

MODELOS CLINICO-DOSIMETRICOS
PREDICTIVOS DEL CONTROL BIOQUIMICO
DEL CANCER DE PROSTATA
ORGANO-CONFINADO TRATADO MEDIANTE
IMPLANTES RADIOACTIVOS PERMANENTES

Alfredo Polo Rubio
Universidad de Barcelona
Departamento de Ciencias Clínicas
Año 2007

8. DISCUSION

8.1 DE LA VALIDEZ DEL METODO EMPLEADO

∴ Los estudios retrospectivos son relativamente fáciles de realizar (si se comparan con los estudios prospectivos controlados) y se adaptan bien a su presentación como una tesis doctoral ya que el doctorando puede ejercer control sobre todas las fases del proceso.

El objeto principal de nuestro estudio es la serie 1: hemos incluido todos los pacientes tratados con braquiterapia exclusiva con ¹²⁵I en nuestro centro durante el periodo estudiado. Todas la historias clínicas fueron revisadas y se utilizó una base de datos común para almacenar la información.

∴ Lamentablemente existen determinados factores confundidores pre-terapéuticos que pueden inducir a errores o limitar la validez de nuestro estudio. En primer lugar existe una falta de datos de contexto del paciente (comorbilidad, hábitos dietéticos etc.). Por otra parte, una proporción muy grande de pacientes se presentó a nuestra valoración bajo la influencia de algún tipo de tratamiento hormonal. Esto plantea dos problemas: la inadecuada valoración del estadio clínico por nuestra parte y la contaminación de los resultados de supervivencia al introducirse una covariable que puede influir en ella. En tercer lugar, la valoración del grado de Gleason de los tejidos obtenidos por biopsia no se realizó de forma centralizada. Considerando que la mayoría de los pacientes tratados en nuestras series no fueron diagnosticados en nuestra institución, este es un aspecto que debe ser señalado.

∴ Los grupos de estudio se presentaron tras un proceso de selección estricto. Por otro lado, el intervalo de reclutamiento fue relativamente corto (desde septiembre 1999 hasta diciembre de 2003), por lo que las técnicas de braquiterapia aplicadas no presentaron diferencias evidentes. El número de operadores que trataron simultáneamente pacientes en cada momento del intervalo de reclutamiento fue de dos, que aplicaron una técnica de implante muy similar entre si. Y por último, el aspecto más crítico del proceso, de cara a la elaboración de una base de datos coherente, que es la delimitación de los volúmenes de interés, fue llevado a cabo por un solo operador experimentado, de manera que se eliminó la variabilidad inter-operador y se minimizó la variabilidad intra-operador. Por todo ello se puede resumir que el análisis de datos se realizó sobre una serie de pacientes homogénea, tratada con una técnica común según un protocolo terapéutico interno de nuestra institución y evaluados según un criterio consistente.

∴ En conclusión, defendemos la validez del método empleado y de los resultados obtenidos para el caso concreto objeto de nuestro estudio. Nuestros resultados se sitúan entre los publicados en la bibliografía revisada y pensamos que se trata de una aportación objetiva, honesta y útil al cuerpo de conocimientos acerca de la moderna braquiterapia prostática con semillas permanentes.

8.2 DE LA DEFINICION DE CONTROL BIOQUIMICO

∴ El marcador sérico PSA (“prostatic specific antigen”) se usa ampliamente para cribaje, diagnóstico, determinación del pronóstico y selección del tratamiento más apropiado en pacientes afectados de cáncer de próstata localizado. Tras el tratamiento el PSA se usa para determinar la efectividad del mismo. Al inicio de su utilización (inicio de la década de 1990) se establecieron criterios de respuesta basados en los niveles absolutos del PSA. En algunos casos se estableció el criterio de alcanzar un valor normal ($< 4\text{ng/ml}$). En otros casos se establecieron otros niveles ($< 2\text{ng/ml}$ y $< 1\text{ng/ml}$ fueron valores de corte populares). Esto provocó la falta de un criterio común consensuado entre toda la comunidad.

∴ En 1994 la ASTRO comenzó a desarrollar una definición estándar de fallo terapéutico tras la radioterapia externa. En 1996 se estableció un comité para establecer una definición de trabajo del fallo bioquímico tras la radioterapia externa. Dicha definición se publicó en 1997 (68) y permitió que series de diferentes instituciones pudieran ser comparadas.

De forma concisa, la definición de consenso ASTRO define el fallo bioquímico como tres incrementos consecutivos en el valor del PSA tras haber alcanzado su nadir, considerándose la fecha del fallo el punto medio entre la fecha del nadir y la fecha de la primera elevación. Se consideró asimismo fallo bioquímico cualquier elevación que provocara la administración de un tratamiento de rescate.

Esta definición de consenso recomienda que las series de pacientes fueran seguidas al menos 24 meses y que las determinaciones de PSA se obtengan cada 3-4 meses durante los dos primeros años post-tratamiento y cada 6 meses a partir de entonces. De esta forma la estimación de la fecha de recidiva se podría hacer de forma más precisa.

Además, se considera que el fallo bioquímico no es justificación suficiente para iniciar tratamiento adicional ni el fallo bioquímico se ha mostrado como factor predictivo del fallo clínico o de la supervivencia. Sin embargo se considera que es un buen criterio de respuesta para los ensayos clínicos o para reportar resultados.

Pronto se comprobó el éxito de esta definición, y pronto se evidenciaron sus defectos: en primer lugar el fenómeno de “backdating” de difícil traducción al castellano y que quiere decir que el fallo bioquímico se establece en un momento anterior al que se diagnostica (se diagnostica tras el tercer aumento consecutivo pero se establece en el punto medio entre nadir y el primer aumento de PSA). El “backdating” introduce un sesgo que tiende a sobrevalorar el éxito terapéutico cuando el tiempo de seguimiento no es suficientemente largo. El “backdating” hace que las curvas de supervivencia actuarial adquieran un nivel “plateau” precoz, en contraste con las curvas de supervivencia tras cirugía radical, donde la supervivencia libre de recidiva bioquímica continúa declinando con el tiempo, al no usarse el mismo criterio para el fracaso terapéutico. En segundo lugar, la definición de fallo bioquímico no está asociada a supervivencia global, fallo clínico o intervenciones de rescate. Por último, la utilización de otras armas terapéuticas como la hormonoterapia o la braquiterapia cuestionó su validez en estos nuevos contextos.

∴ Para resolver las limitaciones expresadas anteriormente se celebró una segunda conferencia de consenso en enero de 2005 en Phoenix, Arizona para revisar la definición de consenso de la ASTRO. En este caso se reunió un subcomité de la ASTRO con un subcomité de la RTOG (72).

De esta segunda conferencia de consenso emerge un concepto fundamental: las nuevas definiciones se proponen para establecer el fracaso terapéutico en un contexto poblacional, no individual. El contexto individual ha de ser guiado por el criterio clínico. El contexto poblacional es estadístico, se puede traducir en un algoritmo y sirve para comparar resultados tras radioterapia externa o braquiterapia con o sin tratamiento hormonal. Sin embargo la nueva definición no debería ser usada para series quirúrgicas, pacientes tratados con radioterapia de rescate o criocirugía.

La definición de Phoenix dice que un aumento igual o superior a 2ng/ml por encima del nadir puede ser considerado como fracaso bioquímico, independientemente de la utilización de tratamiento hormonal. La fecha de recidiva será en el momento de la constatación de dicho aumento (evitando “backdating”). Los pacientes sometidos a tratamiento de rescate también son considerados en fallo bioquímico.

La reunión de consenso de Phoenix recomienda seguir utilizando el criterio ASTRO tras la administración con radioterapia externa o braquiterapia, si no se ha utilizado tratamiento hormonal. Sin embargo, para evitar los artefactos estadísticos de dicha

Influencia de la definición de fallo bioquímico sobre supervivencia

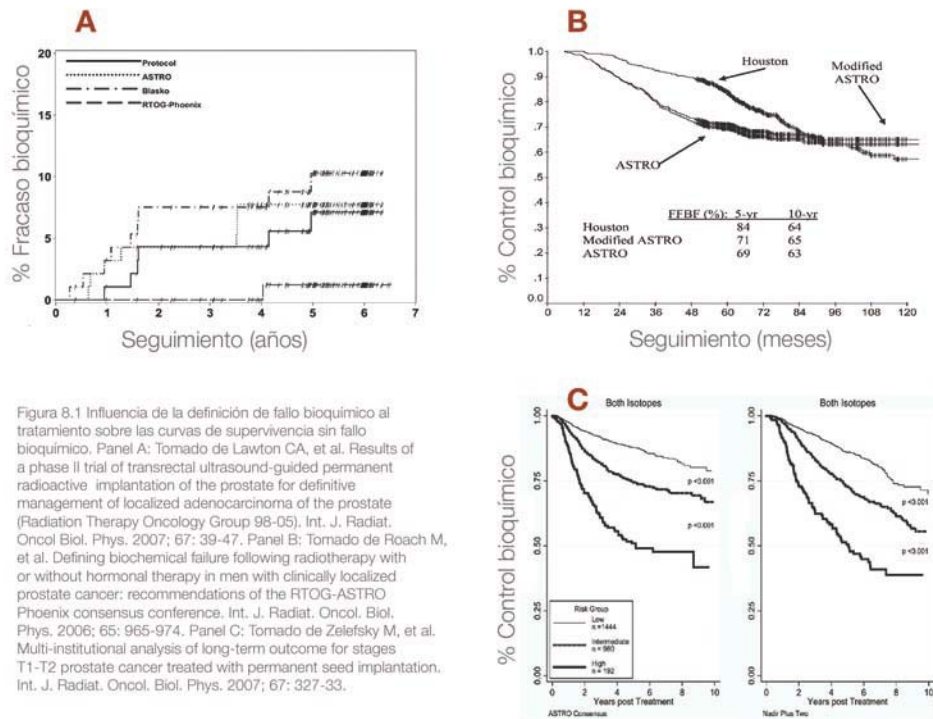


Figura 8.1 Influencia de la definición de fallo bioquímico al tratamiento sobre las curvas de supervivencia sin fallo bioquímico. Panel A: Tomado de Lawton CA, et al. Results of a phase II trial of transrectal ultrasound-guided permanent radioactive implantation of the prostate for definitive management of localized adenocarcinoma of the prostate (Radiation Therapy Oncology Group 98-05). Int. J. Radiat. Oncol. Biol. Phys. 2007; 67: 39-47. Panel B: Tomado de Roach M, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix consensus conference. Int. J. Radiat. Oncol. Biol. Phys. 2006; 65: 965-974. Panel C: Tomado de Zelefsky M, et al. Multi-institutional analysis of long-term outcome for stages T1-T2 prostate cancer treated with permanent seed implantation. Int. J. Radiat. Oncol. Biol. Phys. 2007; 67: 327-33.

definición se recomienda dar resultados de control bioquímico para un intervalo post-tratamiento que sea 2 años menos al de la mediana de seguimiento. Es decir, que si la mediana de seguimiento de la serie es de 5 años, se deberían dar valores de control bioquímico a 3 años. Se este modo se evitará dar resultados sobre-optimistas debido a la falta de seguimiento a largo plazo.

∴ En la figura 8.1 hemos tratado de representar lo que hemos explicado anteriormente de forma gráfica. Se presentan las diferencias en las curvas de supervivencia actuarial al aplicar distintos criterios de fallo bioquímico.

∴ En nuestro caso el análisis de nuestros datos se realizó antes de la publicación del artículo de consenso de Phoenix, por lo que no hemos aplicado el criterio de [nadir +2]. Por otra parte, siguiendo las últimas recomendaciones, al tener nuestra serie un seguimiento medio de 3 años, sólo sería recomendable reportar resultados de control bioquímico a 1 año para evitar el fenómeno de “backdating”. Sin embargo defendemos la validez de nuestro método, ya que las series publicadas adolecen del mismo problema, ya que como hemos dicho las recomendaciones de la reunión de consenso RTOG-Phoenix se publicaron a mitad de 2006. Sin embargo sería recomendable reanalizar nuestros datos en el futuro para incorporar dichas recomendaciones.

8.3 DE LAS CARACTERÍSTICAS PRE-TRATAMIENTO DE LOS PACIENTES

8.3.1 De la histología del tumor

∴ El valor del índice de Gleason es uno de los factores pronósticos más importantes tanto para las series quirúrgicas como rádicas (58, 339, 340, 341, 149, 154, 167). Blasko et al. (167) estudian 230 pacientes tratados con ^{103}Pd como monoterapia encontrando una diferencia significativa ($p = 0.001^*$) al agrupar los pacientes según GS (< 7 vs. ≥ 7). Del mismo modo, el grupo del Mount Sinai de New York (154) analiza 243 pacientes tratados con ^{125}I como tratamiento exclusivo, encontrando diferencias significativas ($p = 0.0003^*$) al estratificar el GS en tres grupos (GS ≤ 6 , GS 7 y GS 8-10).

Nuestra serie 1 es muy homogénea en cuanto a la distribución por el valor del GS. El 87.1% de los pacientes se presentaron con GS 2-6 siendo el GS 6 el valor más abundante (27.7% de los pacientes). El GS 7 se presentó en el 11% de los pacientes y únicamente 2 pacientes (1.8%) se presentaron con GS 8-10. Con este desequilibrio, el análisis estadístico no encontró diferencias significativas en la supervivencia libre de recidiva bioquímica.

∴ El porcentaje de biopsias positivas ha sido exhibido por algunos autores como factor predictivo de recidiva bioquímica tras radioterapia externa o braquiterapia, estableciéndose valores de corte entre 33% y 50% para su asociación con la recidiva bioquímica (126, 342, 124).

En nuestro caso, el porcentaje de biopsias positivas se conoció para 78 pacientes de la serie 1 (69%). No existió sesgo en cuanto a la distribución por estratos de riesgo, pero el control de la procedencia fue difícil de seguir en algunos casos, y por otra parte el porcentaje de histologías revisadas por un patólogo centralizado fue muy escasa. En cualquier caso no se pudo establecer asociación estadística entre el porcentaje de biopsias positivas y la supervivencia libre de recidiva bioquímica.

8.3.2 Del PSA inicial

El PSA inicial (valor de PSA más elevado antes de la intervención terapéutica con braquiterapia) es uno de los factores pronósticos de recidiva y progresión más importantes en la literatura, tanto para las series de RTE como de braquiterapia (343, 344). En nuestro caso se realizó un análisis del PSA como variable continua (ajuste logístico) que mostró significación estadística para el grupo general y para el grupo de riesgo intermedio. A partir del ajuste logístico se establecieron puntos de corte que se introdujeron en el análisis univariado de supervivencia. Para el grupo general el valor de corte fue de 10 ng/ml y para el grupo de riesgo intermedio, 13 ng/ml. En la tabla 8.3_1, que combina datos mostrados anteriormente se muestran las diferencias de supervivencia en el análisis univariado según el grupo de riesgo y el valor de corte. Se presentan los valores para al grupo de bajo riesgo para comparación.

Grupo de riesgo y corte de PSA		2 años	4 años	5 años	p
Bajo riesgo	≤ 10 ng/ml	94.56	89.42	86.95	-
Grupo general	< 10 ng/ml	94.49	89.31	79.08	
	> 10 ng/ml	74.00	65.77	-	
Riesgo intermedio	≤ 13 ng/ml	94.44	86.57	86.57	0.0110*
	> 13 ng/ml	46.88	46.88	46.88	

Tabla 8.3_1 Supervivencia libre de recidiva bioquímica según grupo de riesgo y valor de PSA

En el análisis multivariante de la supervivencia, el PSA inicial fue variable significativa tanto para el grupo general ($p = 0.0091^*$) como para el grupo de riesgo intermedio ($p = 0.0014^*$).

Se comparó la supervivencia libre de recidiva bioquímica para el grupo de pacientes de bajo riesgo ($n = 77$) y el grupo de pacientes con el PSA como único factor de riesgo intermedio que tuvieran valores de PSA > 10 y ≤ 13 ng/ml ($n = 11$) sin encontrar diferencias mesurables por el test de Log-Rank ($p = 0.8553$), de lo que se deduce que para nuestra serie, el grupo de mas bajo riesgo se extiende a pacientes con valor de PSA inicial de hasta 13 ng/ml ($T \leq T2b$, $GS < 70$ y $PSA \leq 13$ ng/ml).

Dentro del grupo de bajo riesgo y teniendo en cuenta la gran homogeneidad para la distribución del PSA (media de 6.65 ng/ml con el 95% de los valores entre 6.20 y 7.11 ng/ml), no se estableció una asociación estadística entre valor del PSA y supervivencia libre de recidiva bioquímica.

La densidad del PSA es una variable muy fácil de determinar, y algunos autores han recomendado su uso fundamentalmente en la etapa de diagnóstico (43) pero también como predictor de recidiva tras prostatectomía radical (345). En nuestro caso la densidad de PSA mostró significación para el grupo general en el análisis univariante, pero no retuvo significación en el multivariante ni en el análisis por subgrupos.

∴ Stock et al. estudian una serie de 97 pacientes tratados mediante braquiterapia exclusiva en la que se presentan los resultados según la estratificación del PSA en 3 grupos ($PSA \leq 10$ ng/ml, $[10 > PSA \leq 20$ ng/ml] y $PSA > 20$ ng/ml). La supervivencia libre de recidiva bioquímica a 2 años fue del 83%, 82% y 58% respectivamente para cada uno de los grupos ($p = 0.02^*$) (196). Nuestros resultados con monoterapia se superponen a los presentados para el grupo de $PSA \leq 10$ ng/ml. En nuestro caso los

valores de PSA > 10 ng/ml se analizan en un solo grupo por lo que los resultados no son comparables.

∴ D'Amico et al. estudian 218 pacientes tratados mediante braquiterapia exclusiva con ^{103}Pd (150). Se establecen tres grupos de PSA (PSA \leq 10ng/ml, [10> PSA \leq 20ng/ml] y PSA > 20ng/ml). La supervivencia libre de recidiva bioquímica a 5 años fue del 85%, 30% y <10% respectivamente para cada uno de los grupos. Nuestros resultados con monoterapia son similares para el grupo de PSA \leq 10 ng/ml, y no son comparables para los valores mayores por la misma razón expresada anteriormente.

∴ Dattoli et al. estudia 73 pacientes tratados mediante combinación de RTE (41 Gy) seguidos de un implante de ^{103}Pd (80 Gy). El valor de PSA se estratificó en dos grupos con un valor de corte de 15 ng/ml. La supervivencia a 18 meses fue del 90% para pacientes con PSA \leq 15ng/ml y de 70% para los pacientes con PSA > 15ng/ml (171). Se ha elegido este trabajo por que muestra un valor de corte similar al obtenido en nuestro análisis (13 ng/ml), aunque usa terapia combinada y tiene un seguimiento inferior al de nuestra serie.

∴ Ragde et al. presentan 152 pacientes tratados con braquiterapia exclusiva (98p) o con tratamiento combinado (54p) en función de la combinación de estadio clínico y valor de GS (151). El valor de PSA inicial no se incluyó en la estratificación por lo que está distribuido entre los dos grupos (PSA > 10 ng/ml en 20 pacientes del grupo 1 y 22 pacientes del grupo 2). La supervivencia libre de recidiva bioquímica a 120 meses fue del 70% en los pacientes con PSA < 10 ng/ml y del 60% en los pacientes con PSA \geq 10 ng/ml. Nuestros resultados son difíciles de comparar con esta serie, por varias razones, entre ellas la definición de fallo bioquímico empleado, pero cabe esperar un mayor número de eventos con el paso del tiempo, ya que en la serie de Ragde no se observa presencia de "plateau" evidente.

8.3.3 Del estadio clínico

La serie 1 incluyó solamente pacientes con alta probabilidad de enfermedad órgano-confinada, por lo que no se incluyó ningún paciente con estadio T3. Hay que tener en cuenta que un paciente con tumor en estadio clínico T3 con PSA \leq 10 ng/ml y GS < 7 se incluiría en un grupo de riesgo intermedio para determinadas clasificaciones, mientras que para otros la presencia de T3 clasificaría al paciente directamente en un grupo de alto riesgo.

Por ello, los pacientes con estadio T2b (tumor palpable bilateral) constituirían el único subgrupo de pacientes de riesgo intermedio por valor de T. En nuestra serie representan el 4.42% (5 pacientes).

Todo lo anterior introduce un sesgo de selección por lo que el análisis estadístico no ha mostrado diferencias significativas en la supervivencia libre de recidiva bioquímica al analizar la serie por estadio clínico.

8.3.4 De la distribución por grupos de riesgo

Como hemos visto anteriormente, la estratificación por grupos de riesgo constituye una herramienta muy potente para reducir la dimensionalidad de las variables pronósticas al combinar tres variables (dos ordinales y una continua) en una sola

ordinal. La estratificación por grupos de riesgo ha sido utilizada en el análisis de las series modernas de braquiterapia y su validez ha sido bien demostrada previamente (148, 149, 150, 155). Su uso permite la comparación multicéntrica de los resultados obtenidos con la técnica y constituye una herramienta de segmentación de bases de datos muy útil.

En nuestro caso hemos utilizado la estratificación del MSKCC descrita inicialmente por Zelefsky para clasificar pacientes tratados mediante RTE (82). La distribución por estratos de riesgo mostró asociación estadística con la supervivencia libre de recidiva bioquímica tanto en el análisis univariante ($p = 0.0326^*$) como en el multivariante ($p = 0.0108^*$). La supervivencia libre de recidiva bioquímica para el grupo de bajo riesgo fue del 94.56%, 89.42% y 86.95% a 2, 4 y 5 años respectivamente. Para el grupo de riesgo intermedio los valores correspondientes fueron del 85.79%, 73.60% y 68.17% respectivamente. En la tabla 8.3_2 se muestra una revisión de series tratadas con braquiterapia como tratamiento exclusivo que reportan resultados a 5 años estratificados por grupos de riesgo.

Autor	Bajo riesgo	Riesgo intermedio
Stock ⁽¹⁴⁹⁾	88.00%	60.00%
D'Amico ⁽¹⁵⁰⁾	85.00%	33.00%
Potters ⁽³⁴⁶⁾	92.00%	74.00%
Zelefsky ⁽¹⁵⁵⁾	88.00%	77.00%
Nuestra serie	86.95%	68.17%

Tabla 8.3_2 Supervivencia libre de recidiva bioquímica según grupo de riesgo

En conclusión, nuestros valores para el grupo de pacientes de bajo riesgo se comparan positivamente con los resultados publicados previamente. Los resultados del grupo de pacientes de riesgo intermedio son más difíciles de comparar por que los criterios no son los mismos, y probablemente se trata de una enfermedad más heterogénea.

8.3.5 Del tratamiento hormonal

En la serie 1, el 61.94% de los pacientes (70/113) habían recibido tratamiento hormonal previo a la braquiterapia. De los 70 pacientes sometidos a bloqueo hormonal, 41 (58.57%) lo fueron con intención citorreductora, y 29 recibieron tratamiento como parte de la estrategia terapéutica de los centros de referencia (que por otra parte no era la del IEO). En el grupo de bajo riesgo el 57.14% de los pacientes estaban bajo la influencia de tratamiento hormonal en el momento de la braquiterapia, mientras que este porcentaje aumentaba hasta el 61.94% en los pacientes de riesgo intermedio.

Esta situación indudablemente conduce a una situación de sesgo, ya que por una parte la serie no se presenta pura a nuestras manos sino alterada por una intervención terapéutica discutible en la mayoría de los casos. Por otra parte no es posible analizar sólo el subconjunto de pacientes que no han recibido tratamiento hormonal por que ello supondría renunciar a más de un tercio de los pacientes, y la pérdida de potencia estadística en el análisis. Por ello se analizó detalladamente la serie con una combinación de técnicas estadísticas (RPA y análisis de la supervivencia) para controlar el efecto de los grupos de riesgo y el valor del PSA inicial (ver el apígrafe 7.4.4 y las figuras 7.24 y 7.25).

De la comparación de las curvas ROC se desprende que el árbol de clasificación que tiene como primer nodo el PSA inicial tiene mejor comportamiento que el que

realiza el primer corte por el valor del grupo de riesgo (0.7255 vs. 0.6784), lo cual es sorprendente y hubiera pasado desapercibido de haber usado un análisis univariante de la supervivencia clásico. Sin embargo el análisis de todas las posibles combinaciones no rindió una situación donde la presencia o no de tratamiento hormonal neoadyuvante haya influido en el control bioquímico de la enfermedad. Por ello se analizó la serie como un todo homogéneo sin análisis por subgrupos que hubieran restado potencia estadística.

∴. Hasta la fecha no existen estudios randomizados que establezcan la utilidad del tratamiento hormonal en el manejo de los pacientes que reciben braquiterapia. Potters et al. (347) publicaron en el año 2000 un estudio de casos-control sobre 620 pacientes que no mostró beneficio de la hormonoterapia para ningún grupo de riesgo, valor de PSA o índice de Gleason o estadio clínico. Sin embargo, el 71% de los pacientes sometidos a tratamiento hormonal recibieron un curso de 3-4 meses, por lo que dicho estudio no es capaz de establecer el valor de la hormonoterapia a largo plazo en combinación con la braquiterapia permanente.

∴. Merrick et al. (348) estudian 425 pacientes tratados con braquiterapia exclusiva o asociación de braquiterapia y radioterapia externa a los que se añade o no tratamiento hormonal. Estableciendo el tratamiento hormonal como variable de estudio se deduce que no hay diferencias de supervivencia libre de recidiva bioquímica entre los pacientes que reciben tratamiento hormonal y los que no. Cuando se hace el mismo análisis por subgrupos los autores refieren que no existen diferencias significativas entre los pacientes sometidos a tratamiento hormonal y los que no lo reciben para ninguno de los grupos de riesgo. Sin embargo, para el subgrupo de alto riesgo la diferencia, aunque no es significativa ($p = 0.09$) muestra una tendencia al beneficio de la adición de tratamiento hormonal (supervivencia libre de recidiva bioquímica a 5 años del 97.1% para el grupo que recibió hormonas y 79.9% para los pacientes que no recibieron hormonas).

∴. En un estudio posterior, el mismo autor reporta sus resultados sobre una serie de 350 pacientes de riesgo intermedio o alto riesgo tratados con braquiterapia exclusiva o asociación de braquiterapia y radioterapia externa a los que se añade o no tratamiento hormonal (349). En el grupo de bajo riesgo no se encontraron diferencias significativas. En el grupo de alto riesgo de encontró una diferencia estadísticamente significativa en la supervivencia libre de enfermedad según el estado hormonal (79% para pacientes sin tratamiento hormona, 94% para hormonoterapia citorreductora y 92% para hormonoterapia adyuvante, $p = 0.046$). Cuando se realizó un análisis según la calidad del implante se mostró que para los pacientes de alto riesgo con valores de $D90 > 140\text{Gy}$ para 125I o bien 100Gy para 103Pd esta diferencia perdía significación, mientras que para los pacientes con dosimetría sub-óptima la significación aumentó.

∴. El grupo del Mount Sinai publicó sus hallazgos acerca del papel de la hormonoterapia en pacientes de riesgo intermedio o alto riesgo tratados con braquiterapia exclusiva (350). Estudiaron 201 pacientes (66 pacientes de riesgo intermedio y 135 pacientes de alto riesgo) de los cuales el 66% recibieron hormonoterapia y 34% no (el 66% de los pacientes de alto riesgo y el 68% de los pacientes de riesgo intermedio recibieron tratamiento hormonal). En el análisis multivariante sobre todos los pacientes la hormonoterapia fue el factor más significativo para predecir la supervivencia libre de recidiva bioquímica a 5 años. También mostraron significación el grupo de riesgo y el valor de PSA inicial. Para los pacientes de riesgo intermedio la hormonoterapia no

resultó ser una variable significativa ($p = 0.6$) para la predicción del control bioquímico (supervivencia libre de recidiva bioquímica a 5 años del 85% vs. 63% para pacientes sometidos o no a hormonoterapia). Sin embargo en los pacientes de alto riesgo la supervivencia libre de recidiva bioquímica fue del 74% en pacientes sometidos a hormonoterapia y del 46% en pacientes sin hormonoterapia ($p = 0.001$). Dentro del grupo general los pacientes con $D90 > 140$ Gy para ^{125}I o bien 100 Gy para ^{103}Pd exhibieron una independencia de la adición de tratamiento hormonal.

∴ Beyer et al. (351) han publicado una amplia experiencia sobre 2378 pacientes tratados con braquiterapia exclusiva (de los que 464 recibieron un ciclo corto de hormonoterapia neoadyuvante). La supervivencia causa específica a 10 años para los dos grupos fue del 89% vs. 81% ($p = 0.133$).

∴ El grupo de Leeds también publicó una larga experiencia sobre 667 pacientes tratados con braquiterapia exclusiva (352). El 51.9% recibieron tratamiento hormonal. La supervivencia libre de recidiva bioquímica fue del 76.1% para los pacientes bajo hormonoterapia y 72.6% para el resto ($p = 0.107$). En el análisis por grupos de riesgo no se detectaron diferencias significativas.

8.4 DE LOS RESULTADOS DE LA TÉCNICA

8.4.1 De la validez del postplanning

La dosimetría post-implante es una herramienta standard para determinar la calidad del implante prostático con semillas permanentes. Tanto la ABS como la AAPM recomiendan la realización de dosimetría post-implante en todos los pacientes que reciban implante permanente (140, 215).

Independientemente de su reconocida utilidad en la evaluación de los implantes permanentes, existen muchas opiniones críticas acerca de la utilidad de la dosimetría post-implante para la evaluación multicéntrica de los resultados de la braquiterapia prostática con semillas permanentes (256). Por una parte, a diferencia de la RTE, la braquiterapia es un procedimiento complejo y dependiente de las habilidades del operador, de su equipamiento y del resto del personal que participa, de forma que la técnica no es homogénea entre instituciones. En segundo lugar, la evaluación post-implante depende de varios factores (entre ellos la modalidad de imagen utilizada y el momento de realización) que también influyen en los resultados. Por todo ello es difícil la creación de un depósito de datos en la bibliografía que permita comparar los nuevos resultados con los publicados previamente, sin la presencia de sesgo por técnica y experiencia del operador, modalidad de imagen o “timing” de obtención del postplanning.

Nosotros somos partidarios de la realización de dosimetría post-implante individualizada en el contexto de un procedimiento validado en los Servicios de Oncología Radioterápica y Física y Protección Radiológica. Existen una serie de requerimientos para su realización que son: mismo intervalo postoperatorio para todos los pacientes, mismo operador para la definición de volúmenes y mismo protocolo de adquisición de imágenes. La realización de dosimetría inmediatamente tras el implante puede ser muy útil, a la luz de los resultados que presentaremos más adelante. Con estos requisitos, la dosimetría post-implante es una herramienta de control de calidad muy útil para la propia institución, y puede servir para construir herramientas predictivas internas a la institución que permitan optimizar los protocolos terapéuticos.

8.4.2 De la modalidad de imagen

El estudio de la influencia de la modalidad de imagen sobre los resultados dosimétricos fue objeto de estudio de un artículo que nuestro grupo publicó en 2004 (242). El objeto de dicho artículo fueron los pacientes de la serie 2 y los resultados mostraron claramente que aplicando dos modalidades de imagen diferentes (TC vs. TC-RNM), tanto los volúmenes prostáticos obtenidos como los parámetros de calidad calculados eran significativamente diferentes.

∴ Los volúmenes prostáticos obtenidos mediante TC (día 0) sobreestimaban los de la ecografía transrectal pre-implante por un factor del 38%. Estos resultados estuvieron en los rangos obtenidos por otros autores (254, 277, 272, 276): Merrick (+25% a día 0 post-implante), Waterman (+53% a día 0 post-implante) o Dogan (30% a día 1 post-implante). Por otra parte, la fusión TC-RNM resultó en una sobreestimación del 8% respecto a los volúmenes obtenidos por ecografía transrectal. Las diferencias entre los volúmenes obtenidos por TC y fusión TC-RNM fueron del 36%, lo que está en consonancia con los resultados publicados (244, 252, 253): Roach (32%), Kagawa (+26%) y Rasch (+43%).

Por otra parte, algunos estudios de imagen han sido validados por confrontación con series quirúrgicas (257, 260, 262, 261, 263). Sosna et al, en una serie de 20 pacientes encontró una fuerte correlación entre los volúmenes calculados mediante RNM y los especímenes patológicos (263), usando software de planimetría ($r^2 = 0.65$) y “rendering” de volúmenes ($r^2 = 0.86$). Sin embargo, Terris et al, (261) en una serie de 150 pacientes estudiados con ecografía transrectal previa a prostatectomía encontró una correlación aún mayor usando la fórmula del elipsoide ($r^2 = 0.90$) y el software de planimetría ($r^2 = 0.93$). Tewari et al, comparando volúmenes obtenidos por ecografía transrectal y RNM encontró resultados muy similares con ambas modalidades y los especímenes quirúrgicos (260). En conclusión, tanto la ecografía transrectal como la RNM son modalidades de imagen que representan adecuadamente el volumen prostático real y son adecuadas y recomendables como herramientas para el cálculo de la dosimetría post-implante.

∴ Se encontró una dependencia entre los valores de los parámetros de control de calidad y la modalidad de imagen empleada. Tanto V100 como D90 fueron infravalorados mediante la dosimetría post-implante basada en TC (tabla 8.4_1).

	TC	Fusión	p
D90 (% de DP)	91.70	114.11	< 0.001*
V100 (% de volumen)	82.48	86.81	0.0143*

Tabla 8.4_1 Valores para D90 y V100 según modalidad de imagen

Otros autores han evaluado la dosimetría post-implante usando fusión TC-RNM. Crook et al. revisan 241 pacientes de bajo riesgo tratados con ^{125}I como tratamiento exclusivo evaluados mediante fusión TC-RNM 1 mes post-implante (353). Los valores medios para V100 y D90 fueron de 92% y 153 Gy respectivamente. En el análisis multivariante, el valor del edema y la adición de hormonas mostraban significación con el valor de V100.

Taussky et al. revisan 20 pacientes de bajo riesgo evaluados mediante fusión TC-RNM (278). Los valores medios para V100 y D90 fueron de 93.6% del volumen prostático y 105% de la dosis de prescripción respectivamente.

Sin embargo también existen resultados que contradicen los anteriores. McLaughlin et al estudian 5 pacientes sometidos a implante permanente y evaluados mediante TC y fusión TC-RNM (354). Sus resultados no muestran una variación significativa en los valores de D90 y V100 aunque la dosis absorbida por los órganos de riesgo (uretra, cuello vesical, esfínter urinario inferior y bulbo peneano) sí que difería en función de la modalidad de imagen aplicada.

Nuestros valores correspondientes a la dosimetría basada en fusión TC-RNM se sitúan en los rangos reportados en la bibliografía estudiada y han de considerarse los más representativos de los valores reales por las consideraciones que hemos hecho anteriormente. Las relaciones dosis-respuesta basadas en estos valores serían más precisas que las obtenidas mediante TC y su utilidad en el conocimiento de la radiobiología del cáncer de próstata, mayor. Lamentablemente la utilización de la fusión de imagen complica el protocolo terapéutico, lo encarece y enlentece notablemente, por lo que su aplicación rutinaria, aunque recomendable, dista mucho de ser factible.

8.4.3 De la importancia del “timing” y del edema

Edema y tiempo transcurrido tras el implante son dos variables relacionadas que hay que estudiar conjuntamente y que conjuntamente afectan a los parámetros dosimétricos y a la supervivencia, según hemos mostrado previamente y discutiremos más adelante.

Desde el inicio del programa de braquiterapia permanente en el IEO nos propusimos estudiar la influencia que tuviera el edema sobre la técnica y sus resultados y por ello establecimos un protocolo con dos dosimetrías post-implante (día 0 y día 50). Asimismo establecimos un estudio de TC seriados sobre un pequeño grupo de pacientes (serie 3), para evaluar en detalle la resolución temporal del edema. Los resultados obtenidos pueden ser agrupados en tres categorías diferentes: impacto del edema sobre los parámetros dosimétricos, predicción de la dosimetría a día 50 a partir de la dosimetría del día 0 y por último, cálculo del periodo de semidesintegración del edema (T_{edema}).

Influencia del edema en los parámetros dosimétricos

Como hemos visto anteriormente, la sobrestimación del volumen en la serie 1 tiene dos componentes:

$$\Delta\text{volumen} = \text{Edema} + \text{Incertidumbre}$$

Debido a que no podemos cuantificar la incertidumbre asociada a la utilización del TC en esta serie hemos de manejar todo el $\Delta\text{volumen}$ (ΔV) como si fuera edema. A nuestro juicio esto no es del todo erróneo: en primer lugar la bibliografía consultada así lo hace. En segundo lugar los resultados obtenidos se refieren únicamente a nuestra serie y no pretendemos extrapolarlos a otros contextos y por último, al estar la variabilidad inter-observador reducida a cero los valores calculados presentan una gran homogeneidad.

∴ En la serie 1 el valor de ΔV_0 fue de 1.46. A día 50 el valor de ΔV fue de 1.16 (mediana 1.10). Estos valores se sitúan en torno a los encontrados en la bibliografía revisada. En la tabla 8.4_2 se presenta un resumen de los mismos.

Autor	PRIMERA TC		SEGUNDA TC	
	Días post-implante	ΔV	Días post-implante	ΔV
Waterman (277)	0	53%	30	10%
Dogan (276)	1	30%	28	10%
Taussky (278)	1	31%	30	5%
Prestidge (273)	0	41%	30	17%
Reed (355)	0	43%	30	21%
Solhjem (270)	0	13%	-	-
Merrick (356)	0	25%	-	-
Lee (255)	-	-	30	31%
Nuestra serie	0	38%	50	16%

Tabla 8.4_2 Valores para el edema según distintos autores

∴ El edema se asoció significativamente al volumen prostático pre-implante, a la actividad implantada y al número de semillas implantadas. Aunque estas variables están claramente relacionadas entre sí, existen dos aspectos llamativos: en primer lugar, el edema es proporcional al volumen prostático pre-implante, lo que contradice algunas observaciones que sugerían que las próstatas de menor volumen presentaban un edema proporcionalmente mayor (278). Por otra parte el número de agujas empleadas en el implante no es predictor de edema aunque sí lo es el número de semillas implantadas. Probablemente se deba a que el factor pronóstico no sea el número de agujas sino el número de punciones que se emplean para depositar las semillas. Este dato no está recogido en nuestra base de datos pero nos parece una explicación plausible que subyugaría por otra parte la importancia de una técnica precisa para obtener una dosimetría adecuada.

∴ La presencia de edema provocó cambios en los valores de los parámetros dosimétricos en función del momento temporal en que se realizó la dosimetría. Los parámetros que presentaron diferencias significativas según el día de evaluación fueron D90, D80, V80, V100 y V150. Otros autores también han encontrado diferencias en este sentido. En la tabla 8.4_3 se presenta un resumen de las mismas y la comparación con nuestros propios valores.

Autor	PRIMER POSTPLAN		SEGUNDO POSTPLAN	
	V100 (%)	D90	V100 (%)	D90
Waterman (277)	88	153 Gy	93	176 Gy
Prestidge (273)	73	-	83	-
Dogan (276)	-	84%	-	93%
Taussky (278)	93	105%	96	110%
Reed (355)	95	169 Gy	95	175 Gy
Nuestra serie	85	136 Gy (93%)	90	148 Gy (102%)

Tabla 8.4_3 Diferencias en los parámetros dosimétricos según el “timing”

∴ Hemos mostrado que el momento en el que se obtiene la dosimetría para la evaluación de la calidad de los implantes permanentes influye en los resultados obtenidos, y que esa diferencia es estadísticamente significativa (figura 8.2). A nuestro juicio la presentación de los resultados dosimétricos debería ir acompañada de la reseña del día en que se obtienen las imágenes correspondientes y del valor de ΔV correspondiente. De esta forma la comparación multicéntrica de los resultados podría incorporar estas dos variables en el análisis. El momento óptimo para la obtención de la dosimetría post-implante sería probablemente en torno al día 30 post-implante, cuando la mayor parte del edema se ha eliminado, como han señalado diferentes autores anteriormente (281).

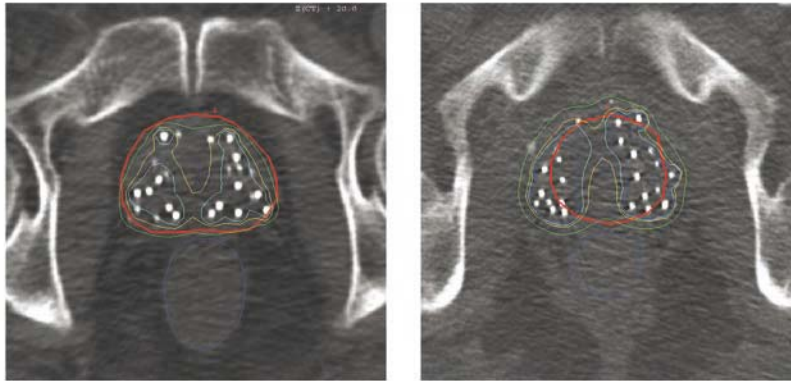


Figura 8.2 Aspecto post-implante utilizando imágenes de TC. A la izquierda se ve una imagen tomada inmediatamente post-implante. A la derecha, un corte del mismo paciente, aproximadamente a la misma altura, 50 días post-implante.

Predicción de la dosimetría de día 50 a partir de la de día 0

Diversos autores han explorado la posibilidad y la utilidad de conocer los parámetros dosimétricos en cualquier momento temporal a partir de los datos de la dosimetría del día 0 (357, 277).

En la serie 1 la variación de ΔV se acompañó de un incremento de V100 de un 4.17% entre el día 0 y el día 50. El valor de D90 se incremento en 12.94 Gy (9.51%) entre los días 0 y 50. A partir de todos estos valores se construyó un modelo de regresión lineal multivariante con la intención de predecir los valores del día 50 a partir de los del día 0 y el valor del edema post-implante. Este análisis generó la siguientes expresiones estadísticamente significativas para D90 ($R^2 = 0.29$) y V100 ($R^2 = 0.23$):

$$D90_{50} = 24.74 + 0.67 * D90_0 + 27.35 * Edema + 0.60 * VOLPROST_PRE$$

$$V100_{50} = 40.39 + 0.48 * V100_0 + 6.62 * Edema + 0.14 * VOLPROST_PRE$$

Waterman et al. (277) establecen su propio modelo, con significación estadística y valores de R^2 mas fuertes, tanto para D90 ($R^2 = 0.69$) como para V100 ($R^2 = 0.68$). Sus ecuaciones son las siguientes:

$$D90_{30} = -87.41 + 1.10 * D90_0 + 62.56 * Edema$$

$$V100_{30} = 11.70 + 0.72 * V100_0 + 11.10 * Edema$$

En nuestro caso, los valores predichos por el modelo y los valores reales no presentaban diferencias estadísticamente significativas. Sin embargo, las predicciones presentaban una desviación respecto a los valores reales de ± 19 Gy para D90 y de $\pm 5.78\%$ para V100 (± 24 Gy y $\pm 4\%$ para el modelo de Waterman). Respecto a los valores de R^2 más débiles en nuestro modelo se explican por el método de planificación empleado: el grupo de Waterman definió el CTV como el volumen prostático delimitado por ecografía transrectal pre-implante añadiendo un margen variable de 3-5mm. De este modo el análisis post-implante del ETV (“evaluation target volume”) es menos sensible a pequeñas diferencias de contorno (± 1.5 mm), como queda demostrado en su artículo. En nuestro caso, el CTV se definió como próstata sin mérgenes, por lo que los valores sobre el ETV post-implante resultaron mucho más sensibles a las pequeñas

variaciones de delineación. En cualquier caso, aunque no hemos podido cuantificarlas por problemas de método, estas diferencias no han de ser muy grandes debido a la baja variabilidad intra-observador esperada.

Para finalizar, hacemos nuestra la conclusión de Frank Waterman y no creemos que el modelo de regresión lineal múltiple propuesto (ni en su caso ni en el nuestro) sea lo suficientemente preciso como para servir de sustituto a la dosimetría diferida a 30 ó 50 días.

Periodo de semidesintegración del edema y cálculo de D100_corr

En el epígrafe 7.3 se ha realizado un completo estudio de cálculo de T_{edema} y validación de la variación de la dosimetría en función de sus valores. Del análisis de la serie 3 hemos obtenido un valor para T_{edema} de 14.84 días (que corresponde a un paciente promedio de los cinco pacientes estudiados). El valor más bajo fue de 7.9 días y el valor mayor de 19.8 días. Estos valores se sitúan en los rangos previamente publicados por otros autores (tabla 8.4_4)

Autor	Valor de T_{edema} (días)	Rango (días)
Waterman (272)	9.3	4 – 25
Merrick (275)	10.9	8.6 – 14.3
Taussky (278)	16 (mediana 14.7)	6.4 – 21.5 (rango IQ)
Nuestra serie	14.84	7.9 – 19.8

Tabla 8.4_4 Valor para el periodo del edema

Antes de utilizar el valor obtenido en ulteriores cálculos dosimétricos sobre la serie 1 se procedió a la validación de los valores obtenidos sobre una predicción teórica⁽²⁸³⁾. El valor medio del error teórico fue del 11.69% y para los datos experimentales del 8.8%. La diferencia de las medias no fue estadísticamente significativa al aplicar el test U de Mann-Whitney ($p = 0.4527$). De este modo se aceptó el valor de 14.84 días como periodo de semidesintegración (semi-eliminación) del edema.

Una vez que se obtuvo el valor de T_{edema} se procedió al cálculo de la dosis mínima periférica corregida según la magnitud del edema del primer día. Dicho valor de dosis se denominó D100_corr. Para ello se integraron los datos de la base de datos con el paquete de cálculo Wolfram Matemática 5.2 ®. La nueva variable así obtenida se devolvió a la base de datos y se maneja de forma convencional.

Para hacernos una idea del significado de esta variable imaginemos un implante que no presenta edema ($\Delta V = 0$) y que recibe como dosis mínima periférica 145 Gy sobre una TC obtenida el día 0. En este caso D100_0 correspondería a la dosis de prescripción (DP). Por tanto,

$$D100_0 = DP = 145 \text{ Gy}$$

En este caso, la dosis absorbida por la totalidad de la próstata a tiempo t post-implante será:

$$D = d_0 * \int_0^t e^{(-\lambda * t)} dt$$

donde d_0 es la tasa de dosis inicial que viene dada por la expresión:

$$d_0 = \lambda * D0$$

donde λ es la constante de semidesintegración del ^{125}I :

$$\lambda = \text{Ln}(2) / T_{\text{edema}}$$

Y donde T_{edema} es el periodo de semidesintegración del ^{125}I .

En la tabla 8.4_5 vemos el comportamiento de este implante (primera columna de la tabla) y lo podemos comparar con otros implantes en los que D0 es igualmente 145 Gy pero en los que se ha inducido un edema traumático por el implante.

Tiempo post-implante	Dosis (Gy)			
	Edema = 1	Edema = 1.1	Edema = 1.3	Edema = 1.5
+ 1d	1.6654	1.2948	0.8320	0.5690
+ 2d	3.3118	2.5890	1.6779	1.1549
+ 3d	4.9392	3.8818	2.5370	1.7573
+ 4d	6.5479	5.1727	3.4087	2.3761
+ 5d	8.1382	6.4612	4.2923	3.0107
+ 6d	9.7102	7.7467	5.1871	3.6610
+ 7d	11.2642	9.0288	6.0925	4.3263
+ 14d	21.6533	17.8698	12.6703	9.3683
+ 21d	31.2353	26.3909	19.5083	14.9498
+ 28d	40.0730	34.5072	26.4078	20.8707
+ 60d	72.5000	65.6901	55.3201	47.7640
+ 180d	126.8750	119.8310	109.0020	100.9980
+ 365d	142.8620	135.8180	124.9880	116.984

Tabla 8.4_5 Comportamiento de D100_corr según edema e intervalo post-implante

El efecto queda reflejado gráficamente en la figura 7.17. Se puede observar que para tiempos post-implante cortos la diferencia proporcional en el valor de D100_corr aumenta. Nosotros hemos introducido como tiempo de cálculo 7 días, que viene a ser la mitad del tiempo de semidesintegración del edema, por lo cual se espera que D100_corr se vea impactado de forma más severa y se traduzca con fuerza el efecto del edema. D100_corr se convierte de esta forma en un parámetro muy poderoso por que combina de forma no-lineal tres variables, reduciendo la dimensionalidad de la base de datos y los problemas de colinearidad en los modelos multivariantes.

8.4.4 Impacto del edema y de las variables dosimétricas sobre la SLE

Sobre el grupo general

Del análisis univariante se desprenden como variables asociadas significativamente a la probabilidad de recidiva bioquímica las siguientes: actividad implantada, número de agujas, D100_0, edema normalizado, D100_corr. De ellas sólo D100_corr junto al valor de PSA inicial y el grupo de riesgo conserva significación en el análisis multivariante.

∴ La dosis mínima periférica (D100) es un parámetro poco usado en las series que hemos revisado. Dado que representa la dosis en la periferia de la próstata es muy sensible a pequeñas variaciones en su definición, al edema y al posicionamiento de las semillas. Por ello D100 suele ser significativamente menor que la dosis de

prescripción, aunque el volumen de infra-dosificación no suele exceder el 10-15% del volumen blanco. Por todo ello su utilidad como parámetro para realizar comparaciones multicéntricas es dudosa (221, 358) y la AAPM TG64 recomienda el uso de D90 en vez de D100 (215). En nuestro caso la ausencia de variabilidad inter-observador y la presumida homogeneidad intra-observador, unidas al método de planificación, sin utilizar márgenes han permitido que D100 adquiriera protagonismo. En nuestro caso ha sido uno de los parámetros más interesantes como veremos en lo sucesivo. Por un lado se ha mostrado predictor de recidiva en la serie general. Por otro lado, su conversión en D100_corr se ha mostrado con un parámetro sólido que ha resistido la influencia de otras variables en el análisis multivariante. Y por último nos ha servido para construir la curva de dosis-respuesta para nuestra serie.

En nuestro caso, el valor de corte para D100_0 fue de 78.4 Gy, que corresponde a un 54.06% de la dosis de prescripción. No hemos encontrado otras referencias en la bibliografía revisada que hagan referencia a D100 como parámetro relacionado con la supervivencia libre de recidiva bioquímica aunque hemos encontrado referencia a valores medios en la serie de Merrick et al. con 100 Gy a día 0 (356) y Stock et al. con 102 Gy a día 30 (358). Recordemos que el valor medio de nuestra serie fue de 69 Gy a día 0 y 72 Gy a día 50.

∴ El valor de D100 corregido por la presencia del edema (D100_corr) se presenta como el valor más importante del análisis multivariante entre toda la panoplia de parámetros técnico-dosimétricos. Recordemos que se ha introducido como valor de T_{edema} 14.84 días y que se ha calculado la dosis a 7 días de implante (que resulta ser la mitad de T_{edema}). Ello nos ha dado como valor medio para toda la serie 2.70 Gy. El valor de corte significativo para supervivencia libre de recidiva bioquímica fue de 3 Gy. De la revisión de la bibliografía no hemos obtenido referencias a este parámetro ni a ningún otro similar.

∴ El valor del edema no demostró asociación con la supervivencia libre de enfermedad bioquímica. Sin embargo, su valor normalizado sí mostró relación con la recidiva bioquímica. El valor de corte fue de 0.75. No hemos encontrado referencias al impacto del edema sobre la supervivencia libre de enfermedad bioquímica. Sin embargo tanto el edema como D100_corr se han mostrado relacionados con el control de la enfermedad, lo que puede interpretarse como una relación con la calidad de la técnica de implante.

∴ Se ha encontrado asociación estadística entre la actividad implantada y la supervivencia libre de recidiva bioquímica, por una parte, y entre el número de agujas y la supervivencia libre de recidiva bioquímica por otro. Ninguna de las dos variables alcanzó significación estadística en el análisis multivariante.

∴ No hemos encontrado valores de corte significativos para D90 ni V100 (ni para día 0 ni para día 50) para la serie general. Estos resultados concuerdan con la experiencia publicada por Dan Ash et al. sobre 667 pacientes (46.2% bajo riesgo, 35.7% riesgo intermedio y 18.1% alto riesgo) tratados con ^{125}I como monoterapia (359). Los autores no encontraron diferencias significativas para ningún punto de corte analizando todo el grupo en su conjunto (figura 8.3). Merrick et al. (348) estudian 425 pacientes con dosimetría a día 0 y tampoco encuentran valores de corte significativos para los parámetros de control de calidad derivados del HDV.

Relación entre D90 y bNED (I)

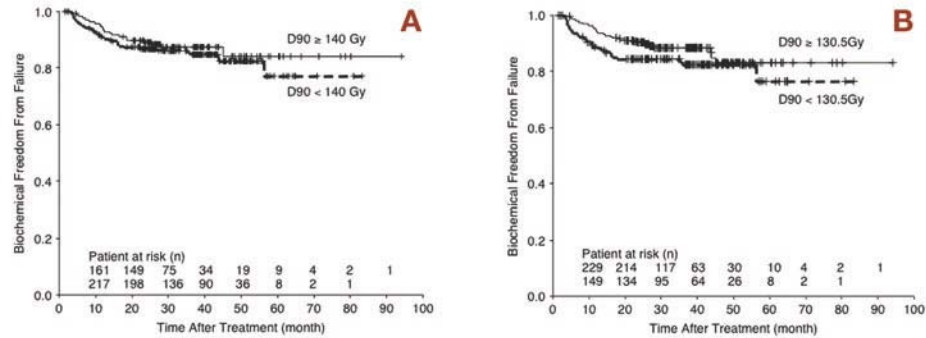


Figura 8.3 Relación entre D90 y supervivencia libre de recidiva bioquímica (bNED). En el panel A se observa la supervivencia libre de recidiva bioquímica para un punto de corte de 140 Gy. Panel B: punto de corte para D90 de 130.5Gy. En ninguno de los dos casos las diferencias adquirieron significación. Tomado de Ash D, et al. The correlation between D90 and outcome for I-125 seed implant monotherapy for localised prostate cancer. Radiother. Oncol. 2006; 79: 185-9

Relación entre D90 y bNED (II)

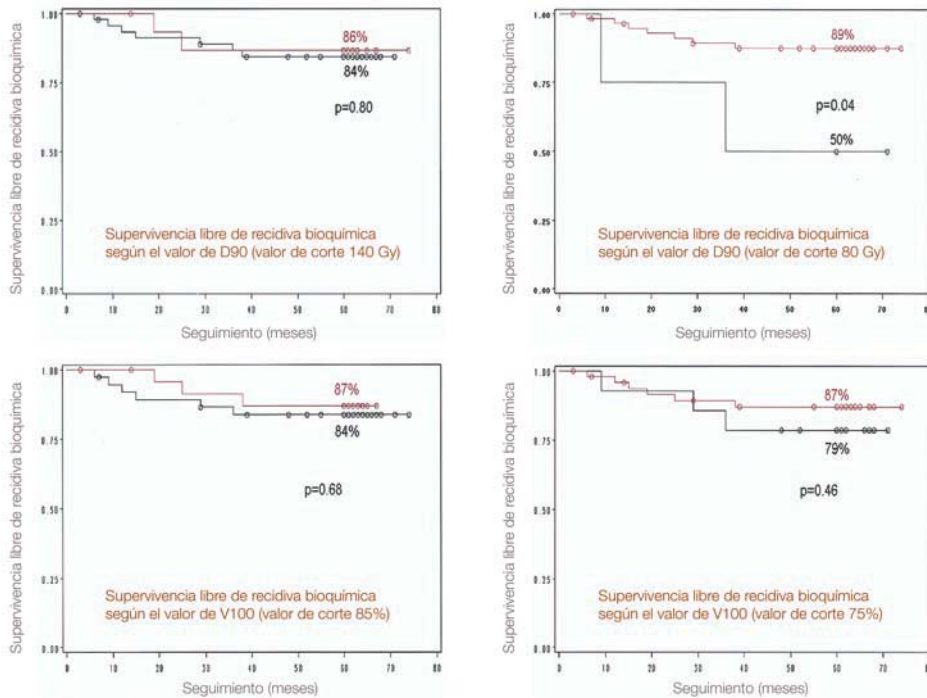


Figura 8.4 Relación entre D90 y V100 y la supervivencia libre de recidiva bioquímica. Tomado de Lee R, et al. Dosimetry and cancer control after low-dose-rate prostate brachytherapy. Int. J. Radiat. Oncol. Biol. Phys 2005; 61: 52-59

Otros autores encuentran valores significativos sólo para dosimetrías muy pobres. Lee et al. publicaron su experiencia sobre 63 pacientes tratados con braquiterapia exclusiva con ^{125}I (360). El seguimiento mediano fue de 5 años y se identificaron puntos de corte con influencia significativa para la supervivencia libre de recidiva bioquímica. Para D90, 80 Gy y para V100, 65% (figura 8.4).

Sin embargo nuestros resultados están en contradicción con algunos resultados de la bibliografía revisada. Papagikos et al. revisan una serie de características similares a la nuestra (361). Un total de 103 pacientes (59% de bajo riesgo y 41% de riesgo intermedio) tratados con braquiterapia exclusiva. El seguimiento medio de la serie fue de 5 años. En su experiencia, tanto D90 como V100 calculados 1 mes post-implante presentaron relación con el control de la enfermedad, con valores de corte de 100 Gy para D90 y 80% para V100 (figura 8.5).

Los investigadores del Hospital Mount Sinai de New York han publicado de forma exhaustiva su experiencia y la relación entre D90 y la supervivencia libre de enfermedad (figura 8.6). En un primer estudio sobre 134 pacientes tratados con braquiterapia exclusiva con ^{125}I y una mediana de seguimiento de 32 meses (rango entre 12 y 72 meses) se describe un punto de corte con significación estadística de 140 Gy (286). La supervivencia libre de recidiva bioquímica a 4 años para el grupo de D90 < 140 Gy fue de 68% mientras que para el grupo de D90 \geq 140 Gy fue del 92% ($p = 0.02^*$). Kollmeier et al. (154) en un estudio sobre 243 pacientes tratados con braquiterapia exclusiva (^{125}I o ^{103}Pd) con un seguimiento mínimo de 5 años (rango entre 61 y 135 meses, media 75 meses) y evaluados a 30 días post-implante mediante dosimetría TC demuestran que la supervivencia libre de recidiva bioquímica a 8 años fue del 82% para pacientes con D90 \geq 140 Gy y del 68% para pacientes con D90 < 140 Gy ($p = 0.007^*$). En el análisis multivariante mostraron significación estadística el PSA inicial ($p < 0.0001^*$), GS ($p = 0.024^*$) y D90 ($p = 0.046^*$). Para el subgrupo de pacientes de bajo riesgo, la supervivencia libre de recidiva bioquímica fue del 94% en los pacientes con D90 \geq 140 Gy y del 75% en pacientes con D90 < 140 Gy ($p = 0.02^*$).

En una revisión multicéntrica sobre 719 pacientes tratados mediante braquiterapia (^{125}I o ^{103}Pd) con o sin adición de RTE, Potters et al. (287) encuentran significación estadística para D90 (valor de corte, 90% de la dosis de prescripción) en las siguientes circunstancias: sobre el grupo general, sobre los subgrupos de ^{103}Pd , ^{125}I , monoterapia, subgrupo tratado con HT y subgrupo sin tratamiento hormonal. En la figura 8.7 se muestra el análisis para el grupo general y para el subgrupo de pacientes tratados con ^{125}I . No alcanzó significación estadística para el subgrupo de pacientes tratados mediante terapia combinada. En la tabla 8.4_6 se muestran los valores de supervivencia libre de recidiva bioquímica a 4 años.

Grupo de pacientes	D90 < 90%	D90 \geq 90%	p
Grupo general	80.4%	92.4%	0.001*
^{103}Pd	83.4%	93.3%	0.01*
^{125}I	63.8%	93%	0.04*
No hormonas	81.0%	93.4%	0.001*
Hormonas	79.3%	92.5%	0.001*
Monoterapia	74.3%	94.6%	0.001*

Tabla 8.4_6 Resultados según valor de corte para D90 (Potters et al.)

Relación entre D90 y bNED (III)

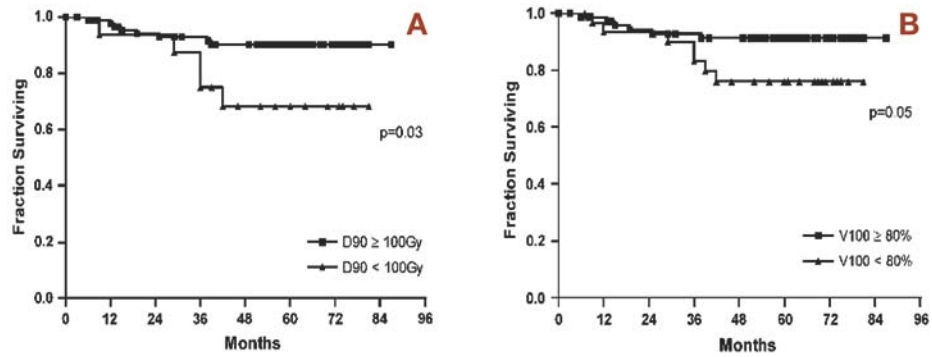


Figura 8.5 Relación entre D90 y supervivencia libre de recidiva bioquímica (bNED). En el panel A se observa la supervivencia libre de recidiva bioquímica según D90. Panel B: supervivencia libre de recidiva bioquímica según V100. Tomado de Papagikos MA, et al. Dosimetric quantifiers for low-dose-rate prostate brachytherapy: is V100 superior to D90? Brachytherapy 2005; 4: 252-258

Relación entre D90 y bNED (IV)

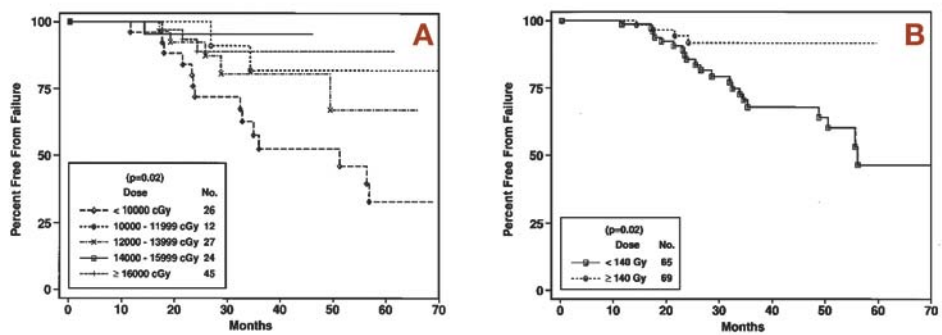
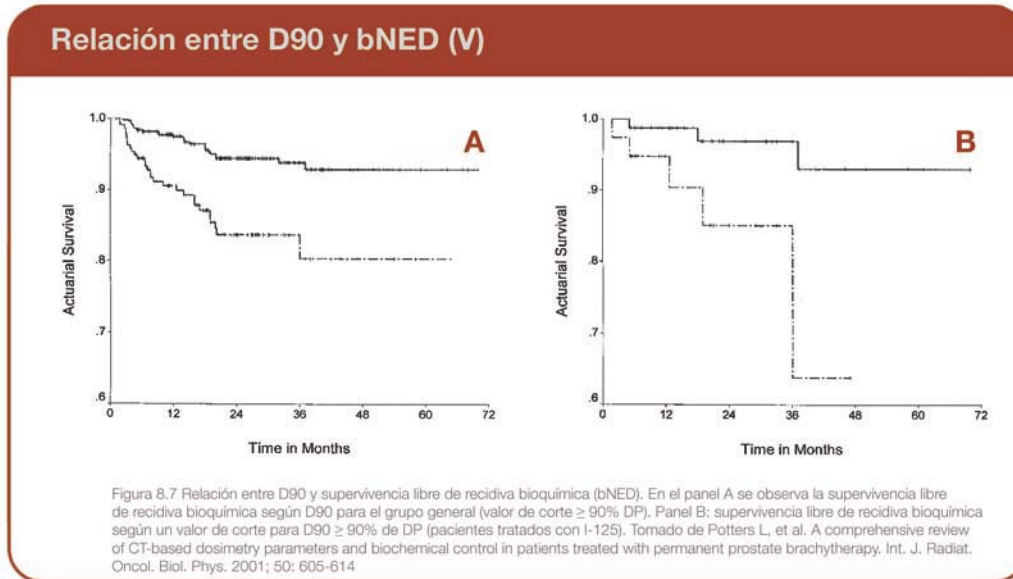


Figura 8.6 Relación entre D90 y supervivencia libre de recidiva bioquímica (bNED). En el panel A se observa la supervivencia libre de recidiva bioquímica según cinco segmentos de dosis para D90. Panel B: supervivencia libre de recidiva bioquímica según un valor de corte para D90 de 145 Gy. Tomado de Stock R, et al. A dose-response study for I-125 prostate implants. Int. J. Radiat. Oncol. Biol. Phys. 1998; 41: 101-108



Sobre el grupo de riesgo intermedio

De las variables dosimétricas mostraron significación estadística en el análisis univariante las siguientes: D100_0, D90_0, V100_0, edema, edema normalizado, D100_corr. Las siguientes variables mostraron significación en el test de Wilcoxon pero no en el de Log-Rank, y no se consideraron significativas: actividad implantada y número de agujas empleadas. En el análisis multivariante sólo D90_0 mostró significación (junto con el PSA inicial, como hemos visto anteriormente).

∴ La dosis mínima periférica (D100_0) también ha tenido significación sobre la supervivencia libre de recidiva bioquímica en el grupo de pacientes de riesgo intermedio. El valor de corte empleado para el grupo general (78.4 Gy, que corresponde a un 54.06% de la dosis de prescripción) también fue significativo para el grupo de riesgo intermedio.

∴ En el grupo de riesgo intermedio se encontró asociación estadísticamente significativa entre la supervivencia libre de recidiva bioquímica y D90_0. En el estudio publicado por Ash et al. citado previamente no se han encontrado diferencias significativas para el valor de D90 en el grupo de riesgo intermedio (359). En el resto de estudios mencionados anteriormente no se hace referencia explícita a los resultados para este grupo de riesgo.

∴ El valor de V100 (punto de corte, 90%) también se asoció significativamente con la supervivencia libre de recidiva bioquímica. No existen resultados para la asociación entre V100 y el grupo de riesgo intermedio. Los valores de corte referidos anteriormente para el grupo general eran: 90% (Potters), 80% (Papagikos) y 65% (Lee).

∴ Los valores asociados con el edema han mostrado gran significación estadística con la supervivencia libre de recidiva bioquímica en el grupo de riesgo intermedio. Sin embargo los problemas de correlación entre ellas han restado validez al modelo multivariante. Por otra parte, el modelo reducido eliminando edema y edema normalizado y dejando D100_0 tampoco ha capturado significación para esta última.

La fuerte asociación de estas variables con la recidiva bioquímica se puede explicar de la siguiente forma: como hemos visto anteriormente, el edema provoca una migración centripeta de las fuentes, lo que origina una disminución de la tasa de dosis en la periferia de la próstata (fenómeno que se ha capturado con el cálculo de D100_corr) y una pérdida de eficacia de la radiación sobre los clonógenos. Esta disminución de la eficacia biológica de la radiación se prolongará en función de la duración del edema (de su constante específica). Al final del tiempo efectivo de irradiación la dosis física administrada habrá sido de 145 Gy, pero la topografía de dicha irradiación y su tasa de dosis inicial serán decisivas en el control efectivo de la enfermedad.

No hemos encontrado bibliografía que soporte la discusión que hemos establecido sobre el valor de D100_corr, de modo que ha de considerarse una hipótesis de trabajo para estudios posteriores, o para un análisis matemático más preciso de los datos de este subgrupo de pacientes.

Sobre el grupo de bajo riesgo

En el grupo de bajo riesgo se han producido 5 eventos durante el tiempo de seguimiento de la serie. Es llamativo por tanto que se hayan podido establecer variables con significación estadística. Sin embargo, nuevamente los parámetros relacionados con el edema han mostrado capacidad predictiva. Sin embargo, como hemos dicho anteriormente, los problemas de correlación entre ellas han restado validez al modelo multivariante, sin que el modelo reducido haya ganado significación estadística.

Como hemos visto, los valores de corte que dan significación para las variables son ligeramente distintos a los vistos en el grupo de riesgo intermedio, pero pueden interpretarse en el mismo sentido hecho anteriormente. Valores de edema mucho mayores (> 1.96) afectarían la supervivencia libre de recidiva bioquímica en el grupo de bajo riesgo. Probablemente la disminución de la tasa de dosis periférica que hemos citado anteriormente no tenga tanta importancia dado que el número de clonógenos es menor o su radiosensibilidad mayor. Por ello valores más bajos de D100_corr han conseguido el control de la enfermedad en este grupo de riesgo.

La hipótesis expresada para el valor del edema y D100_corr ha de ser verificada ulteriormente mediante estudios ad hoc o mediante el análisis sistemático del desplazamiento relativo de las semillas a lo largo del tiempo. En cualquier caso, no hemos encontrado referencias a resultados similares en la búsqueda bibliográfica realizada, por lo que consideramos interesante realizar esfuerzos para intentar comprender la interrelación de estos parámetros con el control local de la enfermedad.

8.4.5 De la curva de dosis-respuesta

Bajo el epígrafe 4.1.3 se ha comentado el paradigma del escalado de dosis en la radioterapia del cáncer de próstata y hemos repasado alguno de los hitos en su desarrollo. El beneficio del escalado de dosis se sugirió a partir de los estudios de biopsia post-radioterapia y se confirmó a partir de series clínicas. Por una parte se pudo comprobar que el porcentaje de biopsias positivas después de irradiación radical era mayor cuanto más avanzado es el estadio clínico, sugiriendo la necesidad de aumentar la dosis por encima de 70 Gy en estos estadios. Por otra parte se comenzaron a publicar los resultados de control bioquímico tras radioterapia externa que demostraron el beneficio bioquímico del escalado de dosis. A partir de los datos clínicos de escalado de dosis, algunos autores han construido sus propias curvas de dosis respuesta para los pacientes tratados con radioterapia externa.

∴ Cheung et al. (362, 312) describen la curva de dosis respuesta para la serie de pacientes del MD Anderson Cancer Centre tratados con radioterapia externa (figura 8.8). Los autores estudian 235 pacientes de bajo riesgo, 387 pacientes de riesgo intermedio y 363 pacientes de alto riesgo. Los criterios para la clasificación en los grupos de riesgo fueron los siguientes:

Bajo riesgo: [estadio clínico \leq T2a] & [PSA \leq 10ng/ml] & [GS \leq 6]

Riesgo intermedio: [estadio clínico T2b-T2c] & [10 < PSA \leq 20 ng/ml] & [GS 7]

Alto riesgo: [estadio clínico \geq T3] o [PSA > 20 ng/ml] o [GS \geq 8]

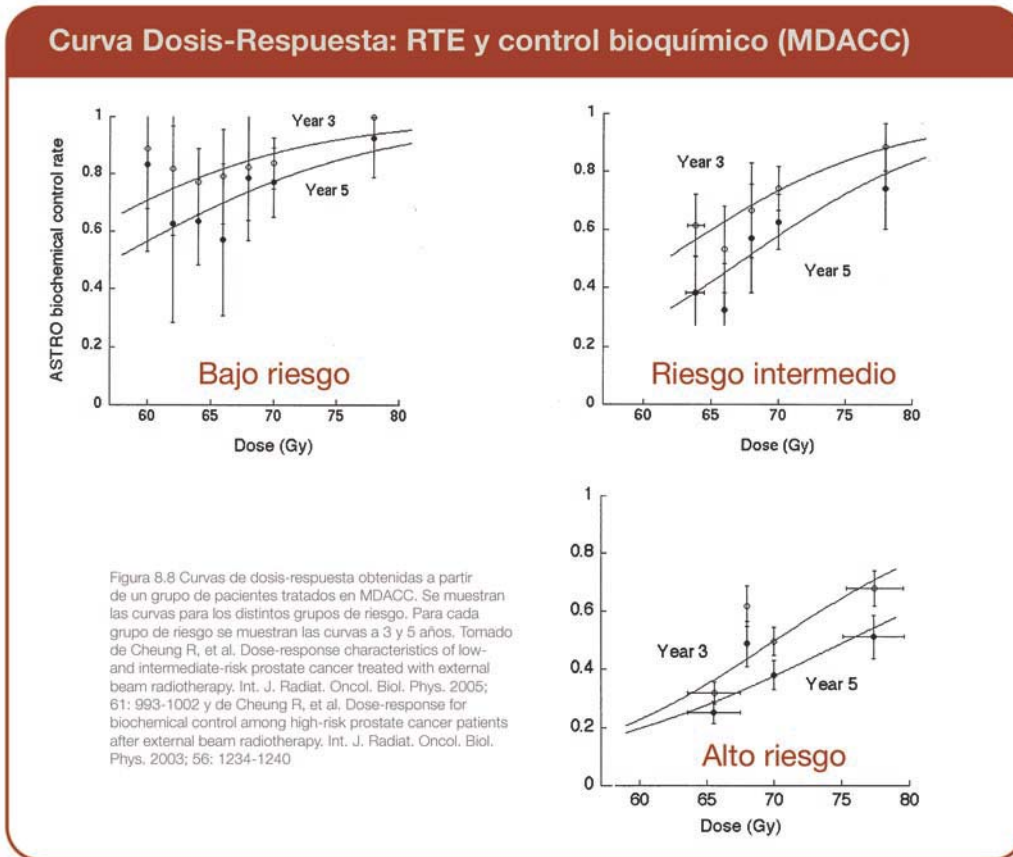
El criterio para evaluación del fracaso terapéutico fue el definido por el consenso ASTRO, aunque se usaron definiciones alternativas a efectos de comparación.

Las curvas se obtuvieron a partir de segmentos de dosis definidos por la evolución histórica de la dosis de prescripción en el MDACC, que varió entre 66 y 78 Gy entre 1987 y 1998. A partir de estos puntos se realizó un ajuste usando un modelo logístico de 4 parámetros. Posteriormente se realizó un estudio usando el algoritmo de RPA para identificar subgrupos donde la relación dosis-respuesta se hiciera significativa. Para los grupos de bajo riesgo y riesgo intermedio, el algoritmo RPA identificó dos subgrupos en función del valor de PSA (valor de corte, 7.5 ng/ml) para los que se hacía evidente el beneficio del escalado de dosis. Para el grupo de alto riesgo se identificarán dos subgrupos según el valor de PSA (valor de corte, 15 ng/ml) donde se hizo evidente el efecto dosis-respuesta (diferencia de 15 Gy para el valor de TCD50). En la tabla 8.4_7 pueden verse los parámetros de dicha curva según los grupos de riesgo.

Grupo de riesgo	TCD50 (Gy) CI95%	σ 50	γ 50 CI95%
Bajo riesgo	57.3 (47.6 – 67.0)	-	1.4 ([-0.1] – 2.9)
Riesgo intermedio	67.5 (65.5 – 69.5)	-	2.2 (1.1 – 3.2)
Alto riesgo	75.5 (70.7 – 80.2)	-	1.7 (0.7 – 2.7)

Tabla 8.4_7 Parámetros para la curva de dosis-respuesta del grupo del MDACC. El valor de σ 50 no se pudo calcular debido a que se hizo una reparametrización de la que no se dieron detalles en el artículo.

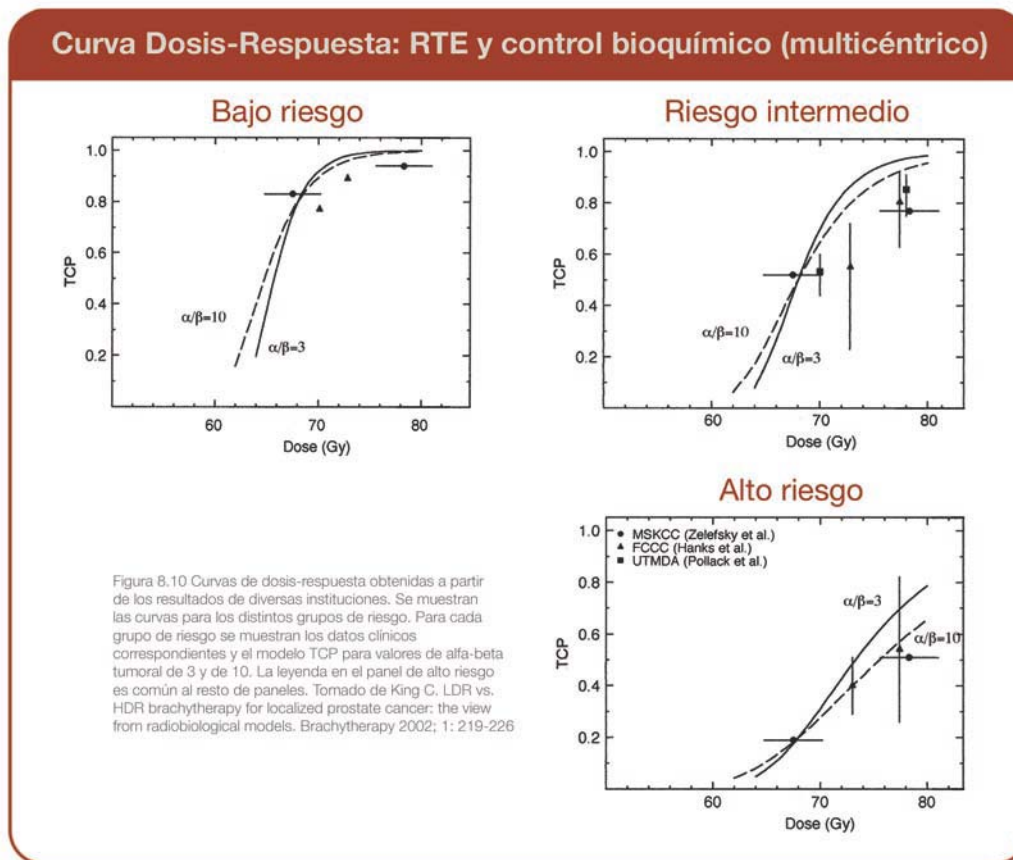
∴ Levegrün et al. (363) analizan las biopsias prostáticas post-radioterapia en una serie de pacientes pertenecientes al MSKCC y realizan un ajuste logístico y un modelado radiobiológico para establecer la relación dosis-respuesta de la serie. Se estudiaron 103 pacientes tratados con radioterapia externa, para los que se dispuso de biopsia prostática a un intervalo de tiempo mayor de 30 meses post-radioterapia. La biopsias positivas se consideraron fracasos terapéuticos mientras que las biopsias negativas o dudosas se consideraron éxito terapéutico. Se obtuvo una curva de dosis respuesta que se muestra en la figura 8.9. En la tabla 8.4_8 se muestran los valores de ajuste de los datos observados a un modelo logístico y a distintos modelos radiobiológicos



Grupo de riesgo	TCD50 (Gy)	σ 50	γ 50
Bajo riesgo	64.05 ± 0.30	-0.1713 ± 0.0090	-
Riesgo intermedio	69.03 ± 0.28	-0.3207 ± 0.0280	-
Alto riesgo	74.96 ± 1.17	-0.3251 ± 0.1148	-
Modelo radiobiológico_1			
Bajo riesgo	64.99	-	2.93
Riesgo intermedio	67.75	-	3.60
Alto riesgo	75.70	-	3.33
Modelo radiobiológico_2			
Bajo riesgo	66.41	-	3.52
Riesgo intermedio	70.38	-	6.28
Alto riesgo	76.16	-	4.23

Tabla 8.4_8 Parámetros para la curva de dosis-respuesta del grupo del MSKCC. El valor de γ 50 no se pudo calcular debido a que se hizo una reparametrización de la que no se dieron detalles en el artículo.

∴ También existe experiencia en el cálculo de curvas de dosis-respuesta en pacientes sometidos a braquiterapia. King et al. (364) realiza un ejercicio de construcción de un modelo radiobiológico específico para la braquiterapia (LDR y HDR) teniendo en cuenta el valor de los parámetros radiobiológicos convencionales (α/β , λ_{isotopo} , tasa de dosis inicial y constante de reparación del daño subletal) posteriormente revisa datos procedentes de varias series de braquiterapia prostática, para comparar los distintos ajustes obtenidos y elaborar una curva de dosis-respuesta por grupos de riesgo. En el artículo no se ofrecen datos respecto a TCD50 ni σ_{50} y no se pueden calcular de forma inversa. La curva de dosis-respuesta resultante se muestra en la figura 8.10.



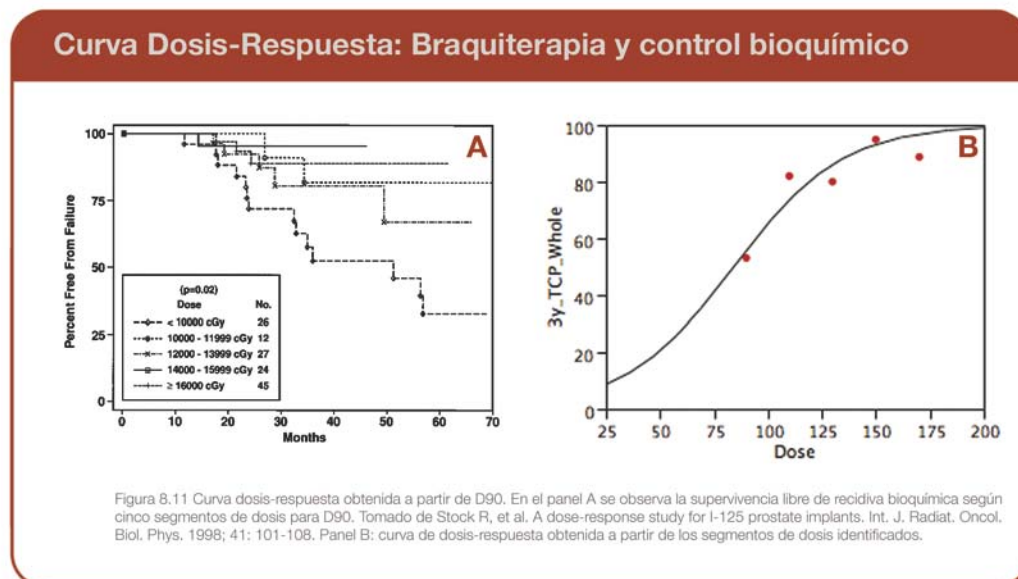
∴ Stock et al. (286) ofrecen una serie excelente sobre 134 pacientes tratados con braquiterapia exclusiva (^{125}I) y evaluados mediante TC un mes post-implante. El fracaso terapéutico se definió como dos aumentos consecutivos del PSA tras el nadir, o bien un valor para el nadir por encima de 1 ng/ml. Se obtuvo una segmentación de la base de datos en cinco grupos de valores para D90:

- Grupo 1: $D_{90} < 100$ Gy
- Grupo 2: $100 < D_{90} < 119.9$ Gy
- Grupo 3: $120 < D_{90} < 139.9$ Gy
- Grupo 4: $140 < D_{90} < 159.9$ Gy
- Grupo 5: $D_{90} \geq 160$ Gy

La supervivencia libre de recaída bioquímica para cada uno de los grupos a 4 años fue del 53, 82, 80, 95 y 89%. La diferencia entre las curvas fue estadísticamente significativa según el test de Log-Rank ($p = 0.02$). Dividiendo el grupo en dos segmentos (valor de corte 140 Gy) se obtuvo una fuerte relación dosis-respuesta. Los pacientes que recibieron $D90 < 140$ Gy tuvieron una supervivencia libre de recidiva bioquímica del 68% comparado con el 92% en los pacientes que recibieron $D90 \geq 140$ Gy. La diferencia fue estadísticamente significativa ($p = 0.02$).

Se realizó un análisis por nivel de PSA. En los pacientes con $PSA > 10$ ng/ml el efecto dosis-respuesta fue mas pronunciado. Los pacientes que recibieron $D90 < 140$ Gy tuvieron una supervivencia libre de recidiva bioquímica a 4 años del 51% comparado con el 100% en pacientes con $D90 \geq 140$ Gy ($p = 0.009$). Para los pacientes con $PSA \leq 10$ ng/ml la supervivencia libre de recidiva bioquímica a 4 años del 82% para pacientes con $D90 < 140$ Gy comparado con el 88% en pacientes con $D90 \geq 140$ Gy ($p = 0.6$).

El ajuste del modelo logístico sobre los 5 grupos de dosis obtenidos a partir de los datos publicados en dicho artículo nos ofrece los siguientes valores para los parámetros del modelo: $TCD50 = 84.19$ Gy (± 6.99), $\sigma_{50} = [-0.0397]$ (± 0.0100). Se representa gráficamente en la figura 8.11.



∴ A partir de los datos existentes en nuestra base de datos se pudo construir una curva que relacionara la dosis administrada mediante braquiterapia con los resultados de control bioquímico de la enfermedad. En la tabla 8.4_9 se muestra el resumen del ajuste del modelo logístico sobre los grupos de dosis generados a partir de los valores de D100_0, D100_corr y D90_50.

Grupo de riesgo	TCD50 (Gy)	σ 50 (Gy-1)	γ 50
D100_0			
Grupo general	40.06 (\pm 4.64)	-0.0557 (\pm 0.0211)	2.07
Bajo riesgo	36.31 (\pm 3.21)	-0.0920 (\pm 0.0239)	3.34
Riesgo intermedio	69.73 (\pm 2.38)	-0.1044 (\pm 0.0309)	7.27
D100_corr			
Grupo general	1.19 (\pm 0.2538)	-1.6125 (\pm 0.4760)	1.91
Bajo riesgo	1.37 (\pm 0.1271)	-3.0884 (\pm 0.5420)	4.23
Riesgo intermedio	2.20 (\pm 0.2018)	-1.4983 (\pm 0.3466)	3.29
D90_50			
Grupo general	80.60 (\pm 27.29)	-0.0312 (\pm 0.0166)	2.51
Bajo riesgo	106.92 (\pm 40.00)	-0.0266 (\pm 0.0229)	2.84
Riesgo intermedio	139.62 (\pm 2.50)	-0.1191 (\pm 0.0128)	16.62

Tabla 8.4_9 Parámetros del ajuste logístico para nuestra serie.

Del método utilizado

Existen diversos métodos para la construcción de las curvas de dosis-respuesta. Por una parte, el criterio para definir la “respuesta” puede variar. Algunos autores han usado la recidiva bioquímica (bien sea por criterio ASTRO o ASTRO_Phoenix) mientras que otros han usado el porcentaje de biopsias positivas post-radioterapia. Por otra parte, la definición de dosis es muy variable: dosis en el isocentro, dosis en periferia del volumen blanco, efecto de los cambios de técnica, etc. Y por último, la segmentación realizada y su significación estadística.

En nuestro caso hemos usado un criterio estricto, definiendo curvas distintas para distintas definiciones de “dosis” (D100, D100_corr, D90) y una sola definición de “respuesta” (recidiva bioquímica según criterio de la ASTRO). Por otra parte se ha utilizado un método interactivo de segmentación que nos ha permitido realizar distintas tentativas para la construcción de grupos de dosis. Por último, se realizó un análisis de la supervivencia y se consideró adecuada la curva de dosis respuesta que ofreciera resultados significativos para el análisis de la supervivencia (la mayoría de los autores no aplican este criterio estricto a la construcción de sus curvas).

8.5 DE LOS MODELOS PREDICTIVOS CLINICO-DOSIMETRICOS

La relación entre las variables clínicas y dosimétricas para nuestro grupo de pacientes es muy complicada por la alta dimensionalidad de la base de datos. Por ello y por el número relativamente bajo de eventos tras el tratamiento, el análisis multivariante reveló pocas relaciones significativas. Por otro lado, si se analizan pares de pacientes correspondientes al mismo grupo de riesgo y con valores de D90 y V100 similares (por poner un ejemplo), veremos que no siempre se comportan igual clínicamente, y que hay otras muchas variables en las que difieren.

Por esta razón decidimos estudiar la aplicación de modelos más complejos pero que aportasen más información de la combinación de todos los parámetros clínico-dosimétricos de la base de datos. Para hacerlo decidimos explorar un modelo estadístico muy conocido (el de regresión logística) y compararlo a dos modelos alternativos: la

red neuronal artificial y el modelo de partición por el algoritmo de RPA. Nuestro objetivo no ha sido la construcción de modelos aplicables de forma rutinaria en otro ámbito clínico asistencial, ya que para ello nuestro análisis hubiera debido incluir mecanismos de validación externa. Por el contrario, hemos pretendido evaluar los distintos modelos, su aplicabilidad, realizar una validación interna y comparar sus resultados.

8.5.1 Modelo de regresión logística (LR)

Los modelos de LR son muy conocidos y ampliamente utilizados en el ámbito de la Oncología (365, 366, 367, 368). Se trata de un modelo no-lineal de clasificación que usa un subconjunto de datos para derivar los coeficientes de una ecuación que calcula la probabilidad de que un nuevo caso pertenezca a una cierta clase.

- ∴ Las aplicaciones de los modelos de LR en el ámbito del cáncer de próstata son muy variadas. Forman et al. (369) establecen un modelo con capacidad para predecir qué pacientes necesitarían tratamiento adyuvante con hormonoterapia tras prostatectomía y radioterapia adyuvante, a partir de variables de la biología del tumor (estadio clínico, respuesta bioquímica e índice de Gleason).

- ∴ También se han empleado los modelos de LR para analizar la respuesta medida por biopsia post-radioterapia. En un estudio sobre 132 pacientes tratados en el MSKCC, Levengrün et al. (370) establecen una serie de modelos de LR para calcular la probabilidad de obtener una biopsia positiva usando un conjunto de variables clínicas y dosimétricas. El primer modelo (modelo_1) emplea D20, índice de Gleason y TCP^{Niemierko} (ROC = 0.73). El modelo_2 utiliza Dmean, índice de Gleason y estadio T (ROC = 0.70). El modelo_3 utiliza Dprescribed, índice de Gleason y estadio clínico (ROC = 0.67).

- ∴ Otros estudios han empleado los modelos LR para predecir la toxicidad rectal en pacientes sometidos a radioterapia. Nuevamente el grupo del MSKCC presenta sus resultados para predecir toxicidad rectal crónica usando un modelo de LR que incorpora variables clínicas y dosimétricas (208). Las variables clínicas incluidas fueron: edad, estadio clínico, índice de Gleason, PSA inicial, historia de diabetes mellitus y hormonoterapia neoadyuvante. Las variables dosimétricas consideradas fueron Dmax_rectum, Dmax_PTV, Dmean_rectum y Disocentro. Se incorporaron variables de la geometría del plan de tratamiento y un modelo NTCP para cálculo de probabilidad de toxicidad. El modelo para pacientes tratados en el segmento de dosis entre 70.2 – 75.6 Gy rindió un ROC de 0.69, mientras que el modelo para pacientes tratados por encima de 75.6 fue de 0.81. El elemento más interesante de este artículo es la equilibrada combinación de factores clínicos, dosimétricos y técnicos incluidos en el modelo. Los autores refieren en el texto que están realizando una validación del modelo sobre una serie independiente.

En este sentido también existe una experiencia española que merece ser destacada. El grupo del Hospital de la Princesa reporta su experiencia con la utilización de modelos de LR para predecir la toxicidad rectal (\geq G2) en pacientes sometidos a radioterapia externa por neoplasia de próstata (371). Los autores analizan 107 pacientes tratados a tres niveles de dosis (70, 72 y 75.6 Gy). Se construyó un modelo de regresión logística multivariante para evaluar las variables asociadas con el sangrado rectal y se calculó el área bajo la curva ROC para analizar la capacidad de respuesta del modelo. Cuando se incluyeron todas las variables en el modelo, sólo VR60 (volumen rectal que recibió

mas de 60 Gy) se mostró significativo [$\exp(\beta) = 1.105$; CI = 1.036 – 1.179; $p = 0.002$; área bajo ROC = 0.889] y cuando este último fue excluido, la dosis mediana en recto (Dmean_rectum) emergió como significativa [$\exp(\beta) = 1.268$; CI = 1.084 – 1.482; área bajo ROC = 0.892].

∴ Pisansky et al. (372) han presentado un modelo de LR para establecer los criterios de selección óptimos para la braquiterapia prostática a partir de una serie de 2905 pacientes tratados quirúrgicamente. Las variables introducidas en el modelo fueron estadio clínico, índice de Gleason y valor de PSA. El modelo calculó la probabilidad de encontrar afectación de vesículas seminales o ganglios linfáticos pélvicos. La capacidad para realizar una predicción correcta se evaluó mediante re-muestreo de 1000 muestras sobre la base de datos. Del modelo se dedujo que existe un subgrupo de pacientes con baja probabilidad de invasión de vesículas seminales o ganglios linfáticos pélvicos, que constituirían el grupo donde la braquiterapia permanente estaría indicada. Los criterios de inclusión derivados del modelo de regresión logística son:

[T1c – T2a] & [GS 2 - 6] & [PSA ≤ 16 ng/ml]

OR

[T1c – T2a] & [GS 7 - 10] & [PSA ≤ 4 ng/ml]

OR

[T2b – T2c] & [GS 2 – 6] & [PSA ≤ 6 ng/ml]

∴ En general, el aspecto más difícil y que más tiempo consume es la construcción del modelo, es decir, decidir qué variables se incluirán en el mismo. Para ello hay que tener cierto conocimiento de las posibles relaciones de causalidad entre las variables y el resultado final. En nuestro caso la base de datos contenía más de 150 variables, entre aquellas nativas y las transformaciones que luego se realizaron de algunas de ellas. Nosotros hemos usado el algoritmo “stepwise” de selección de variables, obteniendo un modelo de 5 variables: edad, densidad del PSA (PSAD), número de agujas por centímetro cúbico de próstata (Ag_cc), V80_0 y D100_corr. Este modelo de LR se considera el modelo de referencia al que se compararán los otros dos. En la revisión de la bibliografía que hemos realizado no hemos encontrado ninguna referencia relativa a la construcción de modelos de regresión logística incorporando variables clínico-dosimétricas a series de pacientes tratados con braquiterapia.

8.5.2 Modelo de partición por el algoritmo de RPA

En la sección correspondiente se ha descrito el algoritmo de RPA (“Recursive Partitioning Analysis”) y su aplicación en el ámbito de la Oncología. A lo largo de nuestro trabajo hemos hecho un uso intensivo de la plataforma de partición del paquete JMP6. Es una herramienta muy útil para realizar una segmentación supervisada o no-supervisada de grandes bases de datos y la interfaz de usuario es muy intuitiva. La hemos usado para analizar el papel de la hormonoterapia en la supervivencia. También ha sido muy importante para la construcción de las curvas de dosis-respuesta. Y por último hemos creado un modelo predictivo sobre su algoritmo.

∴ A nuestro conocimiento existe una única publicación que construya un modelo pronóstico sobre series de braquiterapia prostática por el algoritmo de RPA. Zelefsky et al. (166) en un estudio retrospectivo multi-institucional sobre 2693 pacientes tratados mediante braquiterapia exclusiva (sin tratamiento hormonal) procedentes de once Grupo 1: [PSA < 10 ng/ml] & [T1 – T2a] & [Dosis ≤ 71.79 Gy] & [cualquier GS]

instituciones diferentes. Del total de pacientes, el 68% fueron tratados con ^{125}I (dosis de prescripción: 144 Gy) y el 32% con ^{103}Pd (dosis de prescripción: 130 Gy). La mediana de seguimiento fueron 63 meses. Se definió el fallo terapéutico de acuerdo con la definición de la ASTRO y con la revisión de Phoenix para establecer comparaciones. La supervivencia libre de recidiva bioquímica a 8 años según los distintos grupos de riesgo fue de: 82%, 70% y 48% según ASTRO ($P < 0.001^*$) y del 74%, 61% y 39% según Phoenix ($p < 0.001^*$).

Los autores realizan asimismo un análisis de la supervivencia en función de los datos de la dosimetría post-implante. En el subgrupo de pacientes de ^{125}I se encontró un punto de corte significativo para D90 ($D90 \geq 130$ Gy) que se asoció con la supervivencia libre de recidiva bioquímica a 5 años según los criterios ASTRO (93% vs. 76%, $p < 0.001^*$) y Phoenix. En el subgrupo de pacientes tratados con ^{103}Pd se encontró un punto de corte significativo para D90 ($D90 \geq 115$ Gy) que se asoció con la supervivencia libre de recidiva bioquímica a 5 años según el criterio ASTRO (92% vs. 83%, $p = 0.01^*$). Sin embargo no hubo un punto de corte significativo al aplicar el criterio de Phoenix.

En el análisis multivariante los factores con asociación significativa con la supervivencia libre de recidiva bioquímica fueron PSA inicial, estadio tumoral, índice de Gleason, año de tratamiento y valor de D90 (aunque estas dos variables tenían una correlación muy alta entre si).

Se aplicó el algoritmo CART (una variante de RPA) para definir grupos pronósticos clínico-dosimétricos. Se hizo un análisis separado para cada isótopo. Para el grupo de pacientes tratados con ^{125}I el CART encontró tres subgrupos de riesgo. La supervivencia libre de recidiva bioquímica a 8 años para los tres grupos fue de 96%, 82% y 68%. A continuación se muestran las reglas que generan dichos grupos:

- Grupo 1: [PSA < 4.8 ng/ml]
- Grupo 2: [PSA \geq 4.8 ng/ml] & [D90 > 120 Gy]
- Grupo 3: [PSA \geq 4.8 ng/ml] & [D90 < 120 Gy]

Para el grupo de pacientes tratados con ^{103}Pd el CART encontró dos subgrupos de riesgo. La supervivencia libre de recidiva bioquímica a 5 años para los dos grupos fue de 94% y 82%. A continuación se muestran las reglas que generan dichos grupos:

- Grupo 1: [GS \leq 6] + [D90 > 115 Gy]
- Grupo 2: [Cualquiera, no incluido en grupo 1]

Este artículo es el primero en mostrar una segmentación de una serie larga de pacientes tratados con braquiterapia usando un algoritmo de partición. Para el grupo de ^{125}I se encontraron dos variables dominantes (PSA y D90) y se construyó un árbol de partición con tres grupos finales. De los resultados presentados en el artículo no podemos establecer si las diferencias de supervivencia son significativas o no. Para el grupo de ^{103}Pd se encontraron dos variables dominantes que generaron un árbol con dos grupos finales. Tampoco podemos establecer si las diferencias son significativas. Sin embargo los autores no exploran árboles más complejos que combinen un espectro más amplio de variables clínicas y dosimétricas, y no dan información acerca de la validación interna de su modelo ni de su capacidad de discriminación.

∴ Hemos encontrado un excelente artículo que presenta un modelo pronóstico sobre una serie de pacientes tratados mediante prostatectomía radical (317). En efecto, los autores analizan 485 pacientes tratados mediante prostatectomía radical entre 1991 y 1995. Se incluyen en el modelo las variables generales y de biología del tumor, así como los resultados del espécimen de la cirugía. Aplicando la variante del algoritmo CART para análisis de datos censurados construyen un primer árbol de clasificación con 6 hojas terminales. Estos grupos se combinaron para crear los 3 grupos pronósticos finales. Los primeros 6 grupos se seleccionaron minimizando el valor del test Log-Rank entre los subgrupos resultantes de la partición, pero no todas las comparaciones resultaban estadísticamente significativas. Combinando subgrupos se llegó a una partición final donde el test de Log-Rank mostró diferencias significativas entre grupos. Del árbol final se extrajeron las reglas que definían cada grupo:

Grupo 1: [Stage = CON, SM, EPE] + [GS ≤ 7] + [PSA ≤ 22.7ng/ml]
 Grupo 2: [Stage = CON, SM, EPE] + [GS ≤ 7] + [PSA > 22.7ng/ml]
 Grupo 3: [Stage = CON, SM, EPE] + [GS > 7]
 Grupo 4: [Stage = SV, NODE] + [PSA ≤ 24.1ng/ml] + [Age > 65]
 Grupo 5: [Stage = SV, NODE] + [PSA ≤ 24.1ng/ml] + [Age ≤ 65]
 Grupo 6: [Stage = SV, NODE] + [PSA > 24.1ng/ml]

CON: organ confined disease. SM: positive surgical margin. EPE: extrarostatic extensión. SV: seminal vesicle invasion. NODE: lymph node metastases.

Los grupos 2, 3 y 4 se agruparon en uno solo. Los grupos 5 y 6 se agruparon en uno solo. El resultado final son tres grupos de riesgo. Las tasas de recurrencia para cada uno de los grupos fueron del 13%, 53% y 80% respectivamente para los grupos 1 (buen pronóstico), 2 (riesgo intermedio) y 3 (mal pronóstico).

A pesar de que el método para la construcción del árbol es muy interesante y emplea una variante del algoritmo que permite manejar datos censurados, lo que la hace muy útil para analizar datos de supervivencia, los autores no realizan una validación del modelo (ni interna ni externa) lo que le da una validez limitada por el riesgo de sobre-ajuste (“overfitting”) que se puede producir. Por otra parte los autores no realizan un análisis de la capacidad discriminatoria del modelo respecto a los datos reales observados. En conclusión se trata de un buen ejemplo de clasificación de una base de datos usando el modelo RPA, pero tiene una utilidad limitada como modelo predictivo.

∴ Williams et al. (316), presentan un estudio retrospectivo sobre 256 pacientes tratados en una sola institución mediante radioterapia externa (dosis mediana: 64 Gy). Se aplicó un algoritmo de RPA usando parámetros biológicos del tumor (estadio clínico, valor inicial del PSA e índice de Gleason). El algoritmo RPA estableció 5 grupos pronósticos definidos por las siguientes reglas:

Grupo 1: [GS 2 - 6] & [PSA ≤ 10.6 ng/ml]
 Grupo 2: [GS 2 - 6] & [10.6 < PSA ≤ 20 ng/ml]
 Grupo 3: [GS 2 - 6] & [20 < PSA ≤ 42.4 ng/ml]
 Grupo 4: [GS 7-10] & [PSA < 42.4]
 Grupo 5: [cualquier GS] & [PSA > 42.4 ng/ml]

Las tasas de control bioquímico a 5 años (criterio ASTRO) para cada uno de los grupos de riesgo fueron las siguientes: grupo 1 [75.1%], grupo 2 [55%], grupo 3 [39%], grupo 4 [30.5%], grupo 5 [11.1%].

•• Shipley et al. (373), en un estudio retrospectivo multi-institucional (6 centros) sobre 1765 pacientes tratados con radioterapia externa (dosis mediana: 69.4 Gy), estudian los factores pronósticos usando parámetros biológicos del tumor (estadio clínico, valor inicial del PSA e índice de Gleason). El algoritmo RPA estableció 4 grupos pronósticos definidos por las siguientes reglas:

Grupo 1: [cualquier GS] & [PSA < 9.2 ng/ml]
 Grupo 2: [cualquier GS] & [9.2 ≤ PSA < 19.7]
 Grupo 3: [GS 2-6] & [PSA ≥ 19.7]
 Grupo 4: [GS 7-10] & [PSA ≥ 19.7]

Las tasas de control bioquímico a 5 años (criterio ASTRO) para cada uno de los grupos de riesgo fueron las siguientes: grupo 1 [81%], grupo 2 [69%], grupo 3 [47%], grupo 4 [29%].

•• Por otra parte, un reciente estudio multicéntrico sobre 4537 pacientes estudia la utilidad del algoritmo de RPA para la definición de grupos de riesgo usando exclusivamente parámetros biológicos del tumor (313). Los pacientes fueron divididos en tres grupos de dosis ([60 – 67 Gy], [67 – 70 Gy] y [70 – 75.7 Gy]) con 1862, 1288 y 1136 pacientes respectivamente. En cada uno de los tres grupos de dosis se aplicó el algoritmo de RPA para definir grupos de riesgo. En el grupo de [70 – 76 Gy] se definieron las siguientes reglas que definen los grupos de riesgo.

Grupo 1: [GS 2 - 7] & [PSA 0 - 9] & [cualquier T]
 Grupo 2: [GS ≤ 6] & [PSA 9 - 17] & [cualquier T]
 Grupo 3: [GS 7] & [PSA 9 - 17] & [cualquier T]
 Grupo 3: [GS ≤ 6] & [PSA 17 - 53] & [cualquier T]
 Grupo 4: [GS 7 - 10] & [PSA 17 - 53] & [T1b – T2a]
 Grupo 4: [GS 8 - 10] & [PSA 0 - 17] & [cualquier T]
 Grupo 5: [GS 7 - 10] & [PSA 17 - 53] & [T2b – T2c]
 Grupo 5: [cualquier GS] & [PSA 53 - 297] & [cualquier T]

Las tasas de control bioquímico a 5 y 8 años para cada uno de los grupos de riesgo fueron las siguientes: grupo 1 [94 y 87%], grupo 2 [86 y 78%], grupo 3 [77 y 60%], grupo 4 [69 y 45%], grupo 5 [40 y 19%].

•• El grupo del Fox Chase Cancer Center realizan un estudio similar sobre 939 pacientes, incluyendo factores clínicos y dosimétricos y presentando una clasificación en cuatro estratos de riesgo (314). Se usaron cuatro co-variables (valor de PSA inicial, estadio clínico, índice de Gleason y dosis de prescripción). Se establecieron 12 nodos terminales que se agruparon en 4 grupos de riesgo. Se definieron las siguientes reglas que definen los grupos de riesgo.

Grupo 1: [PSA < 20 ng/ml] & [T1 – T2a] & [Dosis > 72.35 Gy] & [cualquier GS]
 Grupo 1: [PSA < 20 ng/ml] & [T2b – T3] & [Dosis > 76.29 Gy] & [cualquier GS]
 Grupo 2: [PSA < 10 ng/ml] & [T2b – T3] & [Dosis < 71.79 Gy] & [cualquier GS]
 Grupo 2: [PSA < 10 ng/ml] & [cualquier T] & [71.79 < Dosis < 71.79 Gy] & [cualquier GS]
 Grupo 2: [10 < PSA ≤ 19.9 ng/ml] & [cualquier T] & [Dosis ≤ 72.35 Gy] & [cualquier GS]
 Grupo 2: [PSA < 20 ng/ml] & [T2b – T3] & [72.35 < Dosis < 76.29 Gy] & [cualquier GS]
 Grupo 2: [PSA > 20 ng/ml] & [T2b – T3] & [Dosis > 74.82 Gy] & [GS 2-6]
 Grupo 2: [PSA < 20 ng/ml] & [cualquier T] & [Dosis > 77.42 Gy] & [GS 7-10]
 Grupo 3: [PSA ≤ 20 ng/ml] & [cualquier T] & [Dosis > 74.82 Gy] & [cualquier GS]
 Grupo 3: [PSA ≤ 20 ng/ml] & [cualquier T] & [74.82 < Dosis < 77.42 Gy] & [GS 7-10]
 Grupo 4: [PSA ≥ 20 ng/ml] & [T1 – T2a] & [Dosis > 74.82 Gy] & [GS 2-6]

La supervivencia libre de recidiva bioquímica a 5 años para cada uno de ellos fue de 84%, 41%, 16% y 67%. Las diferencias entre los grupos fue estadísticamente significativa ($p < 0.0001^*$).

Este artículo demuestra por una parte que combinando variables clínicas y dosimétricas es posible construir modelos predictivos con significación estadística. Por otra parte, utilizando las técnicas de RPA la dosis emerge de nuevo como variable con significación estadística para la supervivencia libre de recidiva bioquímica.

•. En nuestro caso se ha obtenido un modelo de partición mediante un algoritmo de cálculo ligeramente distinto ya que la diferencia entre grupos se calcula minimizando el valor de la entropía resultante de la partición en cada nodo. Se definieron 7 grupos de riesgo. Estos grupos vienen definidos por las siguientes reglas:

Grupo 1: [IR] & [D100_0 < 78.4 Gy] & [PSAD ≥ 0.6157]
 Grupo 2: [IR] & [D100_0 < 78.4 Gy] & [PSAD < 0.6157]
 Grupo 3: [IR] & [D100_0 ≥ 78.4 Gy]
 Grupo 4: [LR] & [D100_0 < 52.3 Gy] & [PSAD ≥ 0.1584]
 Grupo 5: [LR] & [D100_0 < 52.3 Gy] & [PSAD < 0.1584]
 Grupo 6: [LR] & [D100_0 ≥ 52.3 Gy] & [PSAD ≥ 0.2096]
 Grupo 7: [LR] & [D100_0 ≥ 52.3 Gy] & [PSAD < 0.2096]

Estos siete grupos pronósticos se amalgamaron en tres grupos finales, que resultaron de las siguientes combinaciones de reglas obtenidas comparando las supervivencias actuariales por el método de Kaplan-Meier y calculando las diferencias en la supervivencia por el test de Log-Rank:

Grupo de bajo riesgo: Grupo 3 + Grupo 5 + Grupo 7
 Grupo de riesgo intermedio: Grupo 2 + Grupo 6
 Grupo de alto riesgo: Grupo 1 + Grupo 4

A diferencia de los modelos de partición anteriores el nuestro incorpora variables clínicas y dosimétricas. A diferencia del estudio de Zelefsky mencionado anteriormente las curvas de supervivencia presentan diferencias en el test de Log-Rank. De nuestro modelo se deriva la importancia de la técnica para obtener el control bioquímico de la enfermedad. Esto se ha visto también a la hora de analizar la curva de dosis-respuesta.

Por otra parte, hemos incorporado un método de validación interna del modelo mediante un algoritmo de “k-means crossvalidation”. El ajuste del modelo se realiza con todos los datos, y posteriormente se extraen al azar k subgrupos, sobre los que se aplica el modelo y se calcula la bondad del ajuste. En nuestro caso, como hemos visto anteriormente este proceso de validación interna se ha completado satisfactoriamente y el análisis de la capacidad de discriminación ha mostrado unos resultados muy buenos y que son comparables a los del modelo de LR (tabla 8.5_2).

La ventaja sobre el modelo de LR es que no necesitamos un conocimiento a priori de las relaciones entre las variables, por lo que se puede explorar la base de datos de forma interactiva, empleando un método mixto supervisado y no-supervisado. El inconveniente del modelo es que al no existir un conocimiento de las relaciones, en ocasiones el algoritmo en modo no-supervisado realiza particiones que no tienen sentido clínico, o bien que van contra el razonamiento clínico o bien que encuentran valores de corte poco convencionales para variables no convencionales. Por ejemplo en nuestro caso la densidad de PSA mostró gran fuerza clasificadora en el modo no-supervisado por lo que se incluyó en el modelo a pesar de que no existe bibliografía publicada que relacione esta variable con la supervivencia. Esto da idea por otra parte de la compleja dimensionalidad de las bases de datos médicas y de cómo es necesario a nuestro juicio la aplicación de métodos de razonamiento alternativos empleando las herramientas necesarias para ello.

El algoritmo de RPA presenta algunas ventajas adicionales que lo hacen muy útil. Puede usarse como herramienta para seleccionar atributos para modelos LR, más ortodoxos desde el punto de vista del razonamiento estadístico (374). Asimismo determinados autores lo han usado para definir valores de corte para variables como el PSA o el GS para definir subgrupos en pacientes con estadio clínico T1c (315). Otros autores lo han usado para analizar la cinética del PSA (311). En conclusión, los modelos de RPA son fáciles de implementar sobre nuestra plataforma de análisis de datos, son intuitivos, pueden funcionar supervisados o no-supervisados y presentan un comportamiento similar a los modelos de regresión logística.

8.5.3 Modelo de red neuronal artificial

A diferencia de los modelos obtenidos por RPA, las redes neuronales gozan de gran popularidad como herramienta de análisis de grandes bases de datos en el ámbito de la medicina (320), la oncología (329, 328, 321) y la uro-oncología en particular (330, 375, 333).

∴ Existen diferentes trabajos que comparan modelos de regresión logística con redes neuronales para predecir el estadio quirúrgico a partir de variables pre-quirúrgicas. El grupo del John Hopkins Hospital de Baltimore aplican los dos modelos a un grupo de 756 pacientes sometidos a prostatectomía radical (376). El modelo de red neuronal presentó una precisión en la clasificación mayor que el modelo de regresión logística (96% vs. 40% respectivamente). Dentro del ámbito nacional, el grupo del Hospital Miguel Servet de Zaragoza presentan un estudio similar, comparando cinco modelos de regresión logística con seis modelos de red neuronal. En su caso la comparación del área bajo la curva ROC no presentó diferencias significativas (0.84 para LR vs. 0.88 para ANN).

∴ Sin embargo pueden resultar más interesantes los trabajos que aplican redes neuronales a la predicción de la recidiva bioquímica tras el tratamiento oncológico, ya sea con cirugía, radioterapia externa o braquiterapia. Han et al. (337) del John Hopkins Hospital de Baltimore comparan un modelo de regresión logística con un modelo de red neuronal sobre una muestra de 600 pacientes sometidos a prostatectomía radical. La metodología de trabajo es muy similar a la que hemos presentado nosotros aunque en el artículo no se dan detalles de la construcción del modelo de regresión logística, sólo sus resultados. Para el modelo de red neuronal se utilizó la mitad de los pacientes para su construcción, un 20% para probar el modelo (validación interna) y un 30% para su validación (validación externa). La variable de la neurona de salida fue dicotómica (recidiva bioquímica vs. no-recidiva bioquímica). La comparación entre los dos modelos se realiza de forma cualitativa sin realizarse inferencia sobre las curvas ROC. El área bajo la curva ROC para el modelo de regresión logística fue de 0.68 y de 0.81 para el modelo de red neuronal.

∴ Porter et al. (377) presentan un trabajo más completo desde el punto de vista metodológico, aunque recurren a técnicas de re-muestreo (resampling) para contrarrestar el número de 175 pacientes disponibles para el análisis, número que es ligeramente insuficiente para obtener “hazard ratios” que den significación para el número de comparaciones que proponen. Los autores proponen un modelo de red neuronal para predecir el riesgo de recidiva bioquímica. La novedad metodológica es que comparan un modelo de análisis de componentes principales, un árbol de decisión (similar a nuestro modelo RPA) y un modelo de regresión logística para determinar las variables que entraron en la red neuronal. De este modo la red neuronal de mejor comportamiento mostró un área bajo la curva ROC de 0.80, con una sensibilidad del 74% y una especificidad del 78%, un VPP de 71% y un VPN de 81%.

∴ Gulliford et al. (338) describen la aplicación de un modelo basado en una red neuronal para predecir el riesgo de recidiva bioquímica en un grupo de 126 pacientes tratados mediante dos esquemas de radioterapia en el marco de un ensayo randomizado. Los resultados de la red neuronal fueron comparados con un modelo biológico para control tumoral (TCP) y otro para toxicidad (NTCP). La red neuronal fue entrenada usando todos los datos y posteriormente se realizó una validación interna por el algoritmo de “k-means crossvalidation”. Se construyeron tres redes neuronales para predecir el control bioquímico, la toxicidad rectal y la toxicidad urinaria. En la primera red neuronal se obtuvieron valores de sensibilidad y especificidad del 66.83% y 64.17%. Para la segunda de 66.07% y 56.39% y para la última de 70.19% y 52.73%. La conclusión para los autores es positiva en el sentido de que consideran que las redes neuronales son capaces de aprender de los datos de entrenamiento para establecer previsiones aceptables sobre los datos de validación. Su utilidad es para los autores superior a los modelos TCP – NTCP, aunque el pequeño número de pacientes y de eventos impiden generalizar sus resultados.

∴ Después de revisar la bibliografía de forma repetida y exhaustiva, a nuestro conocimiento sólo ha llegado un trabajo que aplique redes neuronales a los resultados de la braquiterapia prostática. Dicho trabajo proviene del grupo de Seattle y analiza los resultados de una red neuronal sobre una población de 162 pacientes sometidos a braquiterapia exclusiva o tratamiento combinado con radioterapia externa y braquiterapia (151). Se utilizó para el estudio una variante de red neuronal conocida como “turboprop”, que usa un algoritmo de retro-propagación (“back-propagation”)

que resulta más rápido en su modo de aprendizaje. Se dividió la base de datos en dos grupos: el grupo de entrenamiento con el 84% de los pacientes y el grupo de validación con el 16% restante. Una vez decidida la estructura de la red neuronal (número óptimo de neuronas ocultas) y tras realizar el entrenamiento de la misma se utilizó dicha red para conocer sus predicciones sobre el grupo de validación. Los resultados de la red se compararon con los datos reales mediante un análisis de la contingencia. Las variables incluidas en la red fueron: edad, estadio clínico, PSA inicial, GS y adición de RTE (si-no). La red neuronal se ejecutó dos veces, una considerando la RTE y otra sin considerarla. De esta forma se pretendió estudiar el papel de la misma en concomitancia con la braquiterapia permanente. Los resultados de la red neuronal se compararon a los de un modelo de regresión logística clásico. La red neuronal fue capaz de predecir la recidiva bioquímica a 10 años con una precisión del 76% (66% para el modelo LR). Los valores de sensibilidad y especificidad fueron del 55% y 90% respectivamente (15% y 94% para el modelo LR). El valor predictivo positivo fue del 76% (59% para el modelo de LR) y el valor predictivo negativo del 82% (64% para el modelo de LR).

∴. A diferencia de los modelos de red neuronal revisados anteriormente el nuestro incorpora variables clínicas y dosimétricas. Al igual que para el modelo de partición y el modelo de regresión logística, esto nos permite estudiar la influencia que la técnica tiene sobre el control de la enfermedad. En la tabla 8.5_1 se presenta una comparación de los modelos revisados previamente junto con los datos del nuestro.

Autor	Precis (%)	Sens (%)	Esp (%)	VPP (%)	VPN (%)	ROC
Han (337)	NA	NA	NA	NA	NA	0.81
Porter (377)	NA	74	78	71	81	0.80
Gulliford (I) (338)	NA	66	64	NA	NA	NA
Gulliford (II)	NA	66	56	NA	NA	NA
Gulliford (III)	NA	70	52	NA	NA	NA
Ragde (151)	76	55	90	76	82	NA
Nuestro modelo	97	84	98	91	97	0.85

Tabla 8.5_1 Capacidad de discriminación de los distintos modelos de red neuronal estudiados

Hemos hecho un intento por evaluar la robustez del modelo de red neuronal mediante un algoritmo de validación de “k-means crossvalidation”. El ajuste del modelo se realizó con todos los datos, y posteriormente se extrajeron al azar k subgrupos, sobre los que se aplicó el modelo y se calculó la bondad del ajuste. A falta de una validación externa, los resultados del modelo se comparan favorablemente tanto al resto de la literatura como a los del modelo de partición y de regresión logística.

∴. La desventaja de las redes neuronales es que se comportan como “cajas negras” de las cuales se extrae una información pero de las que no se puede extraer el mecanismo de razonamiento. Podemos conocer su predicción pero no aportan un razonamiento, y por tanto no podemos construir a partir de ellas un modelo de explicación causal

sino un modelo empírico. Por otra parte, las redes neuronales no pueden manejar información cualitativa en su mecanismo de cálculo. Finalmente las redes neuronales no son aplicadas de forma más intensiva en el ámbito médico debido a que en muchas ocasiones presentan resultados muy optimistas en cuanto a su capacidad de predicción que no se ven reproducidos por otros autores. Ello es debido al fenómeno del sobreajuste (“overfitting”) que se produce cuando no se seleccionan adecuadamente las variables o cuando no se realiza un validación adecuada interna y externa.

Sin embargo las redes neuronales presentan ventajas en su aplicación al análisis de bases de datos médicas. Su capacidad de adaptación las hacen especialmente útiles en problemas de clasificación que manejen datos cuantitativos. Las redes neuronales pueden ser fácilmente reprogramadas. Otro aspecto esencial que aumenta su utilidad es el hecho de que son capaces de realizar un razonamiento no-lineal sin necesidad de imponer reglas a priori por parte del operados. Por último, el resultado final es una ecuación con múltiples coeficientes conocidos, que se puede aplicar de forma sencilla a los casos nuevos estudiados. Todo ello convierte a las redes neuronales en una herramienta de primer orden en el análisis de bases de datos médicas.

8.5.4 Resumen de los modelos predictivos.

De la revisión de la bibliografía respecto a modelos predictivos en braquiterapia prostática se pueden establecer dos conclusiones fundamentales: la primera es que no existe experiencia amplia en la creación de modelos combinando variables clínicas y dosimétricas. La segunda es que los modelos encontrados se comportan bastante bien, independientemente de la metodología de construcción, y rinden valores para el área bajo la curva ROC entre 0.85 y 0.95. En nuestro caso, la tabla 8.5_2 resume el comportamiento de los modelos generados. Como vemos el modelo de red neuronal consigue una precisión mayor, aunque el modelo de partición por RPA consigue una especificidad y valor predictivo positivo del 100%.

Parámetro	LR	ANN	RPA
Area bajo la curva ROC	0.95	0.85	0.90
Precisión (“accuracy”)	92%	97.10%	92.95%
Sensibilidad	55.55%	84.00%	38.46%
Especificidad	96.96%	98.00%	100.00%
Valor predictivo positivo	71.42%	91.66%	100.00%
Valor predictivo negativo	94.11%	97.89%	92.15%

Tabla 8.5_2 Capacidad discriminatoria de los distintos modelos clínico-dosimétricos construidos

Consideración respecto al uso de los modelos previos.

La primera parte de nuestro trabajo ha estado dedicada al examen retrospectivo de los resultados de la braquiterapia con semillas en nuestra base de datos de pacientes. Las conclusiones que se han establecido a partir del análisis estadístico estaban implícitas en los datos. Sin embargo, la parte de nuestro trabajo dedicada a la elaboración de modelos predictivos tiene otro enfoque distinto, ya que a partir de estos datos implícitos hemos construido unos modelos matemáticos que pretenden predecir los acontecimientos de pacientes que no estaban incluidos originariamente en nuestra base de datos. Por eso escribimos estas líneas de consideración final al respecto. Hay que decir que el objetivo inicial de nuestra tesis no es el de construir este tipo de modelos, como se ha especificado en la descripción de objetivos sino el de establecer el valor y la importancia de los parámetros de la dosimetría post-implante en los

resultados de control de la enfermedad. Ocurre sin embargo que estos modelos son un subproducto muy interesante de los datos recogidos y probablemente merecería la pena que fueran estudiados en un trabajo adicional. Para que el capítulo de modelado tal y como lo hemos planteado tenga validez formal debería establecerse un sistema más cuidadoso de selección de variables (probablemente construyendo centenares o miles de modelos distintos que incluyeran cada uno un subconjunto de variables) y un sistema más estricto de validación que incluyera por supuesto la validación interna que hemos realizado pero también una validación temporal de casos de la misma institución y una validación externa con datos de pacientes provenientes de otras instituciones. De esta forma se podría calibrar el modelo de forma más precisa y sus resultados serían más fiables.

Lo dicho anteriormente no resta valor a nuestra sección de modelos predictivos sino que la sitúan en su lugar. Los modelos construidos se han comportado como excelentes herramientas de clasificación de nuestra base de datos y complementan los resultados obtenidos por las técnicas de inferencia clásicas como el análisis de la supervivencia y los modelos de Cox. Pero nuestros modelos no pueden ser aplicados fuera de nuestra base de datos sin el proceso de calibración y validación referido anteriormente.