bien existe una relación evidente entre ambas variables coincidiendo en un amplio número de casos. Utilizando el valor de corte de 6,35 y la variable FGMCesc como referencia encontramos 11 casos que serían falsos positivos y 4 que serían falsos negativos, estando los 43 casos restantes igualmente clasificados por ambas variables. Por tanto aunque la significación de la prueba entre las variables binarias derivadas PASA_FG6,35MCesc y PASA_FG6,35CMPM del punto de corte es 0,003, ambas discrepan en la clasificación del 25% de casos. Esto indica que la tabla 7.IV sobre todo sirve para resaltar aquellas variables binarias que se pueden considerar totalmente independientes, es decir, métricas de las que se están obteniendo resultados contradictorios.

7.5.1. Correlaciones para Delta4

Si nos centramos en las variables binarias asociadas a Delta4, PASA_DELTA y PASA_DELTAOBJ, con respecto a CMPM vemos que la independencia es descartable con respecto a PASA_MCescOBJ y PASA_CMPMOBJ y en el caso de PASA_DELTA también con respecto a PASA_TCPMOD2 y PASA_TCP5CMPM. El resultado deja tablas de contingencia que suponen que los mismos pacientes evaluados con la misma métrica con dos sistemas de medida independientes no siempre conducen a los mismos resultados. La tabla de contingencia para Delta4 y CMPM con la métrica de diferencias basada en objetivos se presenta como la 7.V.

Tabla 7.V Tabla de contingencia PASA_CMPMOBJ * PASA_DELTAOBJ				
		PASA_DELTAOBJ		Total
		0	1	0
PASA_CMPMOBJ	0	9	8	17
	1	3	16	19
Total		10	26	36

En la tabla 7.V se aprecia que en el 30% de los casos Delta4 y Compass no clasifican igual, si bien las variables, al menos, no son independientes dado que la significación de la Chi-cuadrado es 0,018. Cabe señalar que Delta4 y Compass son muy diferentes en la reconstrucción de dosis 3D, siendo en la TC de planificación para Compass y sobre un maniquí cilíndrico en el caso de Delta 4. También es diferente la generación de HDV a partir de los volúmenes importados y las dosis reconstruidas, usándose un modelo más simple en

el caso de Delta4. Como se ha dicho anteriormente, Compass tiene una puesta en marcha parecida a un SPT mientras que Delta4, tal como se empleó, sólo precisó de una calibración en dosis del detector.

Además, la medida del plan con cada uno de los sistemas se realizó en momentos distintos, lo que supone un estado diferente del MLC y de la dosis por UM del acelerador. No obstante, se ha comprobado que, corrigiendo los histogramas de CMPM por el PPBC obtenido el día que se irradió cada plan, la tabla de contingencia no cambia, y el test de Chi-cuadrado sigue dando los mismos valores (0,018 la significación con Chi-cuadrado=5,573). Este análisis se ha hecho a través del escalado proporcional, tomando como desviación de dosis media en el PTV la que se obtiene en promedio usando la recta de la figura 7.14, escalando los HDV de CMPM como si se hubieran medido en un estado del MLC caracterizado por un parámetro PPBC igual que cuando se hicieron las medidas de Delta4. Dentro de la variable PASA_CMPMOBJ solo hay 4 planes que cambian de inaceptable a aceptable, mientras que hay 2 planes que pasan de aceptable a inaceptable, siempre dentro de las limitaciones aproximación del escalado proporcional (ver apartado 7.3.3). De esos 6 planes ninguno pertenece al conjunto de 36 planes comunes entre Delta4 y CMPM. Por tanto, las razones de las diferencias entre Delta4 y Compass no se encuentran en el estado del MLC. Sí que existen diferencias debidas al algoritmo de cálculo. Delta4, en el modo en que se usa, toma la matriz 3D planificada con SCC y la perturba por la medida, mientras que CMPM lleva implícito el algoritmo de cálculo CMPC, existiendo diferencias de partida entre CMPC y SCC sin introducir ningún tipo de medida en primera instancia. Este hecho será analizado más adelante.

Entre PASA_DELTAOBJ y PASA_MCescOBJ las cosas funcionan un poco mejor. Ambas variables clasifican distinto solo un 19% de los planes coincidiendo en una amplia mayoría. La tabla de contingencia es la 7.VI.

		PASA_MCescOBJ		Total
		0	1	0
PASA_DELTAOBJ	0	8	3	10
	1	4	21	26
Total		11	25	36

La razón para esta mejoría podría estar en que MC y SCC tienen un mejor acuerdo que CMPC y SCC, y, como hemos mencionado, el punto de partida de Delta4 es la matriz de SCC.

7.5.2. Correlación entre MC escalado y CMPM

En cuanto a MCesc y CMPM, el análisis es importante, ya que ambos clasificadores se pueden tomar como referencia y ambos usan las medidas hechas con Compass para generar las métricas. Es importante que se descarte su independencia a la hora de clasificar planes, tal como confirma la tabla 7.IV, con una significación de 0,02 para la Chi-cuadrado entre PASA_MCescOBJ y PASA_CMPMObj. En el fondo esa comparación de PASA_MCescOBJ y PASA_CMPMObj es una comparación entre el algoritmo MC y el algoritmo CMPC, ya que MCesc está escalado precisamente por la diferencia entre CMPM y CMPC. La tabla 7.VII muestra que ambos clasificadores discrepan en la clasificación del 32% de los planes.

		PASA_MCescOBJ		Total
		0	1	0
PASA_CMPMObj	0	11	13	25
	1	6	28	34
Total		17	41	58

En la tabla 7.VIII podemos observar resumida la correlación encontrada entre MC y CMPC referente a los parámetros de la tabla 6.I extraídos para PTVs y OARs. La correlación para tronco, quiasma, vías ópticas, recto y vejiga es mala, mientras que para PTVs, médula, parótidas y cristalinos se muestra significativa. En la figura 7.28 se presentan las correlaciones extraídas para la DM y la D95 del PTV principal. Lo que más llama la atención de las gráficas presentadas es el desacuerdo entre los resultados obtenidos con MC y CMPC a la hora de calcular las diferencias en dosis media y cobertura del PTV con respecto a SCC, donde la correlación entre incrementos encontrada es pobre y hay muchos casos, especialmente en D95, que se sitúan en cuadrantes con signo opuesto. Si a esto añadimos el mal comportamiento de las correlaciones en determinados OARs, es necesario preguntarse directamente por la tabla de contingencia entre MC y CMPC para la función de diferencias basada en objetivos (tabla 7.IX). Se construyen las variables binarias PASA_MCOBJ y PASA_CMPCOBJ, que aunque no están descritas es 6.3, es fácil ver que son análogas a PASA_MCescOBJ y PASA_CMPMObj pero sin usar medidas.

Tabla 7.VIII. Coeficientes de regresión y su significación entre paréntesis para la comparación de diferencias encontradas con SCC para los algoritmos indicados. PTVs, médula y parótida derecha y cristalinos mantienen una correlación significativa entre CMPC y MC mientras que entre MCesc y CMPM ocurre para PTVs, médula, parótidas, tronco y cristalinos.

	Δ DM PTVp	Δ D95 PTVp	Δ DM PTVs	Δ D95 PTVs	Δ D1 MED	Δ DM PAR_I	Δ DM PAR_D	Δ D1 TRC	Δ D1 QUI	Δ D1 VOP_I	Δ D1 VOP_D	Δ D1 CRI_I	Δ D1 CRI_D	Δ D20R	Δ D30V
CMPC- MC	0,550 (0,000)	0,426 (0,001)	0,709 (0,000)	0,495 (0,001)	0,579 (0,001)	0,264 (0,159)	0,472 (0,01)	0,215 (0,441)	0,32 (0,226)	0,08 (0,769)	0,311 (0,259)	0,526 (0,01)	0,491 (0,017)	0,38 (0,401)	0,145 (0,816)
MCesc- CMPM	0,756 (0,000)	0,626 (0,000)	0,747 (0,000)	0,662 (0,000)	0,522 (0,004)	0,516 (0,005)	0,555 (0,003)	0,731 (0,005)	0,344 (0,192)	0,182 (0,5)	0,378 (0,164)	0,503 (0,015)	0,525 (0,023)	0,267 (0,562)	0,156 (0,802)

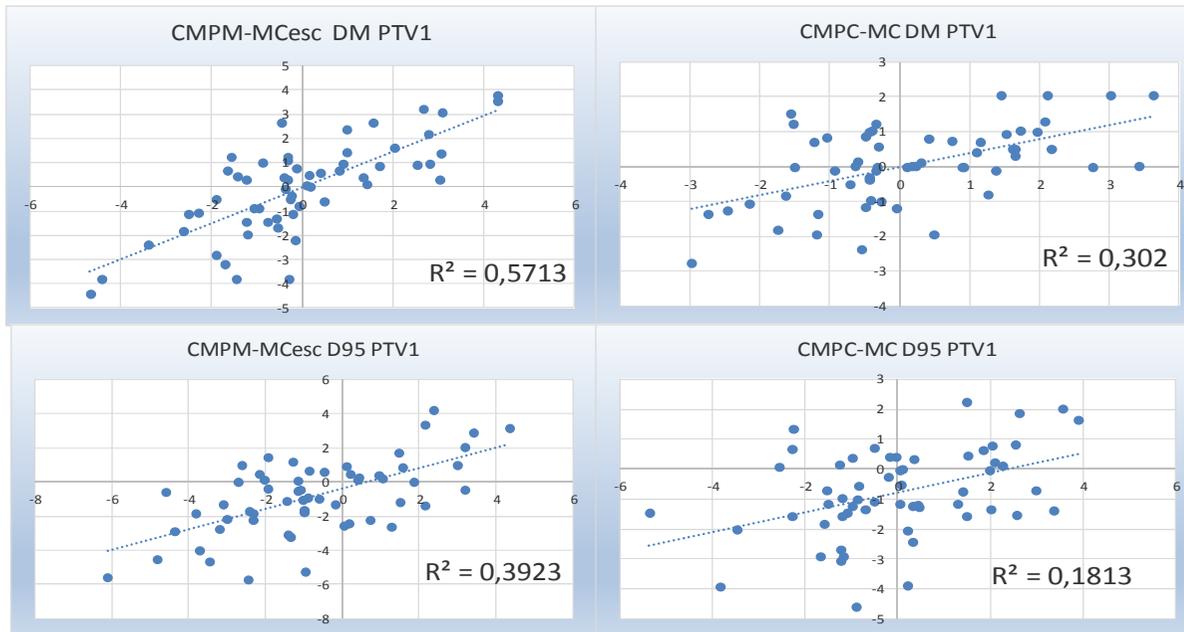


Figura 7.28.- Gráficas de correlación para Δ DM y Δ D95 en el PTV principal para los algoritmos CMPM y MCesc (izquierda), así como CMPC y MC (derecha). Cada algoritmo se compara con SCC y los resultados se correlacionan. Las coeficientes de determinación entre CMPC y MC reflejan un desacuerdo importante entre los dos algoritmos de cálculo redundante.

Tabla 7.IX. Tabla de contingencia PASA_CMPCOBJ * PASA_MCOBJ				
		PASA_MCOBJ		Total
		0	1	
PASA_CMPCOBJ	0	2	12	14
	1	3	41	44
Total		5	53	58

En la tabla 7.IX se pone de manifiesto el principal problema de la comparación entre los resultados calculados por MC y CMPC: los de CMPC con la métrica de diferencias de dosis basada en objetivos muestran muchos planes que no cumplen criterios con respecto al cálculo de SCC, sólo como cálculo redundante, un 24%, mientras que MC no acepta el 8% de los planes. Esta diferencia es la base principal de las que luego se encuentran entre MCesc y CPM, dado que en las clasificaciones realizadas con ambos, influye la diferencia que existe entre CPM y CMPC. Además, los planes que se rechazan con CMPC y MC simultáneamente son tan solo dos, existiendo discrepancia de criterio en 15 (25%).

Este resultado coincide con el de algunos autores (49), que indican que el utilizar un algoritmo independiente de cálculo para la reconstrucción de la dosis que se utiliza en la verificación puede conducir a discrepancias adicionales en los resultados de la misma, debidas a las incertidumbres asociadas a la puesta en marcha y cálculo de dosis del sistema redundante.

7.5.3. Correlaciones para las medidas convencionales

Todos los planes investigados han sido medidos con cámara de ionización en un punto junto con la comprobación de la dosis relativa en determinados planos, bien con EPID, placas radiocrómicas o MatriXX en el maniquí Multicube. De los 58 planes solo 4 fueron inaceptables para tratamiento según la métrica convencional, y tuvieron que ser modificados variando el número de unidades de monitor de los campos para compensar la diferencia de dosis medida con cámara. 11 planes más resultaron dudosos a la vista de los resultados, pero se aceptaron por estar justo en el límite de tolerancia de las medidas con cámara (8 de ellos) y cumplir tolerancias los resultados de la comparación de distribuciones de dosis en planos, o bien porque, tras dar positivo en la verificación con EPID, posteriormente la verificación con Multicube fue aceptable. Precisamente, en la verificación

convencional, una de las razones para hacer verificación con Multicube (6 planes) fue que el análisis previo no fuera satisfactorio.

Tal como se refleja en la tabla 7.IV las únicas variables que no se muestran independientes de PASA_CONVENCIONAL son PASA_CMPMOBJ y PASA_FG6,35MCesc, ya que presentan valores de Chi-cuadrado con una significación por debajo de 0,05. La tabla de contingencia entre la verificación convencional y la de CMPM en sus dos modalidades (diferencias en dosis y basada en objetivos) se muestra en la 7.X.

A la vista de la tabla se observa que los 4 planes de tratamiento que no fueron aceptados por la verificación convencional y que supusieron una acción correctora en su día, tampoco son aceptados por CMPM y en ello radica la significación encontrada en la tabla de contingencia.

Tabla 7.X. Tabla de contingencia PASA_CMPM * PASA_CONVENCIONAL				
		PASA_CONVENCIONAL		Total
		0	1	
PASA_CMPM	0	4	23	27
	1	0	31	31
Total		4	54	58
Tabla de contingencia PASA_CMPMOBJ * PASA_CONVENCIONAL				
		PASA_CONVENCIONAL		Total
		0	1	
PASA_CMPCOBJ	0	4	20	24
	1	0	34	34
Total		4	54	58

Sin embargo, como ya se había anticipado, se observa que hay al menos 20 planes que no pasan las métricas relacionadas con CMPM. La prevalencia de los planes inaceptables es muy alta tanto para la métrica de diferencias como para la métrica basada en objetivos. Si pormenorizamos en la clasificación que realiza PASA_CMPM, de los 27 planes que no se aceptan, 7 de ellos lo hacen por el estado del colimador, estando el parámetro PPBC fuera de tolerancia el día de la verificación (6 por exceso y uno por defecto), 10 son por pérdida de cobertura en los PTV principales, 6 sobrepasan la diferencia permitida para la dosis máxima al PTV, 3 no cumplen restricciones en los cristalinos y 1 en tronco y médula. En lo referente a

problemas con PPBC, cabe señalar que las medidas convencionales y las realizadas con Compass se han hecho en fechas diferentes. Las de Compass se hicieron en fechas posteriores al tratamiento siendo las medidas convencionales las que se hicieron de forma previa al tratamiento. Estas últimas se realizaron en su día con un estado de colimador diferente, por lo general con PPBC inferior, por lo que es fácil entender que estos planes en su día pasaran la verificación convencional. En cuanto a los casos rechazados por pérdidas de cobertura, de los 10 casos en tan sólo 1 de ellos también está fuera de tolerancia (según PASA_CMPM) la dosis media en el PTV, estando en los demás casos por encima del -3%. En general las diferencias negativas en dosis media se amplifican en términos de cobertura (D95). Las medidas con cámara de ionización suelen realizarse en el centro del PTV (o, si hablamos de un maniquí, en el centro de gravedad de la isodosis de prescripción), por lo que probablemente han de estar correlacionadas más con la dosis media al PTV que con su dosis de cobertura.

Para poder esclarecer más el asunto se han realizado medidas simultáneas con una cámara pin-point y con el sistema Compass en 18 de los 58 planes estudiados. En la figura 7.29 se representan cuatro correlaciones por medio de la recta de regresión entre las diferencias de dosis obtenidas con cámara pin-point y la DM en el PTV principal, la D95 y la dosis calculada con CMPM sobre el maniquí IMRT Phantom, en el mismo punto en que se puso la cámara pin-point para medir y también en la TC del paciente. Los coeficientes de determinación se pueden observar también en la figura y están lejos de 1 incluso para la correlación de la medida con cámara frente a la dosis calculada con CMPM en el mismo punto en el mismo maniquí. La significación (bilateral) de la correlación es menor de 0,05 para el cálculo de la dosis CMPM en el mismo punto donde se mide con cámara tanto en el maniquí (cubo o pelvis IMRT phantom) como en el paciente, siendo esta última correlación más débil, indicando que las heterogeneidades y forma del paciente presentes en el cálculo CMPM introducen perturbaciones en la correlación entre ambas diferencias. La pendiente de las rectas es también inferior a 1, que sería la deseable en un modelo ideal donde las medidas de Compass y las medidas hechas con cámara coincidieran plenamente. Las diferencias encontradas con CMPM parecen amplificadas con respecto a las que se encuentran con cámara en el mismo punto. Estos resultados no son diferentes a otros de la literatura; en concreto en (76), con objeto de validar Compass para su uso en la verificación

de planes de arcoterapia volumétrica modulada (VMAT), se miden 12 planes creados en un maniquí con una cámara de ionización y se comparan los valores obtenidos con CMPC y CPM. Los autores no presentan la correlación entre las diferencias encontradas con el SPT, pero sí los datos obtenidos en dosis absorbida. Es posible realizar por tanto la correlación, y al hacerla comprobamos que el coeficiente de determinación es 0,47, inferior incluso al valor de 0,591 obtenido en este trabajo (figura 7.29). Sin embargo, la conclusión de los autores es que existe un buen acuerdo entre Compass y cámara dado que las diferencias no superan el 3% de la dosis prevista por el SPT, lo que indica un nivel de exigencia bajo en cuanto a la correlación que debería existir entre los dos sistemas de medida. También en (78) se presentan resultados que comparan CPM y medidas con cámara en un maniquí cúbico para 12 pacientes de VMAT, y analizando los datos que presentan se obtiene que la correlación entre diferencias cámara-SPT y diferencias CPM-SPT es de 0,48, con un comportamiento muy parecido al caso anterior. Al igual que en nuestro caso, en los dos extraídos de la literatura se observa que la dosis obtenida con CPM está por encima de la medida con cámara en el mismo punto.

En la figura 7.29 también se comprueba que en lo referente a las diferencias en dosis media y en D95 en el PTV principal, la correlación no ha resultado significativa al comparar las diferencias obtenidas con cámara y las obtenidas con CPM. Por tanto, y volviendo por tanto a lo expuesto más arriba, parece que la medida de dosis en un solo punto con cámara de ionización no parece tener ningún poder predictor sobre las diferencias que luego se encontraría en dosis media y cobertura del PTV principal usando los modelos de la figura 7.29, o, en otras palabras, los resultados de las medidas de cámara apenas son explicados por las diferencias en dosis media y cobertura obtenidos con Compass y viceversa. Casi todas las razones que conducen al rechazo de planes en la medida con CPM no son detectables por la medida de dosis en un punto.

Si centramos la atención de nuevo en la verificación convencional, esto es, la combinación de medidas con cámara con las medidas en un plano, bien sea haz a haz o bien sea la suma de todos los haces en el interior de un maniquí, tiene difícil correlación con las métricas obtenidas a partir de los HDV, como ya ha sido documentado previamente (14,22). Por tanto, tal y como se comprueba, no es de esperar que del uso de la métrica gamma en la comparación de planos aumente la correlación entre la verificación convencional y

PASA_CMPM. En este trabajo y para el MLC estudiado, se ha comprobado que el uso de métricas basadas en HDV aumenta sensiblemente el número de positivos en la verificación. Algunos autores han encontrado resultados similares; en concreto en (49) se analizan 20 planes de tratamiento en nasofaringe y se comparan métricas gamma con métricas basadas en HDV, obteniéndose discrepancias importantes en éstas últimas que habían pasado desapercibidas en las primeras. Este tipo de correlaciones débiles entre ambos tipos de métricas también había sido ya anticipado por Stasi (24) y Carrasco (9).

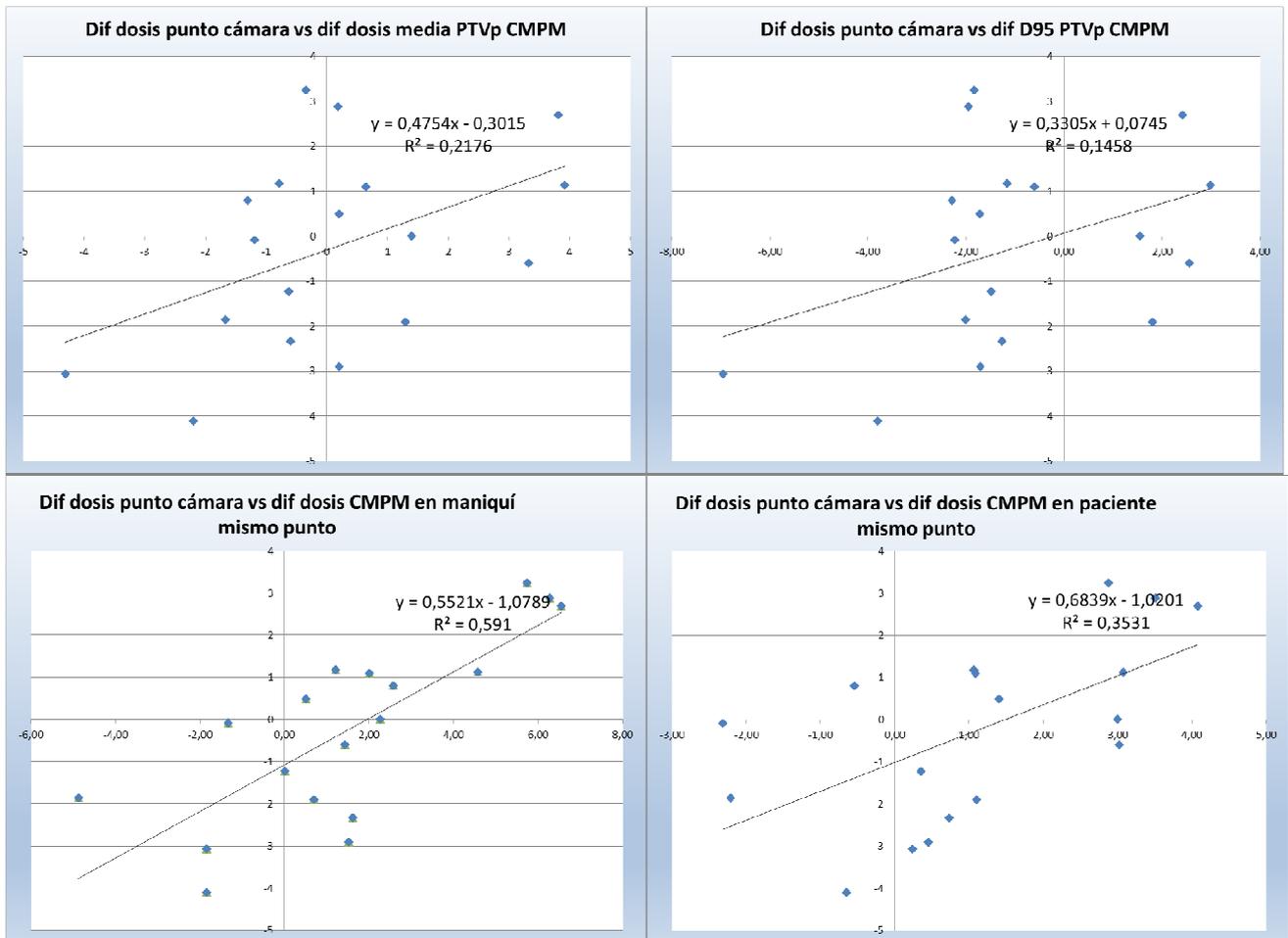


Figura 7.29.- Regresión obtenida a partir 18 planes medidos con cámara y Compass el mismo día, para las diferencias entre la medida con cámara y SCC frente a (de izq a dch y de arriba abajo): la diferencia encontrada CMPM-SCC en la dosis media en el PTV principal, la diferencia encontrada CMPM-SCC en la D95 en el PTV principal, la diferencia CMPM-SCC calculando con CMPM en el mismo punto y en el mismo maniquí en que se mide con cámara, la diferencia CMPM-SCC calculando con CMPM en el mismo punto pero en el paciente (imágenes TC).

7.5.4. Índice de robustez

Se ha construido una variable PASA_IR cuyo valor se ha determinado para cada plan usando los algoritmos SCC y CPM. Para ello se calculan los índices de robustez IRSCC e IRCMPM de acuerdo con la ecuación 6.12. También se calcula la diferencia entre ellos para un mismo plan, ΔIR . PASA_IR es un positivo y el plan es rechazado cuando el valor de IR es menor de 0,05 y $\Delta IR < 0$, o bien, cuando la pérdida ΔIR es superior a 0,05. Se considera además que un plan tiene poca robustez cuando IR es inferior a 0,05. La elección de este último valor se debe al análisis de los índices de los planes estudiados. IR disminuye rápidamente conforme alguno de los valores de restricción está cerca, y es fácil que los valores de IR sean pequeños. La distribución de la robustez de planes es la que se aprecia en la tabla 7.X. En ella se distribuyen los planes robustos y poco robustos con respecto a las variables PASA_TCPMOD2 y PASA_TCPMOD5 provenientes del análisis radiobiológico de los resultados de CPM.

Tabla 7.XI. Distribución de la robustez de planes con respecto a las variables radiobiológicas modificadas, junto con la probabilidad de que el plan resulte positivo según PASA_TCPMOD2 y PASA_TCPMOD5				
	PASA_TCPMOD2		PASA_TCPMOD5	
	NOPASA_TCPMOD2	PASA_TCPMOD2	NOPASA_TCPMOD5	PASA_TCPMOD5
POCOROBUSTO	21	18	3	85,7
ROBUSTO	37	19	18	51
	PASA_TCPMOD2		PASA_TCPMOD5	
	PROBABILIDAD DE NO PASAR %		PROBABILIDAD DE NO PASAR %	
POCOROBUSTO	21	13	8	61,9
ROBUSTO	37	8	29	22

De los 58 planes, 21 se consideran poco robustos y 37 robustos, según IRSCC. Desde el punto de vista radiobiológico la poca robustez de un plan aumenta considerablemente la probabilidad de no ser aceptado por métricas relacionadas con los índices radiobiológicos modificados, con lo que IR puede ser un buen índice a la hora de aceptar un plan realizado por el SPT.

Con respecto a las métricas relacionadas con las diferencias en dosis el índice de robustez carece de capacidad predictiva (tabla 7.XII)

Tabla 7.XII. Distribución de la robustez de planes con respecto a las variables de diferencia en dosis basada en objetivos, junto con la probabilidad de que el plan resulte positivo según PASA_MCescOBJ y PASA_CMPMOBJ

	NOPASA_MCescOBJ		PASA_MCescOBJ	PROBABILIDAD DE NO PASAR %
	POCOROBUSTO	21	5	16
ROBUSTO	37	12	25	32
	NOPASA_CMPMOBJ		PASA_CMPMOBJ	PROBABILIDAD DE NO PASAR %
	POCOROBUSTO	21	10	11
ROBUSTO	37	14	23	38

Se aprecia que un índice de robustez bajo no significa una probabilidad alta de no pasar las métricas MCescOBJ y CMPMOBJ, por lo que no se puede concluir que sea un indicativo de mal pronóstico con respecto a esas métricas.

Con respecto a PASA_IR, los resultados obtenidos se presentan en la tabla 7.XIII Se muestra la tabla de contingencia de PASA_TCPMOD3 con respecto a PASA_IR. Se ha escogido dicha variable porque es la variable que mejor significación mostró en la comparación entre PASA_TCPMPDX y PASA_IR (Chi-cuadrado = 12,389 y significación=0).

Tabla 7.XIII. Tabla de contingencia para la variable PASA_IR frente PASA_TCPMOD3

		PASA_IR		Total
		0	1	0
PASA_TCPMOD3	0	14	13	27
	1	3	28	31
Total		17	41	58

El principal problema de PASA_IR es que las desviaciones globales de dosis en todo el plan juegan un papel contradictorio en la métrica, planes que tienen una pérdida de dosis empeoran la TCP pero mejoran la NTCP con lo que ΔIR puede ser incluso mayor que 0 para planes que tienen realmente problemas de cobertura.

7.6. Propuesta de verificación de planes de IMRT basada en las métricas analizadas

La nueva colección de métricas expuestas en el presente trabajo conduce a un replanteamiento de la verificación de tratamientos de IMRT que tenga en cuenta las particularidades de cada una. Por un lado se cuenta con métricas basadas en diferencias de dosis, por otro lado, con aquellas que tienen en cuenta los índices radiobiológicos. Todas ellas se han tratado de resumir en uno o muy pocos parámetros. Se ha encontrado una

función, en particular, que predice con suficiente efectividad los resultados combinados de todo el proceso. Es la función global de diferencias en dosis (FG). En la tabla 7.XIV se resume su capacidad predictiva, utilizando de nuevo los resultados del análisis ROC.

Tabla 7.XIV. Capacidad predictiva de la variable continua función global de diferencias en dosis con respecto a una representación significativa de variables binarias. Cuando la variable binaria proviene de CPM se usa FGCMPM y cuando proviene de MCesc se usa FGMcsc

VARIABLE BINARIA	AUC	PUNTO DE CORTE	SENSIBILIDAD	ESPECIFICIDAD
PASA_MCesc	0,91	6,375	77,8	90
PASA_CPM	0,962	6,325	88,5	93,7
PASA_MCescOBJ	0,891	6,37	76,5	87,8
PASA_CPMOBJ	0,903	6,32	87,5	88,2
PASA_TCPCPMMOD3	0,884	6,325	81,5	90,3
PASA_TCPCPMMOD5	0,885	6,325	85,7	81,1
PASA_TCPCPM5	0,869	8	92,9	79,5
PASA_TCPCPM9	0,879	8	100	75
PASA_TCPMcsc3	0,75	4,09	84,2	59
PASA_TCPMcsc5	0,808	4,34	100	56
PASA_TCPMcsc9	0,945	9,31	100	85,5
PASA_GAMMA5	0,874	4,22	85,3	70,8
PASA_GAMMA10	0,923	6,62	82,3	88,2

A la vista de la tabla y de los resultados de Chi-cuadrado de la tabla 7.IV para PASA_FG6,35, queda de manifiesto la polivalencia de FG al ser utilizada como variable predictora en diferentes contextos. El establecimiento del punto de corte, como queda dicho en anteriores apartados, depende de la función de coste que se establezca. No obstante, un punto de corte como el establecido, 6,35, parece bastante adecuado.

En la tabla 7.XV se muestra la tabla de contingencia correspondiente a las variables PASA_FG6,35CPM y PASA_CPMOBJ.

Tabla 7.XV. Tabla de contingencia PASA_FG6,35CMPM * PASA_CMPMOBJ				
		PASA_CMPMOBJ		Total
		0	1	0
PASA_FG6,35CMPM	0	21	4	25
	1	3	30	33
Total		24	34	58

El estudio de los planes rechazados por la variable derivada de la función global mostró que 4 planes que no pasan la prueba de FG corresponden a problemas de sobredosificación que no influyen en las restricciones pero que indican que algo va mal en la verificación. El parámetro PPBC estaba fuera de tolerancia en 3 de ellos. Los 3 planes que no pasan PASA_CMPMOBJ pero si pasan PASA_FG6,35CMPM tienen pérdidas de cobertura del PTV principal que están fuera de tolerancia pero muy cerca del -3% y FG no es sensible a ellas. Como ya sabemos estos planes son negativos también para PASA_TCPMOD5, por lo que pueden aceptarse desde un punto de vista radiobiológico.

Otra manera de verlo consiste en utilizar la variable binaria PASA_FG6,35, en lugar de la continua FG, para construir una curva ROC en función del uso de diferentes umbrales X en PASA_TCPX y en PASA_TCPMODX. Se analizaron los dos algoritmos que usan las medidas de Compass, CMPM y MCesc. Los umbrales de rechazo X tomaron los valores 1,2,3,5,7,9,12 y 15 tanto para Δ TCP como para Δ NTCP simultáneamente en todos los volúmenes estudiados. Para el algoritmo CMPM se obtuvo una AUC=0,93 con el punto de corte en 3 y para el algoritmo MCesc una AUC=0,8 y un punto de corte en 2. El resultado puede verse en la figura 7.30. En ella se observa una buena capacidad predictora de PASA_TCPMODX sobre la distribución de positivos y negativos que genera PASA_FG6,35, siendo mejor el resultado para el algoritmo CMPM que para MCesc.

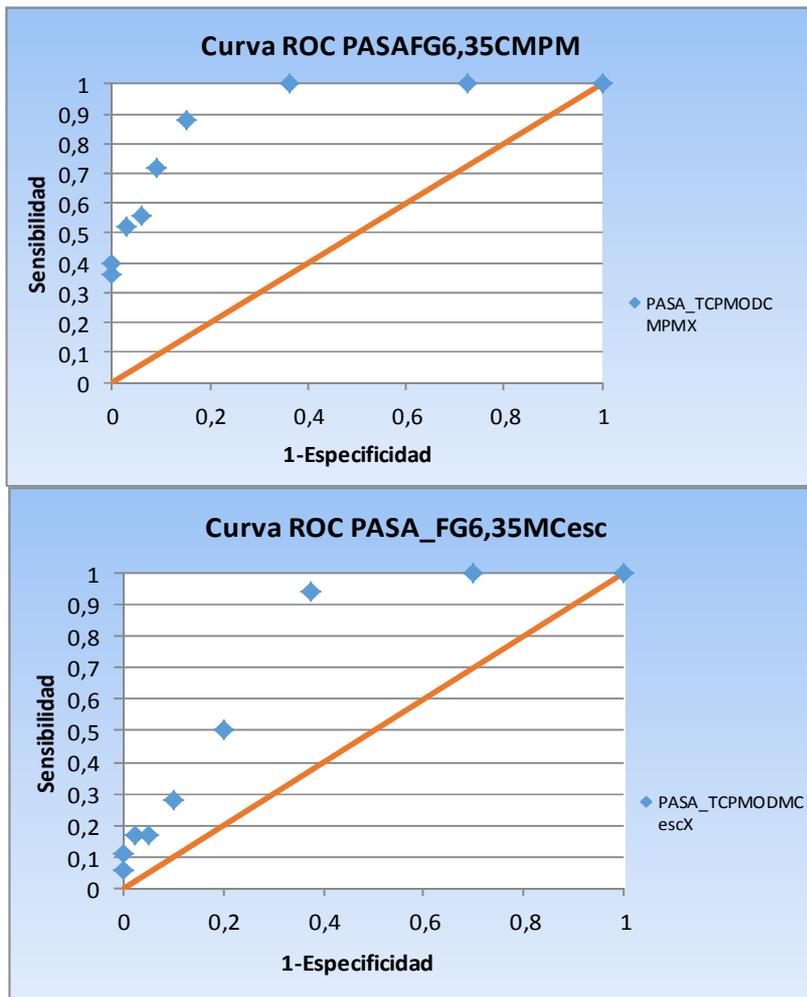


Figura 7.30.- Curvas ROC para la variable PASA_TCPMODX usando PASA_FG6,35 como sistema clasificador, tanto para CPM (arriba) como para MCesc (abajo). Se cambia el umbral X de rechazo pasando por los valores Δ TCP ó Δ NTCP = 1,2,3,5,7,9,12 y 15. Para CPM AUC=0,93 PC=3 y para MCesc AUC=0,8 y PC=2

De este modo FG parece guardar un equilibrio adecuado entre métricas basadas en diferencias de dosis y métricas basadas en índices radiobiológicos modificados, y, por tanto, encajaría muy bien en el papel de parámetro resumen de la verificación, que era lo que se estaba buscando. Además, el uso de FG evita tener que procesar los histogramas para obtener los índices radiobiológicos con el consiguiente ahorro de tiempo. Se prefiere también el uso de FG al de la función gamma global por órganos, que también se ha mostrado muy significativa, porque en general no siempre se dispone de un algoritmo para calcularla.

Resumiendo, los resultados obtenidos con diferentes métricas para el conjunto de planes estudiado, considerado significativo, sugiere un diagrama de flujo ideal tal como el expresado en la figura 7.31. Nótese que el diagrama es independiente del algoritmo utilizado para evaluar los HDV de verificación. El valor de corte para FG, 6,35, es un valor que depende del umbral que se elija en la tabla 6.IV. para las métricas de diferencias de dosis (3% en la tabla).

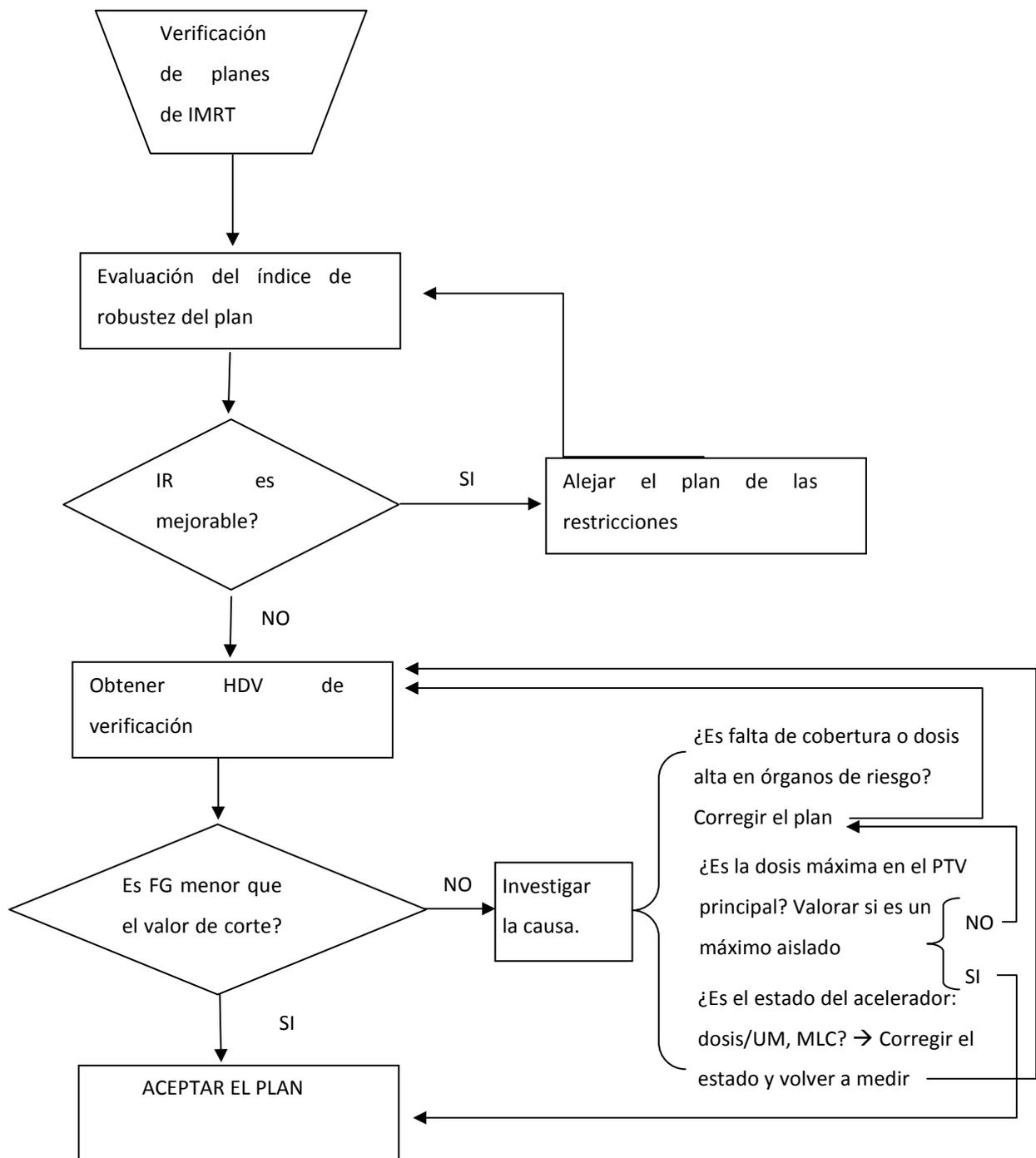


Figura 7.31. Diagrama de flujo para el proceso de verificación de IMRT basado en métricas provenientes del análisis de HDV. El proceso no depende del algoritmo y el valor de corte para la función FG depende de los umbrales escogidos en la tabla 6.IV.

8. Conclusiones

Las conclusiones se diferencian entre aquellas que solo son aplicables a los sistemas y tipo de IMRT utilizados en este trabajo y aquellas que son aplicables de manera general.

8.1. Conclusiones generales

1. Se ha presentado un método novedoso que permite buscar métricas que sean significativas para la verificación de planes de tratamiento de IMRT, construidas a partir de otras conocidas y utilizando el análisis ROC.
2. El algoritmo de cálculo utilizado en la reconstrucción de la dosis en el paciente tiene influencia en el resultado final de la verificación, teniendo una incertidumbre asociada que debe ser tomada en cuenta a la hora de aceptar o rechazar planes. Se encuentran diferencias significativas entre algoritmos de cálculo diferentes.
3. Se ha visto que las métricas de diferencias en dosis pueden ser representadas convenientemente por un parámetro único como el definido en el presente trabajo, función global de diferencias en dosis (FG), con un punto de corte óptimo que puede partir del análisis ROC de la métrica. Dicho valor de corte depende de la tolerancia establecida en la métrica de diferencias en dosis. Para una tolerancia del 3% de desviaciones en dosis global entre histograma de verificación y el proveniente del TPS en determinados puntos de evaluación, el punto de corte para FG resultó ser 6,35.
4. Los puntos de corte obtenidos con el análisis ROC pueden variar en función de la posición del mínimo de una función de coste que se asocie a la diferencia de valor entre un falso positivo y un falso negativo, balance a establecer en cada centro.
5. Las métricas basadas en TCP/NTCP no se correlacionan bien con las basadas en diferencia de dosis. La razón es que, en numerosas ocasiones, las dosis de

restricción impuestas a un plan, tanto para PTVs como para OARs, están lejos de las dosis que suponen cambios significativos en los índices radiobiológicos, y, por tanto, estos son insensibles a diferencias grandes de dosis que puede haber entre la dosis de verificación y la de referencia. Es posible proponer métricas TCP/NTCP modificadas que sitúan el gradiente de la curva de respuesta radiobiológica en el entorno de los valores de dosis de restricción del plan, mejorándose en ese caso su correlación con las métricas de diferencias de dosis.

6. Las métricas basadas en el incumplimiento de objetivos tienen una limitación inherente a su construcción: si el plan aceptado cumple las restricciones impuestas con escaso margen, su robustez es mala y por tanto las posibilidades de incumplir dichos objetivos en la verificación son altas. En ese sentido, una métrica basada en TCP/NTCP modificado describe con exactitud dichas posibilidades de incumplimiento. En cualquier caso, siempre es preferible la elección de planes robustos desde el punto de vista radiobiológico, lo que puede ser tenido en cuenta antes de aceptar un plan. En particular, se ha propuesto en este trabajo un índice de robustez, con valores entre 0 y 1, que puede ser calculado para cada plan y que es un indicativo de la probabilidad de que un plan incumpla métricas relacionadas con los índices radiobiológicos.
7. Las métricas basadas en la función gamma 3D global, para cada uno de los volúmenes considerados en la planificación, tienen una clara correlación con las métricas de diferencias en dosis, y permiten predecir con mucha sensibilidad qué planes resultarán rechazados en las verificaciones. Aunque no son sensibles al signo de la diferencia de dosis este hecho no produce un aumento de falsos positivos con respecto a las métricas de diferencias en dosis. Es recomendable su presencia en el informe de resultados de una verificación.
8. Si se incumplen métricas de diferencias de dosis es posible que el uso del plan para tratamiento no suponga ningún problema, si no se superan

ninguno de los niveles de restricción establecidos para el mismo, por lo que se recomienda el uso en segunda instancia de métricas relacionadas con el incumplimiento de objetivos para aceptar o rechazar un plan.

9. Los sistemas de verificación y medida tienen incertidumbres asociadas. Incluso si se realiza un proceso de puesta en marcha con una verificación cuidadosa de todos los parámetros que puedan afectar a los resultados obtenidos, los planes que son aceptados y rechazados pueden variar según el sistema empleado.
10. Como conclusión acerca de la verificación de planes, se propone un análisis que parta de planes que sean lo más robustos posibles, evaluándose posteriormente la función global de diferencias en dosis, conjuntamente con la investigación, si es el caso, de las circunstancias que producen el rechazo del plan en dicha función.

8.2. Conclusiones particulares

1. El colimador Optifocus utilizado en el presente trabajo, presenta una variabilidad en su comportamiento que tiene gran influencia en la verificación de los planes de tratamiento de IMRT. Es necesario un control exhaustivo y semanal del mismo, junto con acciones correctoras, para que la dosis media de los PTVs principales no desvíe por encima del $\pm 1,5\%$ debido al efecto del posicionamiento de láminas. La tolerancia se establece para un parámetro promedio denominado PPBC que proviene del análisis de un test de campos adyacentes, para el cuál se ha fijado un rango de variación entre el 93% y el 113%.
2. El algoritmo de superposición de cono colapsado (SCC) que permite calcular la dosis en pacientes para su uso clínico en casos de 3DCRT e IMRT estática, que ha sido desarrollado y verificado, tiene una buena exactitud y se

comporta de acuerdo con las características de otros existentes en el mercado.

3. El sistema de verificación Monte Carlo, que incluye las características particulares del colimador Optifocus, ha sido desarrollado y se ha verificado que se comporta de manera adecuada. El módulo de BEAM SMLC, que permite simular la estructura de machihembrado del MLC, ha sido construido y está disponible bajo petición. Además la aplicación de cálculo, denominada MCVerif, que ha sido adaptada para manejar las imágenes TC del paciente y los ficheros de espacio de fase asociados a la IMRT, permite el cálculo automatizado de la dosis usando un algoritmo MC.
4. El software COMPASS se ha mostrado útil a la hora de verificar los tratamientos, ya que proporciona información añadida en forma de histogramas dosis volumen a partir de medidas realizadas con una matriz de cámara de ionización. Estos histogramas pueden ser procesados para obtener diferentes métricas.
5. La operación de escalado de la dosis para transformar los resultados de CMPC en CMPM tiene limitaciones y no produce resultados satisfactorios en los puntos de evaluación de los histogramas de los OARs.
6. Delta4 y Compass, en la manera en que han sido empleados en este trabajo, han discrepado en la clasificación que realizan de los planes, rechazados o no, en un número significativo de casos.

9. Referencias

1. Low D a, Moran JM, Dempsey JF, Dong L, Oldham M. Dosimetry tools and techniques for IMRT. *Med Phys*. 2011;38(3):1313–38.
2. Bayouth JE, Wendt D, Morrill SM. MLC quality assurance techniques for IMRT applications. *Med Phys* [Internet]. 2003 [cited 2011 Jan 12];30(5):743. Available from: <http://link.aip.org/link/MPHYA6/v30/i5/p743/s1&Agg=doi>
3. Klein EE, Hanley J, Bayouth J, Yin F-F, Simon W, Dresser S, et al. Task Group 142 report: Quality assurance of medical accelerators. *Med Phys* [Internet]. 2009 [cited 2014 Jan 21];36(9):4197. Available from: <http://link.aip.org/link/MPHYA6/v36/i9/p4197/s1&Agg=doi>
4. ESTRO. Independent dose calculations. Concepts and models. *Estro booklet n°10*. 2010.
5. Ezzell G a., Galvin JM, Low D, Palta JR, Rosen I, Sharpe MB, et al. Guidance document on delivery, treatment planning, and clinical implementation of IMRT: Report of the IMRT subcommittee of the AAPM radiation therapy committee. *Med Phys* [Internet]. 2003;30(8):2089. Available from: <http://link.aip.org/link/MPHYA6/v30/i8/p2089/s1&Agg=doi>
6. Visser R, Wauben DJL, de Groot M, Godart J, Langendijk J a, van't Veld a a, et al. Efficient and reliable 3D dose quality assurance for IMRT by combining independent dose calculations with measurements. *Med Phys* [Internet]. 2013 Feb [cited 2014 Mar 17];40(2):021710. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23387733>
7. Godart J, Korevaar EW, Visser R, Wauben DJL, Van't Veld a a. Reconstruction of high-resolution 3D dose from matrix measurements: error detection capability of the COMPASS correction kernel method. *Phys Med Biol* [Internet]. 2011 Aug 7 [cited 2014 Mar 17];56(15):5029–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21772084>
8. Bakhtiari M, Parniani A, Lerma F, Reynolds S, Jordan J, Sedaghat A, et al. Evaluation of a software system for estimating planned dose error in patients, based on planar IMRT QA measurements. *Radiol Oncol* [Internet]. 2014 Mar [cited 2014 Mar 17];48(1):87–93. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3908853&tool=pmcentrez&rendertype=abstract>
9. Carrasco P, Jornet N, Latorre A, Eudaldo T, Ruiz A, Ribas M. 3D DVH-based metric analysis versus per-beam planar analysis in IMRT pretreatment verification. *Med Phys* [Internet]. 2012 Aug;39(8):5040–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22894429>
10. Korevaar EW, Wauben DJL, van der Hulst PC, Langendijk J a, Van't Veld A a. Clinical introduction of a linac head-mounted 2D detector array based quality assurance system in head and neck IMRT. *Radiother Oncol* [Internet]. Elsevier Ireland Ltd; 2011 Sep [cited 2014 Mar 17];100(3):446–52. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21963288>

11. Feygelman V, Zhang G, Stevens C, Nelms BE. Evaluation of a new VMAT QA device, or the “X” and “O” array geometries. *J Appl Clin Med Phys* [Internet]. 2011 Jan;12(2):3346. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21587178>
12. Van Elmpt WJC, Nijsten SMJJG, Dekker AL a. J, Mijnheer BJ, Lambin P. Treatment verification in the presence of inhomogeneities using EPID-based three-dimensional dose reconstruction. *Med Phys* [Internet]. 2007 [cited 2014 Mar 17];34(7):2816. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/34/7/10.1118/1.2742778>
13. Olaciregui-Ruiz I, Rozendaal R, Mijnheer B, van Herk M, Mans a. Automatic in vivo portal dosimetry of all treatments. *Phys Med Biol* [Internet]. 2013 Nov 21 [cited 2014 Mar 7];58(22):8253–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24201085>
14. Nijsten SMJJG, van Elmpt WJC, Jacobs M, Mijnheer BJ, Dekker a. L a. J, Lambin P, et al. A global calibration model for a-Si EPIDs used for transit dosimetry. *Med Phys* [Internet]. 2007 [cited 2014 Jan 30];34(10):3872. Available from: <http://link.aip.org/link/MPHYA6/v34/i10/p3872/s1&Agg=doi>
15. Van Elmpt W, Nijsten S, Mijnheer B, Dekker A, Lambin P. The next step in patient-specific QA: 3D dose verification of conformal and intensity-modulated RT based on EPID dosimetry and Monte Carlo dose calculations. *Radiother Oncol*. 2008;86(1):86–92.
16. Lee L, Le Q-T, Xing L. Retrospective IMRT Dose Reconstruction Based on Cone-Beam CT and MLC Log-File. *Int J Radiat Oncol* [Internet]. 2008 Feb [cited 2015 Apr 15];70(2):634–44. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0360301607043751>
17. Zhen H, Nelms BE, Tome W a. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. *Med Phys* [Internet]. 2011 Oct [cited 2014 Feb 4];38(10):5477–89. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21992366>
18. McKenzie EM, Balter P a, Stingo FC, Jones J, Followill DS, Kry SF. Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans. *Med Phys* [Internet]. 2014 Dec [cited 2015 Feb 3];41(12):121702. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25471949>
19. Zhen H, Nelms BE, Tomé W a. On the use of biomathematical models in patient-specific IMRT dose QA. *Med Phys* [Internet]. 2013 Jul [cited 2014 Jan 24];40(7):071702. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23822406>
20. Low D a, Harms WB, Mutic S, Purdy J a. A technique for the quantitative evaluation of dose distributions. *Med Phys* [Internet]. 1998 May;25(5):656–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9608475>
21. Low D a., Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Med Phys* [Internet]. 2003 [cited 2014 Mar 14];30(9):2455. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/30/9/10.1118/1.1598711>
22. Stock M, Kroupa B, Georg D. Interpretation and evaluation of the γ index and the γ index angle for the verification of IMRT hybrid plans. *Phys Med Biol* [Internet]. 2005 Feb 7 [cited

2014 Mar 3];50(3):399–411. Available from: <http://stacks.iop.org/0031-9155/50/i=3/a=001?key=crossref.46f6a8c85ce4269eb09216bc97a7f381>

23. Nelms BE, Zhen H, Tomé W a. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys* [Internet]. 2011 [cited 2014 Mar 19];38(2):1037. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/38/2/10.1118/1.3544657>
24. Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V, Gabriele P. Pretreatment patient-specific IMRT quality assurance : A correlation study. *Med Phys* [Internet]. 2012;39(12):7626–34. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/39/12/10.1118/1.4767763>
25. Kruse JJ. On the insensitivity of single field planar dosimetry to IMRT inaccuracies. *Med Phys* [Internet]. 2010 Jun [cited 2014 Mar 29];37(6):2516–24. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/37/6/10.1118/1.3425781>
26. Waghorn BJ, Meeks SL, Langen KM. Analyzing the impact of intrafraction motion: Correlation of different dose metrics with changes in target D95%. *Med Phys* [Internet]. 2011 [cited 2014 Mar 29];38(8):4505. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/38/8/10.1118/1.3605633>
27. Pulliam KB, Huang JY, Howell RM, Followill D, Bosca R, O’Daniel J, et al. Comparison of 2D and 3D gamma analyses. *Med Phys* [Internet]. 2014 Feb [cited 2014 Feb 26];41(2):021710. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24506601>
28. Rogers DWO. BEAM: A Monte Carlo code to simulate radiotherapy treatment units. *Med Phys* [Internet]. 1995 [cited 2014 Mar 17];22(5):503. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/22/5/10.1118/1.597552>
29. Sempau J, Wilderman SJ, Bielajew a F. DPM, a fast, accurate Monte Carlo code optimized for photon and electron radiotherapy treatment planning dose calculations. *Phys Med Biol* [Internet]. 2000 Aug;45(8):2263–91. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10958194>
30. Tsiakalos MF, Stathakis S, Plataniotis G a, Kappas C, Theodorou K. Monte Carlo dosimetric evaluation of high energy vs low energy photon beams in low density tissues. *Radiother Oncol* [Internet]. 2006 Apr [cited 2014 Mar 17];79(1):131–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16618511>
31. Yamamoto T, Mizowaki T, Miyabe Y, Takegawa H, Narita Y, Yano S, et al. An integrated Monte Carlo dosimetric verification system for radiotherapy treatment planning. *Phys Med Biol* [Internet]. 2007 Apr [cited 2011 Feb 3];52(7):1991–2008. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17374923>
32. Asuni G, van Beek T a, Venkataraman S, Popescu I a, McCurdy BMC. A Monte Carlo tool for evaluating VMAT and DIMRT treatment deliveries including planar detectors. *Phys Med Biol* [Internet]. 2013 Jun 7 [cited 2014 Mar 17];58(11):3535–50. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23640066>

33. Verhaegen F, Seuntjens J. Monte Carlo modelling of external radiotherapy photon beams. *Phys Med Biol* [Internet]. 2003 Nov;48(21):R107–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14653555>
34. Nijsten SMJJG, Mijnheer BJ, Dekker AL a J, Lambin P, Minken AWH. Routine individualised patient dosimetry using electronic portal imaging devices. *Radiother Oncol* [Internet]. 2007 Apr [cited 2014 Mar 17];83(1):65–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17383761>
35. Blake SJ, McNamara AL, Deshpande S, Holloway L, Greer PB, Kuncic Z, et al. Characterization of a novel EPID designed for simultaneous imaging and dose verification in radiotherapy. *Med Phys* [Internet]. 2013 Sep [cited 2014 Mar 7];40(9):091902. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24007153>
36. Sabet M, Rowshanfarzad P, Vial P, Menk FW, Greer PB. Transit dosimetry in IMRT with an a-Si EPID in direct detection configuration. *Phys Med Biol* [Internet]. 2012 Aug 7 [cited 2014 Mar 4];57(15):N295–306. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22801089>
37. Bailey DW, Kumaraswamy L, Bakhtiari M, Malhotra HK, Podgorsak MB. EPID dosimetry for pretreatment quality assurance with two commercial systems. *J Appl Clin Med Phys* [Internet]. 2012 Jan;13(4):3736. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22766944>
38. Xing L, Li JG. Computer verification of fluence map for intensity modulated radiation therapy. *Med Phys* [Internet]. 2000 Sep;27(9):2084–92. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11011737>
39. Laliena V, García-Romero A. Monte Carlo modeling of the Siemens Optifocus multileaf collimator. *Phys Med* [Internet]. Elsevier Ltd; 2015 Feb 13 [cited 2015 Mar 23];31(3):301–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25687417>
40. García-Romero A, Canellas-Anoz M, Lardies-Fleta D. Desarrollo y verificación Monte Carlo de un algoritmo de superposición de cono colapsado para cálculo de haces de fotones en radioterapia. *Rev Física Médica*. 2009;10(3):187–98.
41. Vial P, Greer PB, Oliver L, Baldock C. Initial evaluation of a commercial EPID modified to a novel direct-detection configuration for radiotherapy dosimetry. *Med Phys* [Internet]. 2008 [cited 2015 Mar 26];35(10):4362. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/35/10/10.1118/1.2975156>
42. McCurdy BMC, Greer PB. Dosimetric properties of an amorphous-silicon EPID used in continuous acquisition mode for application to dynamic and arc IMRT. *Med Phys* [Internet]. 2009 [cited 2015 Mar 26];36(7):3028. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/36/7/10.1118/1.3148822>
43. McDermott LN, Nijsten SMJJG, Sonke J-J, Partridge M, van Herk M, Mijnheer BJ. Comparison of ghosting effects for three commercial a-Si EPIDs. *Med Phys* [Internet]. 2006 [cited 2015 Mar 26];33(7):2448. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/33/7/10.1118/1.2207318>

44. Ripol Valentín O, García Romero A, Hernández Vitoria A, Jiménez Albericio J, Cortés Rodicio J, Millán Cebrián E, et al. Caracterización dosimétrica de un dispositivo electrónico de imagen portal (EPID) y desarrollo de un modelo simple de dosimetría portal Dosimetric characterization of an electronic portal imaging device (EPID) and development of a portal dosimetry sim. *Rev Física Médica*. 2010;11(3):199–210.
45. Poppinga D, Schoenfeld a a, Doerner KJ, Blanck O, Harder D, Poppe B. A new correction method serving to eliminate the parabola effect of flatbed scanners used in radiochromic film dosimetry. *Med Phys* [Internet]. 2014 Feb [cited 2014 Mar 1];41(2):021707. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24506598>
46. Paelinck L, De Neve W, De Wagter C. Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry. *Phys Med Biol* [Internet]. 2007 Jan 7 [cited 2014 Mar 17];52(1):231–42. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17183138>
47. Devic S, Seuntjens J, Sham E, Podgorsak EB, Schmidlein CR, Kirov AS, et al. Precise radiochromic film dosimetry using a flat-bed document scanner. *Med Phys* [Internet]. 2005;32(7):2245. Available from: <http://link.aip.org/link/MPHYA6/v32/i7/p2245/s1&Agg=doi>
48. Van Battum LJ, Hoffmans D, Piersma H, Heukelom S. Accurate dosimetry with GafChromic™ EBT film of a 6 MV photon beam in water: What level is achievable? *Med Phys* [Internet]. 2008 [cited 2011 Feb 18];35(2):704. Available from: <http://link.aip.org/link/MPHYA6/v35/i2/p704/s1&Agg=doi>
49. Lin H, Huang S, Deng X, Zhu J, Chen L. Comparison of 3D anatomical dose verification and 2D phantom dose verification of IMRT/VMAT treatments for nasopharyngeal carcinoma. *Radiat Oncol* [Internet]. 2014 Mar 7 [cited 2014 Mar 20];9(1):71. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24606879>
50. Herzen J, Todorovic M, Cremers F, Platz V, Albers D, Bartels a, et al. Dosimetric evaluation of a 2D pixel ionization chamber for implementation in clinical routine. *Phys Med Biol* [Internet]. 2007 Feb 21 [cited 2014 Aug 6];52(4):1197–208. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17264380>
51. Ahnesjö A. Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media. *Med Phys*. 1989;16(Jul/Aug):577–92.
52. Seltzer S, Berger M. Bremsstrahlung Energy Spectra from Electrons with Kinetic Energy 1 keV-10 GeV Incident on Screened Nuclei and Orbital Electrons of Neutral Atoms with Z = 1-100. *At Data Nucl Data Tables*. 1986;35:345–418.
53. Press W, Teukolsky S, Vetterling W, Flannery B. *The art of scientific computing*. Third. Cambridge University Press; 2007.
54. Ahnesjö A, Andreo P. Determination of effective bremsstrahlung spectra and electron contamination for photon dose calculations. *Phys Med Biol*. 1989;34(10):1451–64.

55. Mackie TR, Bielajew a F, Rogers DWO, Battista JJ. Generation of photon energy deposition kernels using the EGS Monte Carlo code. *Phys Med Biol*. 1988;33(1):1–20.
56. Mainegra-Hing E, Rogers DWO, Kawrakow I. Calculation of photon energy deposition kernels and electron dose point kernels in water. *Med Phys* [Internet]. 2005 [cited 2014 Aug 14];32(3):685. Available from: <http://scitation.aip.org/content/aapm/journal/medphys/32/3/10.1118/1.1861412>
57. Ahnesjö a, Aspradakis MM. Dose calculations for external photon beams in radiotherapy. *Phys Med Biol* [Internet]. 1999 Nov;44(11):R99–155. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10588277>
58. Storchi PR, van Battum LJ, Woudstra E. Calculation of a pencil beam kernel from measured photon beam data. *Phys Med Biol*. 1999;44(12):2917–28.
59. Storchi P, Woudstra E. Calculation of the absorbed dose distribution due to irregularly shaped photon beams using pencil beam kernels derived form basic beam data. *Phys Med Biol*. 1996;41(4):637–56.
60. Storchi P, Woudstra E, Verlinde P, Johansson K a, Samuelsson a. Calculation of absorbed dose distributions from dynamic wedges. *Phys Med Biol* [Internet]. 1998 Jun;43(6):1497–506. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9651020>
61. Kawrakow I. Accurate condensed history Monte Carlo simulation of electron transport . I . EGSnrc , the new EGS4 version Accurate condensed history Monte Carlo simulation of electron transport . I . EGS nrc , the new EGS4 version. *Med Phys*. 2000;27(2000):485–98.
62. Walters B, Kawrakow I, Rogers DWO. *DOSXYZnrc Users Manual*. 2011;
63. Sheikh-Bagheri D, Rogers DWO. Sensitivity of megavoltage photon beam Monte Carlo simulations to electron beam and other parameters. *Med Phys*. 2002;29(3):379–90.
64. Tzedakis A, Damilakis JE, Mazonakis M, Stratakis J, Varveris H, Gourtsoyiannis N. Influence of initial electron beam parameters on Monte Carlo calculated absorbed dose distributions for radiotherapy photon beams. *Med Phys*. 2004;31(4):907–13.
65. Almberg SS, Frengen J, Kylling A, Lindmo T. Monte Carlo linear accelerator simulation of megavoltage photon beams: independent determination of initial beam parameters. *Med Phys* [Internet]. 2012 Jan [cited 2014 Jan 24];39(1):40–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22225273>
66. Kairn T, Aland T, Franich RD, Johnston PN, Kakakhel MB, Kenny J, et al. Adapting a generic BEAMnrc model of the BrainLAB m3 micro-multileaf collimator to simulate a local collimation device. *Phys Med Biol* [Internet]. 2010 Sep 7 [cited 2014 Mar 17];55(17):N451–63. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20702922>
67. Kutcher G.J., Burman C, Brewster L., Goitein M, Mohan R. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *Int J Radiat Oncol Biol Phys*. 1991;21:137–46.

68. Niemierko a. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Med Phys*. 1997;24(1):103–10.
69. Group AT, Committee TP. The Use and QA of Biologically Related Models for Treatment Planning Report of AAPM Task Group 166. 2012.
70. Zygmanski P, Kung JH. Method of identifying dynamic multileaf collimator irradiation that is highly sensitive to a systematic MLC calibration error. *Med Phys*. 2001;28(11):2220–6.
71. Olofsson N. Evaluation of IMRT beam complexity metrics to be used in the IMRT QA process. University of Gothenburg; 2011.
72. Siebers J V, Keall PJ, Nahum a E, Mohan R. Converting absorbed dose to medium to absorbed dose to water for Monte Carlo based photon beam dose calculations. *Phys Med Biol* [Internet]. 2000 Apr;45(4):983–95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10795986>
73. Fogliata A, Vanetti E, Albers D, Brink C, Clivio A, Knöös T, et al. On the dosimetric behaviour of photon dose calculation algorithms in the presence of simple geometric heterogeneities: comparison with Monte Carlo calculations. *Phys Med Biol* [Internet]. 2007 Mar;52(5):1363–85. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17301460>
74. Carrasco P, Jornet N, Duch M a., Panettieri V, Weber L, Eudaldo T, et al. Comparison of dose calculation algorithms in slab phantoms with cortical bone equivalent heterogeneities. *Med Phys* [Internet]. 2007 [cited 2010 Nov 16];34(8):3323. Available from: <http://link.aip.org/link/MPHYA6/v34/i8/p3323/s1&Agg=doi>
75. Carrasco P, Jornet N, Duch M a., Weber L, Ginjaume M, Eudaldo T, et al. Comparison of dose calculation algorithms in phantoms with lung equivalent heterogeneities under conditions of lateral electronic disequilibrium. *Med Phys* [Internet]. 2004 [cited 2011 Feb 18];31(10):2899. Available from: <http://link.aip.org/link/MPHYA6/v31/i10/p2899/s1&Agg=doi>
76. Gutiérrez FC, Vara CP. Validación de un sistema de control de calidad pre-tratamiento 3D en técnicas volumétricas basado en matrices bidimensionales de detectores. *Rev Física Médica*. 2015;16(1):11–32.