



Programa de Doctorado 3139 de Medicina

Línea de investigación: Cirugía general

Modelo predictivo de recidiva precoz de carcinoma hepatocelular postrasplante basado en inteligencia artificial (redes bayesianas)

TESIS DOCTORAL

Presentada por

D. Rafael García Domínguez

Dirigida por

Dra. Dña. Eva María Montalvá Orón

Dr. D. Antonio José Serrano López

Dr. D. Rafael López Andújar

Valencia, octubre 2023

Dña. Eva María Montalvá Orón, Médica Adjunta de la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Servicio de Cirugía General y del Aparato Digestivo del Hospital Universitari i Politècnic La Fe de Valencia,

D. Antonio José Serrano López, Profesor Titular del Departamento de Ingeniería Electrónica de la Escuela Técnica Superior de la Universidad de Valencia,

y D. Rafael López Andújar, Jefe de Servicio de Cirugía General y del Aparato Digestivo del Hospital Universitari i Politècnic La Fe de Valencia,

como directores de la Tesis Doctoral titulada **“Modelo predictivo de recidiva precoz de carcinoma hepatocelular postrasplante basado en inteligencia artificial (redes bayesianas)”**

AUTORIZAMOS A D. Rafael García Domínguez, licenciado en Medicina y autor de la misma bajo nuestra dirección, para su lectura y defensa.

Dra. Dña. E. M. Montalvá Orón

Dr. D. A. J. Serrano López

Dr. D. R. López Andújar

Dra. Dña. Eva María Montalvá Orón,

Doctora en Medicina y Cirugía y Médica Adjunta de la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Servicio de Cirugía General y del Aparato Digestivo del Hospital Universitari i Politècnic La Fe de Valencia

CERTIFICA QUE el trabajo de Tesis Doctoral titulado “**Modelo predictivo de recidiva precoz de carcinoma hepatocelular postrasplante basado en inteligencia artificial (redes bayesianas)**”, ha sido realizado íntegramente y bajo mi dirección, compartida con el Prof. D. Antonio José Serrano López y el Dr. D. Rafael López Andújar, por D. Rafael García Domínguez y que puede ser defendida públicamente ante la comisión correspondiente para la obtención del Título de Doctor en Medicina y Cirugía.

Valencia, octubre 2023

Fdo.: Dra. Dña. E. M. Montalvá Orón

Prof. D. Antonio José Serrano López,

Doctor en Física y Profesor Titular del Departamento de Ingeniería Electrónica de la Escuela Técnica Superior de Ingeniería de la Universidad de Valencia

CERTIFICA QUE el trabajo de Tesis Doctoral titulado “**Modelo predictivo de recidiva precoz de carcinoma hepatocelular postrasplante basado en inteligencia artificial (redes bayesianas)**”, ha sido realizado íntegramente y bajo mi dirección, compartida con la Dra. Dña. Eva María Montalvá Orón y el Dr. D. Rafael López Andújar, por D. Rafael García Domínguez y que puede ser defendida públicamente ante la comisión correspondiente para la obtención del Título de Doctor en Medicina y Cirugía.

Valencia, octubre 2023

Fdo.: Prof. D. A. J. Serrano López

Dr. D. Rafael López Andújar,

*Doctor en Medicina y Cirugía y Jefe de Servicio de Cirugía General
y del Aparato Digestivo del Hospital Universitari i Politècnic La Fe
de Valencia*

CERTIFICA QUE el trabajo de Tesis Doctoral titulado “**Modelo predictivo de recidiva precoz de carcinoma hepatocelular postrasplante basado en inteligencia artificial (redes bayesianas)**”, ha sido realizado íntegramente y bajo mi dirección, compartida con la Dra. Dña. Eva María Montalvá Orón y el Prof. D. Antonio José Serrano López, por D. Rafael García Domínguez y que puede ser defendida públicamente ante la comisión correspondiente para la obtención del Título de Doctor en Medicina y Cirugía.

Valencia, octubre 2023

Fdo.: Dr. D. R. López Andújar

Agradecimientos

A los doctores Antonio José Serrano López, Rafael López Andújar y, muy especialmente, a la doctora Eva María Montalvá Orón, por participar como directores en el desarrollo de esta tesis. Por sus sabios consejos en la elaboración de la misma. Son para mí un modelo de excelencia en el ámbito profesional, científico y académico. Su confianza y dedicación constantes han sido esenciales para llevar a cabo este proyecto.

Al Dr. Salvador Pous Serrano, maestro y amigo, cuyas lecciones, tanto quirúrgicas como vitales, guardo con cariño. Su calidad humana y profesional ha sido y será siempre un ejemplo para mí.

A tantos compañeros en la cirugía que merecerían ser aquí nombrados, de quienes siempre he tratado de aprender para ser mejor en mi día a día.

A mi familia y amigos, quienes han sido un pilar fundamental en una labor gratificante, pero que tiene también momentos de dificultad en los que las cosas no fluyen con la facilidad y sencillez que uno desearía. He sentido su aliento todos los días y todas las horas. Sólo ellos saben el tiempo que les debo y que espero compensar con creces.

*A mis padres, Rafa y Marcela,
por su cariño y paciencia.
Por el ejemplo de renunciar
a sus cosas para atender
las de otros,
tantas veces las mías*

*A Ana, por ser
la alegría
de mis días*

*A mi hermano Borja,
por su ejemplo constante
de tenacidad y esfuerzo*

“Conoce todas las teorías. Domina todas las técnicas, pero
al tocar un alma humana, sé **apenas otra alma humana”**

Carl Gustav Jung

Índice

Índice de contenidos

1.- INTRODUCCIÓN.....	3
1.1.- Generalidades.....	3
1.1.1.- Papel del trasplante hepático en el tratamiento del carcinoma hepatocelular.....	3
1.1.2.- Tratamiento pretrasplante de los pacientes con carcinoma hepatocelular.....	10
1.1.3.- Evaluación de la respuesta tumoral al tratamiento pretrasplante del carcinoma hepatocelular.....	17
1.2.- Factores pronósticos de recidiva del carcinoma hepatocelular tras el trasplante hepático.....	22
1.2.1.- Factores morfológicos.....	22
1.2.2.- Estadificación TNM.....	23
1.2.3.- Factores anatomopatológicos.....	23
1.2.4.- Alfafetoproteína.....	25
1.2.5.- Respuesta al tratamiento pretrasplante.....	26
1.2.6.- Serología del receptor.....	28
1.2.7.- Factores de biología molecular.....	29
1.2.8.- Estatus inflamatorio pretrasplante.....	30

1.2.9.- <i>Scores</i> y criterios pronósticos combinados.....	31
1.3.- Modelos matemáticos de predicción.....	42
1.3.1.- Características de un modelo predictivo.....	42
1.3.2.- Modelos gráficos probabilísticos.....	53
1.3.3.- Redes bayesianas.....	56
1.3.3.1.- Funcionamiento de las redes bayesianas..	58
1.3.3.2.- Resolución de problemas mediante redes bayesianas en biomedicina.....	66
1.4.- Estado actual del tema.....	72
2.- HIPÓTESIS DE TRABAJO.....	85
3.- OBJETIVOS DEL ESTUDIO.....	89
4.- MATERIAL Y MÉTODOS.....	93
4.1.- Diseño del estudio.....	93
4.2.- Población del estudio.....	94
4.3.- Criterios de inclusión y exclusión.....	96
4.4.- Protocolo de tratamiento pretrasplante.....	97

4.5.- Evaluación de la respuesta tumoral al tratamiento pretrasplante.....	107
4.6.- Criterios de inclusión en lista de espera de trasplante hepático para pacientes con carcinoma hepatocelular.....	109
4.7.- Seguimiento de los pacientes después del trasplante hepático.....	112
4.8.- Definición de las variables del estudio.....	115
4.9.- Métodos estadísticos del estudio.....	123
4.9.1.- Análisis descriptivo de la serie.....	123
4.9.2.- Estudio analítico univariante.....	124
4.9.3.- Estudio analítico multivariante.....	125
4.9.3.1.- Diseño del modelo predictivo basado en regresión logística.....	126
4.9.3.2.- Diseño del modelo predictivo basado en redes bayesianas.....	128
4.9.4.- Función de utilidad.....	130
4.9.5.- Generalización de los modelos desarrollados.....	134
4.9.6.- Comparación de los modelos desarrollados.....	135
4.10.- Interfaz informática para la aplicación clínica del modelo predictivo.....	137

5.- RESULTADOS.....	141
5.1.- Análisis descriptivo de la población de estudio.....	141
5.1.1.- Análisis descriptivo de las características de los receptores.....	143
5.1.2.- Análisis comparativo de las características de los receptores del grupo de entrenamiento frente al grupo de generalización.....	150
5.1.3.- Estudio descriptivo de supervivencia y recurrencia del carcinoma hepatocelular en los receptores.....	157
5.2.- Estudio analítico del grupo de entrenamiento.....	163
5.2.1.- Estudio univariante de la relación de las variables pretrasplante con la recurrencia precoz postrasplante del carcinoma hepatocelular.....	163
5.2.2.- Estudio multivariante de la relación de las variables pretrasplante con la recurrencia precoz postrasplante del carcinoma hepatocelular.....	166
5.2.2.1.- Modelo de regresión logística.....	166
5.2.2.2.- Modelo de redes bayesianas.....	170
5.3.- Generalización de los modelos desarrollados.....	174
5.3.1.- Modelo de regresión logística.....	174
5.3.2.- Modelo de redes bayesianas.....	177

6.6.- Aplicación clínica del modelo de redes bayesianas. Predicción de la recurrencia precoz postrasplante del carcinoma hepatocelular.....	245
7.- CONCLUSIONES.....	253
8.- BIBLIOGRAFÍA.....	257

Lista de tablas

- Tabla 01 Criterios de indicación de trasplante en el carcinoma hepatocelular*
- Tabla 02 Criterios RECIST*
- Tabla 03 Criterios mRECIST (para “lesiones diana”)*
- Tabla 04 Criterios mRECIST (para “lesiones no diana”)*
- Tabla 05 Grados de diferenciación tumoral y supervivencia global del CHC*
- Tabla 06 Índices y criterios pronósticos combinados*
- Tabla 07 Resumen de criterios y scores pronósticos de trasplante para CHC*
- Tabla 08 División de los pacientes del estudio según grupos para el desarrollo de modelos predictivos*
- Tabla 09 Esquema de tratamiento radioembolización en la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Hospital Universitari i Politècnic La Fe*
- Tabla 10 RETREAT score*
- Tabla 11 Riesgo de recidiva de CHC a 5 años según RETREAT score y seguimiento después del TH*

- Tabla 12 Relación entre el valor del coeficiente K y el grado de concordancia entre 2 medidas*
- Tabla 13 Estudio descriptivo de las características de base de los pacientes incluidos en el estudio*
- Tabla 14 Estudio descriptivo de las variables relacionadas con la hepatopatía de base de los pacientes incluidos en el estudio*
- Tabla 15 Estudio descriptivo de las variables referentes a la carga tumoral de los pacientes incluidos en el estudio*
- Tabla 16 Estudio descriptivo de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (I)*
- Tabla 17 Estudio descriptivo de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (II)*
- Tabla 18 Estudio descriptivo de las variables referentes a la evolución postrasplante de los pacientes del estudio (recurrencia y mortalidad)*
- Tabla 19 Estudio descriptivo comparativo de las variables referentes a las características de base del grupo de entrenamiento frente al grupo de generalización*
- Tabla 20 Tabla 20: Estudio descriptivo comparativo de las variables referentes a la hepatopatía de base del grupo de entrenamiento frente al grupo de generalización*

- Tabla 21 Estudio descriptivo comparativo de las variables referentes a la carga tumoral de los pacientes del grupo de entrenamiento frente al grupo de generalización*
- Tabla 22 Estudio descriptivo comparativo de las variables de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (I)*
- Tabla 23 Estudio descriptivo comparativo de las variables de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (II)*
- Tabla 24 Estudio descriptivo comparativo de las variables de las variables referentes a la evolución postrasplante (recurrencia y mortalidad)*
- Tabla 25 Relación de variables y constante del modelo de regresión logística en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC*
- Tabla 26 Matriz de confusión del modelo de regresión logística para predecir la recurrencia precoz postrasplante del CHC en el grupo de entrenamiento*
- Tabla 27 Matriz de confusión del modelo de redes bayesianas para predecir la recurrencia precoz postrasplante del CHC en el grupo de entrenamiento*
- Tabla 28 Matriz de confusión del modelo de regresión para predecir la recurrencia precoz postrasplante del CHC en el grupo de generalización*

Tabla 29 Matriz de confusión del modelo de redes bayesianas para predecir la recurrencia precoz postrasplante del CHC en el grupo de generalización

Tabla 30 Valores de utilidad otorgados a los diferentes escenarios clínicos según la predicción del modelo

Lista de figuras

- Figura 01 Clasificación Barcelona Clinic Liver Cancer en su versión de 2022*
- Figura 02 Clasificación de las situaciones resultantes tras la aplicación de un modelo predictivo en forma de matriz de confusión*
- Figura 03 Representación gráfica de una curva ROC*
- Figura 04 Utilidad de un modelo predictivo*
- Figura 05 Ejemplos de grafo dirigido acíclico (A) y de grafo no dirigido (B)*
- Figura 06 Representación de las dependencias probabilísticas entre variables aleatorias, con nodos “padres” e “hijos”*
- Figura 07 Naïve Bayes (izquierda) y Tree Augmented Naïve (derecha)*
- Figura 08 Representación gráfica de un modelo general de red bayesiana pronóstica*
- Figura 09 Diagrama de influencia para selección de tratamiento*
- Figura 10 Número de trasplantes realizados anualmente por CHC en la Unidad de Cirugía Hepatobiliopancreática y Trasplante. Hospital Universitari i Politècnic La Fe*
- Figura 11 Curva de supervivencia global postrasplante de los pacientes incluidos en el estudio*

Figura 12 Análisis comparativo de supervivencia global postrasplante de los pacientes del grupo de entrenamiento frente al grupo de generalización

Figura 13 Curva de supervivencia libre de enfermedad postrasplante de los pacientes incluidos en el estudio

Figura 14 Análisis comparativo de supervivencia libre de enfermedad postrasplante de los pacientes del grupo de entrenamiento frente al grupo de generalización

Figura 15 Función de incidencia de recurrencia postrasplante del CHC de los pacientes incluidos en el estudio

Figura 16 Análisis comparativo de la supervivencia global entre el grupo con recurrencia precoz postrasplante (< 2 años) frente al grupo con recurrencia > 2 años y sin recurrencia

Figura 17 Resultados del análisis univariante de las variables cuantitativas relacionadas con la recurrencia precoz del CHC postrasplante

Figura 18 Resultados del análisis univariante de las variables cualitativas relacionadas con la recurrencia precoz del CHC postrasplante

Figura 19 Curva ROC obtenida con el modelo de regresión logística en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

Figura 20 Grafo del modelo de redes bayesianas para predicción de recurrencia precoz postrasplante del CHC

Figura 21 Curva ROC obtenida con el modelo de redes bayesianas en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

Figura 22 Curva ROC obtenida con el modelo de regresión logística en el grupo de generalización para determinar la recurrencia precoz postrasplante del CHC

Figura 23 Curva ROC obtenida con el modelo de redes bayesianas en el grupo de generalización para determinar la recurrencia precoz postrasplante del CHC

Figura 24 Comparación de las curvas ROC obtenida con los modelos de regresión logística y redes bayesianas en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

Figura 25 Comparación de las curvas ROC obtenida con los modelos de regresión logística y redes bayesianas en el grupo de generalización para determinar la recurrencia precoz postrasplante del CHC

Figura 26 Formulario para la inclusión de datos en la aplicación web de acceso abierto: <http://oceano.uv.es:3838/RBRecidiva>

Símbolos, abreviaturas y siglas

AAD	Antivirales de Acción Directa
ABC	Área Bajo la Curva
AFP	Alfafetoproteína
BCLC	Barcelona-Clinic <i>Liver Cancer</i>
CHC	Carcinoma hepatocelular
CMD	Comité multidisciplinar
DCP	Des-gamma-carboxi-protrombina
EASL	<i>European Association for the Study of the Liver</i>
EE	Enfermedad Estable
EORTC	<i>European Organisation for Research and Treatment of Cancer</i>
GDA	Grafo Dirigido Acíclico
HALT-HCC	<i>Hazard Associated with Liver Transplantation for Hepatocellular carcinoma</i>
IC	Intervalo de Confianza
MELD	<i>Model for End-stage Liver Disease</i>
MGP	Modelo Gráfico Probabilístico
MO	Microondas
MORAL	<i>Model of Recurrence After Liver Transplant</i>
mRECIST	<i>modified Response Evaluation Criteria In Solid Tumours</i>
NCI	<i>National Cancer Institute</i>
NCIC-CTG	<i>National Cancer Institute of Canada-Clinical Trials Group</i>
PE	Progresión de Enfermedad
PET	Tomografía por emisión de positrones
QETA	Quimioembolización Transarterial
RB	Redes Bayesianas
RC	Respuesta Completa
RECIST	<i>Response Evaluation Criteria In Solid Tumours</i>

RETA	Radioembolización Transarterial
RF	Radiofrecuencia
RL	Regresión Logística
RM	Resonancia Magnética
RNL	Ratio Neutrófilos / Linfocitos
ROC	<i>Receiver Operating Characteristic</i>
RP	Respuesta Parcial
RPL	Ratio Plaquetas / Linfocitos
SETH	Sociedad Española de Trasplante Hepático
SG	Supervivencia Global
SLE	Supervivencia Libre de Enfermedad
TAN	<i>Tree Augmented Network</i>
TBS	<i>Tumor Burden Score</i>
TC	Tomografía computarizada
TH	Trasplante Hepático
TRAIN	<i>Time - Radiological response - Alpha-phetoprotein – Inflammation</i>
UCSF	Universidad de California San Francisco
UFN	Utilidad Falso Negativo
UFP	Utilidad Falso Positivo
UVN	Utilidad Verdadero Negativo
UVP	Utilidad Verdadero Positivo
VHB	Virus de la hepatitis B
VHC	Virus de la hepatitis C
VN	Verdadero Negativo
VP	Verdadero Positivo
VPN	Valor Predictivo Negativo
VPP	Valor Predictivo Positivo

Introducción

1.- Introducción

1.1.- Generalidades

1.1.1.- Papel del trasplante hepático en el tratamiento del carcinoma hepatocelular

Desde sus inicios, con el trasplante hepático (TH) siempre se ha buscado la curación de tumores hepáticos irresecables tanto primarios como secundarios. De hecho, el primer TH “exitoso” fue realizado por Thomas E. Starzl el 23 de julio de 1967 en un niño con un carcinoma hepatocelular (CHC) irresecable en el contexto de una atresia de vías biliares (1).

Durante décadas, el papel del TH como tratamiento de neoplasias malignas hepáticas fue muy cuestionado debido a los elevados porcentajes de recurrencia, ocasionados por la ausencia de criterios de selección estandarizados (2). La situación cambió con la publicación del estudio de V. Mazzaferro y colaboradores en 1996, que demostraba una supervivencia global a 4 años superior al 75 % y una supervivencia libre de enfermedad (SLE) de 83 % después de trasplante de CHC con unos criterios estrictos: tumores únicos menores de 5 cm o hasta 3 nódulos menores de 3 cm de diámetro (3). Posteriormente estos principios fueron establecidos y conocidos como

Criterios de Milán, que siguen siendo la referencia fundamental para ofrecer los mejores resultados postrasplante en pacientes con CHC y han sido ampliamente aceptados por los grupos de trasplante de todo el mundo (4). Comparativamente con otras terapias con intención curativa como la ablación o la resección quirúrgica, cuyo talón de Aquiles es una elevada recurrencia (70 % a 5 años) (5), el TH con los criterios de Milán ofrece unos resultados de supervivencia a 5 años similares (60-80 %), con una recurrencia mucho más baja, que se sitúa en torno al 15 % (6). Así pues, hoy es posible afirmar que el TH es el mejor tratamiento curativo para el CHC, ya que permite la exéresis completa del tumor y la sustitución del hígado patológico subyacente.

En este contexto, los buenos resultados del TH han animado a muchos grupos a ampliar los criterios de trasplante en los pacientes con CHC, siempre que no suponga un incremento considerable de la recurrencia ni una disminución significativa de la supervivencia a largo plazo. De todos los presentados a continuación, únicamente los Criterios de San Francisco y los *Up-to-seven* (o *Metroticket*) han sido validados externamente.

Criterios de San Francisco

Establecidos por la Universidad de California San Francisco (UCSF) en el año 2001 como la presencia de un único tumor de diámetro $\leq 6,5$ cm o 3 tumores con una suma total de los diámetros ≤ 8 cm. Con estos criterios, la recurrencia publicada es del 11,4 % y la supervivencia a 5 años es comparable a la de los criterios de Milán (7). El estudio inicial que demostró la utilidad de estos criterios se basaba en datos anatomopatológicos obtenidos de los hígados explantados: aquellas piezas con nódulos que excedían discretamente los criterios de Milán tenían supervivencias y recurrencia similar a los pacientes trasplantados con tumores intraMilán. Posteriormente los datos se validaron tanto interna como externamente con la inclusión de pacientes en base al tamaño de los nódulos tumorales en estudios de imagen pretrasplante (8,9). Se estima que la utilización de estos criterios expandidos ha aumentado el acceso de pacientes con CHC a un TH en un 5-20 % (10)

Criterios Up-to-seven

También conocidos como *Metroticket*, y definidos como la suma del diámetro tumoral (cm) del tumor más grande y el número de tumores con un resultado igual o inferior a 7, basados exclusivamente en datos aportados por las pruebas de imagen pretrasplante. Fueron establecidos en 2009 por el grupo

de Milán, a partir de una cohorte de 1556 pacientes trasplantados entre 1984 y 2006 en 36 centros. El TH en base a estos criterios ha demostrado una supervivencia a 5 años de 71,2 %, similar a los criterios de Milán, frente a los pacientes que excedían estos criterios (48,1 %) (11).

Criterios de Navarra

Un único nódulo de diámetro ≤ 6 cm o 2-3 nódulos ≤ 5 cm cada uno. Estos criterios se establecieron en 2007 en la Clínica Universitaria de Navarra y, aunque no han sido validados externamente, la supervivencia a 5 años (73 %) y recurrencia (8 %) de los pacientes trasplantados bajo estos criterios también es similar a los resultados de Milán (12).

Criterios de Valencia

Definidos en 2008 por el grupo del Hospital Universitari i Politècnic La Fe de Valencia como hasta 3 tumores de ≤ 5 cm de diámetro siempre que la suma total de los mismos sea ≤ 10 cm. La supervivencia y recurrencia a los 5 años de los pacientes trasplantados con estos criterios es de 69 % y 11 % frente al 40 % y 43 %, respectivamente, de los pacientes que exceden dichos criterios (13).

Criterios de Seúl

Definidos en 2008 en el Asan Medical Center de Seúl como la presencia de hasta 6 nódulos de tamaño ≤ 5 cm sin invasión macrovascular. La recidiva a los 3 años y la supervivencia a 5 años de los pacientes trasplantados bajo estos criterios es de 9,1 % y 76,3 % respectivamente (14).

Criterios de Toronto

Se publicaron en el año 2011 (15) y son unos criterios menos restrictivos que permiten el trasplante de cualquier CHC limitado sólo al hígado, independientemente del tamaño y número de los nódulos, siempre que el estado general del paciente sea adecuado, no existan síntomas constitucionales relacionados con el CHC y exista una biopsia del nódulo de mayor tamaño que demuestre que el tumor no es pobremente diferenciado. La base de esta política de trasplante es la discrepancia entre los datos aportados por las pruebas radiológicas pretrasplante y el análisis anatomopatológico de los hígados explantados: en este estudio hasta un 30 % de pacientes caracterizados como “dentro de Milán” por las técnicas de imagen previas al trasplante excedían los criterios de Milán, mientras que el 23 % de los pacientes caracterizados como “fuera de Milán” realmente sí cumplían estos criterios.

La utilización de estos criterios expandidos no supone un peor pronóstico comparativamente con los criterios de Milán, con resultados equiparables en términos supervivencia global y libre de enfermedad a 5 años (70 % y 66 %, respectivamente).

Como hemos visto, la mayoría de estos criterios se basan exclusivamente en el número y tamaño tumoral (bien de los nódulos analizados en los hígados explantados o bien de los estudios de imagen pretrasplante). Estos límites morfológicos se han ido incrementando progresivamente desde unos criterios iniciales (Milán) que hoy en día pueden llegar a ser considerados excesivamente restrictivos. Sin embargo, siguen siendo la base de la mayoría de las políticas de trasplante en todo el mundo. Sólo los criterios de Toronto han obviado la morfología tumoral como criterio de trasplante y se basan exclusivamente en factores anatomopatológicos y el estado general del paciente. A modo de resumen, se incluye la siguiente tabla que recoge los citados criterios:

Milán*	- 1 tumor ≤ 5 cm - 2 ó 3 tumores cada uno ≤ 3 cm
Up-to-seven*	- Suma total del diámetro del tumor más grande y del número de tumores ≤ 7
San Francisco (UCSF)*	- 1 tumor $\leq 6,5$ cm - 2 ó 3 tumores cada uno $\leq 4,5$ cm - Suma total del diámetro tumoral ≤ 8 cm
Navarra*	- 1 tumor ≤ 6 cm - 2 ó 3 tumores cada uno ≤ 5 cm
Valencia*	- 1 tumor ≤ 5 cm - 2 ó 3 tumores cada uno ≤ 5 cm - Suma total del diámetro tumoral ≤ 10 cm
Seúl*	- Hasta 6 tumores con diámetro del tumor más grande ≤ 5 cm
Toronto*	- Cualquier número de nódulos y cualquier tamaño - Ausencia de síntomas sistémicos - Biopsia del tumor mayor: NO pobremente diferenciado

* Sin evidencia de invasión macrovascular y de diseminación extrahepática

Tabla 01: Criterios de indicación de trasplante en el carcinoma hepatocelular

Además de las indicaciones de trasplante con criterios *de novo*, en los últimos años se ha propuesto la estrategia de incluir en lista de espera de TH a aquellos pacientes con CHC resecados y que presentan factores anatomopatológicos de mal pronóstico (invasión microvascular, satelitosis...). Los estudios realizados confirman que estos pacientes tienen una supervivencia global y libre de enfermedad comparable a los pacientes incluidos de entrada en lista de espera (16). Esta estrategia, conocida como trasplante *ab initio*, ha sido

validada en estudios prospectivos y el consenso de la Sociedad Española de Trasplante Hepático (SETH) del año 2018 la aceptó como válida para indicación de trasplante (17). Asimismo, algunos grupos de trasplante aceptan como indicación aquellos pacientes en los que se ha realizado una resección con márgenes microscópicos afectos (R1) o cuando se produce una recidiva tumoral tras la resección hepática, siempre que la carga tumoral no exceda los criterios considerados como aceptables.

Además, la práctica clínica y la comprensión del comportamiento tumoral en el manejo del CHC han ido evolucionando desde la publicación de los criterios de Milán, hace más de 25 años. En este sentido, la respuesta a los tratamientos locorregionales pretrasplante, datos anatomopatológicos como la invasión vascular o la utilización de biomarcadores (alfafetoproteína, des-gamma-carboxi protrombina...), se han aplicado con el fin de mejorar las herramientas para evaluar el pronóstico de los pacientes con CHC, tanto antes como después del TH.

1.1.2.- Tratamiento pretrasplante de los pacientes con carcinoma hepatocelular

Con la utilización de criterios expandidos se plantea un incremento en el número de pacientes que se pueden beneficiar del TH como tratamiento curativo. Sin embargo, la relativa escasez de órganos hace que se prolongue el tiempo en lista de espera. Esto

obliga a controlar estrechamente a los pacientes con CHC para evitar la progresión tumoral.

Si bien es cierto que no hay evidencia sobre qué intervalo de tiempo es el mejor para el control de estos pacientes, la mayoría de los grupos de trasplante siguen a los pacientes con CHC mediante tomografía computarizada (TC) o resonancia magnética (RM) y analítica con niveles de alfafetoproteína (AFP) en sangre cada 3 meses. En este sentido, los pacientes candidatos a TH se encuentran en riesgo de progresión tumoral mientras permanecen en lista de espera, lo que a su vez acarrea un riesgo de recurrencia y mortalidad relacionada con cáncer después del trasplante (18).

Una situación posible en pacientes con CHC que esperan un TH es la salida de lista de espera (*drop-out*), bien sea por progresión tumoral u otras causas médicas derivadas de la cirrosis o hipertensión portal. Por ello es de utilidad el empleo de distintos tipos de tratamientos locorregionales neoadyuvantes (previos al trasplante) con el fin de disminuir el riesgo de *drop-out*.

Los tratamientos neoadyuvantes se dividen en 2 tipos, en función de su objetivo:

- a) *Terapias puente*: Son aquellos tratamientos que reciben los pacientes en lista de espera que ya cumplían criterios de trasplante al ser incluidos, con el fin de evitar la progresión tumoral y el *drop-out*. La probabilidad global de salir de lista de espera por progresión tumoral, si no se recibe tratamiento, es del 15-30 %. En cambio, las terapias puente conllevan un

impacto favorable en la disminución de la tasa de salida de lista de espera que alcanza el 21 %, con pacientes que permanecen en lista de espera más de 9 meses (19).

- b) *Downstaging*: este término describe aquellas terapias destinadas a disminuir el tamaño tumoral en pacientes que inicialmente exceden los criterios de trasplante. El objetivo es reducir la carga tumoral hasta un tamaño dentro de los criterios aceptados. Si el paciente presenta una respuesta satisfactoria al tratamiento cumpliendo criterios de trasplante, la mayoría de los grupos suelen exigir un período de tiempo (6 meses) de estabilidad sin tratamiento antes de incluir al paciente en lista de espera para trasplante.

Múltiples estudios han demostrado las ventajas de este tipo de terapias neoadyuvantes, especialmente en la reducción del riesgo de *drop-out* de lista de espera (20). Los protocolos de tratamiento varían mucho en los distintos grupos de trasplante, pero el empleo de terapias como la ablación por radiofrecuencia o la quimioembolización transarterial es casi universal, especialmente en aquellos pacientes en los que se prevé un tiempo en lista de espera superior a 6 meses.

Además, la respuesta a tratamientos puente y de *downstaging* no solo reduce significativamente el porcentaje de *drop-out* en lista de espera, sino que tiene una influencia directa en los resultados de supervivencia postrasplante (21). Generalmente, una respuesta

satisfactoria (completa o parcial) a las terapias neoadyuvantes está relacionada con factores de buen pronóstico (ausencia de invasión microvascular y de satelitosis, tumor de bajo grado) (20). Por tanto, la respuesta a estos tratamientos tiene un importante papel en la predicción de la agresividad del tumor, traduciendo en cierta manera la biología molecular del tumor, de tal forma que constituye en sí mismo una forma de selección de pacientes para el TH (22,23).

En este sentido, el sistema *Barcelona-Clínic Liver Cancer* (BCLC) en su última actualización del año 2022 (24) ya incluye la recomendación de trasplantar a pacientes que han experimentado un tratamiento de *downstaging* satisfactorio.

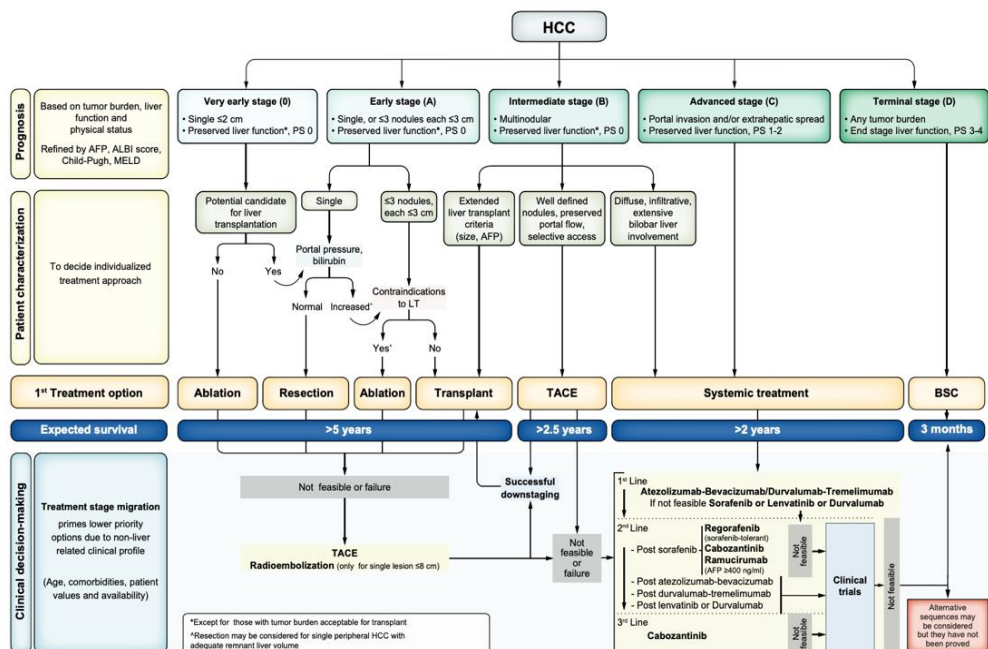


Figura 01: Clasificación Barcelona Clínic Liver Cancer en su versión de 2022 (24)

Los tratamientos radiológicos más utilizados son:

Radiofrecuencia (RF)

La ablación por RF es una de las técnicas más comúnmente empleadas como terapia puente al TH. En general, dentro de las técnicas ablativas se considera superior y de elección frente a la inyección de etanol, ya que en estudios comparativos ha demostrado obtener áreas de necrosis tumoral mayores. En hepatocarcinomas de diámetro inferior a 3 cm, la RF puede conseguir una necrosis tumoral completa y tiene resultados oncológicos comparables a la resección quirúrgica. Además, ha demostrado ser eficaz evitando el *drop-out* de pacientes en lista de espera para TH, con unas cifras que alcanzan el 6 % en series amplias (25,26). Sin embargo, en tumores de más de 3 cm de diámetro, la necrosis tumoral inducida por la RF puede llegar a ser inferior al 50 %, lo que hace desaconsejable el empleo de esta técnica en este escenario, o asociar un tratamiento complementario como es la quimioembolización transarterial (27).

Microondas (MO)

Comparada con la RF, la ablación mediante MO consigue una mayor elevación de la temperatura intratumoral y, por tanto, ablaciones más rápidas y de áreas más extensas. En este

sentido, la ablación mediante MO ha demostrado ser un procedimiento seguro y eficaz como tratamiento puente al TH, ya que consigue cifras similares de necrosis tumoral completa (28) sin diferencias en la recurrencia local o las complicaciones, por lo que en general se considera una alternativa en aquellos nódulos de localización menos favorable para la RF (proximidad de la lesión a vasos o a estructuras como vesícula, estómago, colon u otras vísceras).

Quimioembolización transarterial (QETA)

La QETA es una de las modalidades terapéuticas más empleadas como terapia puente y herramienta de *downstaging* en aquellos pacientes con CHC que inicialmente exceden los criterios de TH. Si bien es cierto que no existen estudios aleatorizados que comparen pacientes tratados con QETA frente a no tratados en lista de espera para TH, la QETA se ha mostrado como una técnica segura en este tipo de pacientes, con unos resultados de *drop-out* aceptables, que oscilan entre el 3 % y el 13 %, según las series (29–31).

Asimismo, los estudios que analizan el papel de la QETA como tratamiento para el *downstaging* adolecen de una gran heterogeneidad en cuanto a los pacientes analizados, y existen escasos estudios prospectivos al respecto. No obstante los datos disponibles sugieren que aquellos pacientes en los que se ha conseguido un *downstaging* alcanzando los criterios de Milán

tienen una supervivencia postrasplante y unas cifras de recurrencia comparables a aquellos pacientes que inicialmente cumplían los criterios de Milán (20,31).

Radioembolización transarterial (RETA)

Si bien es cierto que la experiencia con la RETA sigue siendo limitada, ya existe literatura que sugiere que esta modalidad terapéutica tiene un papel relevante como herramienta para el *downstaging*. En este sentido, algunos estudios apuntan a una mayor proporción de pacientes que consiguen el *downstaging* entre aquellos que se tratan con RETA frente a QETA (32,33). Asimismo, la supervivencia postrasplante de los pacientes que consiguen un *downstaging* efectivo con la RETA es equiparable a la de otras técnicas (34). Aunque no existen estudios aleatorizados prospectivos que avalen estos resultados, en la última actualización del sistema BCLC se incluye la RETA como herramienta terapéutica de *dowstaging*, especialmente en pacientes con nódulos únicos, grandes, de hasta 8 cm de diámetro y pacientes no candidatos a QETA (24).

1.1.3.- Evaluación de la respuesta tumoral al tratamiento pretrasplante del carcinoma hepatocelular

La respuesta tumoral al tratamiento se evalúa mediante los criterios RECIST (*Response Evaluation Criteria In Solid Tumours*), publicados en el año 2000 (35) por el *National Cancer Institute* (NCI) de los Estados Unidos, el *National Cancer Institute of Canada-Clinical Trials Group* (NCIC-CTG) y la *European Organisation for Research and Treatment of Cancer* (EORTC), con el objetivo de unificar los criterios existentes hasta ese momento.

Este modelo se basa en medir radiológicamente el tamaño del tumor viable, clasificando la respuesta en 4 categorías:

Respuesta completa (RC)	Desaparición de todas las lesiones
Respuesta parcial (RP)	Descenso ≥ 30 % de la suma de los diámetros máximos de las lesiones
Enfermedad Estable (EE)	No respuesta parcial ni progresión de enfermedad
Progresión de Enfermedad (PE)	Aumento ≥ 20 % de la suma de los diámetros máximos de las lesiones Aparición de nuevas lesiones.

Tabla 02: Criterios RECIST

Los criterios RECIST fueron diseñados inicialmente para agentes citotóxicos sistémicos y, por tanto, evalúan únicamente el tamaño unidimensional del tumor, por lo que no tienen en cuenta los posibles cambios necróticos en la zona tumoral ni en la viabilidad del tumor. Además, el propio documento inicial RECIST ya tiene en cuenta la necesidad de hacer modificaciones específicas para determinados tratamientos (por ejemplo, las terapias locorregionales) o para tumores que tengan un comportamiento radiológico particularmente complejo (como el CHC). En este sentido, los criterios RECIST no son óptimos para valorar la respuesta tumoral después del tratamiento pretrasplante del CHC. Además, diversos estudios han demostrado una escasa correlación entre la respuesta evaluada por estos criterios y el beneficio clínico obtenido con tratamientos locorregionales o incluso sistémicos como el sorafenib (36).

En el año 2001, un grupo de expertos de la *European Association for the Study of the Liver* (EASL) modificó los criterios de respuesta para que tuviesen en cuenta la necrosis tumoral inducida por el tratamiento (37). De forma subsiguiente, se consideró que el método óptimo para evaluar la respuesta del CHC al tratamiento pretrasplante es una estimación de la reducción del volumen del tumor viable, definida como la detección mediante técnicas de imagen dinámicas (TC o RM) de áreas sin captación de contraste. Estos conceptos de “respuesta” y “tumor viable” propuestos por la EASL han sido refrendados por la *American Association for the Study of Liver Diseases*, de tal forma que están incorporados en sus guías de práctica clínica desde 2005.

Sin embargo, dada la creciente complejidad de los estudios, las terapias disponibles y la necesidad de obtener un marco común de trabajo para poder objetivar los beneficios clínicos de los tratamientos disponibles, en 2010 se publicaron las guías que incluyen la modificación de la respuesta tumoral basada en los criterios RECIST con el objetivo de actualizar el concepto de “tumor viable”. Son los llamados criterios RECIST modificados (mRECIST) (38), que evalúan la respuesta tumoral al tratamiento pretrasplante de la siguiente manera:

Respuesta completa (RC)	Desaparición de cualquier realce o captación de contraste arterial en todas las “lesiones diana”
Respuesta parcial (RP)	Descenso ≥ 30 % de áreas con captación de contraste arterial en la suma de los diámetros de las “lesiones diana” viables
Enfermedad Estable (EE)	No respuesta parcial ni progresión de enfermedad
Progresión de Enfermedad (PE)	Aumento ≥ 20 % de áreas con captación de contraste arterial en la suma de los diámetros de las lesiones diana viables Aparición de nuevas lesiones.

Se consideran “lesiones diana” para ser evaluadas por los criterios mRECIST aquellas bien delimitadas y con captación de contraste arterial

Tabla 03: Criterios mRECIST (para “lesiones diana”)

En las guías de la EASL para el manejo del CHC del año 2018 se recomienda la utilización de estos criterios para evaluar la respuesta tumoral a los tratamientos locorregionales pretrasplante (39).

Asimismo, estos criterios sólo son aplicables con la utilización de pruebas de imagen dinámicas (TC y/o RM con contraste en varias fases). Por otra parte, es primordial que las imágenes sean evaluadas por radiólogos con experiencia y emplear, a ser posible, la misma técnica de imagen para que los resultados sean comparables (38).

Además, los criterios mRECIST también tienen en cuenta el comportamiento de las “lesiones no diana” y la aparición de nuevas lesiones a la hora de evaluar la respuesta tumoral.

Lesiones diana	Lesiones NO diana	Nuevas lesiones	Respuesta global
RC	RC	No	RC
RC	EE / Respuesta incompleta	No	RP
RP	No PE	No	RP
EE	No PE	No	EE
PE	Cualquiera	Sí o No	PE
Cualquiera	PE	Sí o No	PE
Cualquiera	Cualquiera	Sí	PE

Tabla 04: Criterios mRECIST (para “lesiones no diana”)

RC: Respuesta completa

RP: Respuesta parcial

EE: Enfermedad estable

PE: Progresión de enfermedad

Distintos estudios han demostrado que la respuesta al tratamiento, objetivada mediante los criterios mRECIST, predice la supervivencia de pacientes que han recibido tratamientos locorregionales (40).

1.2.- Factores pronósticos de recidiva del carcinoma hepatocelular tras el trasplante hepático

Múltiples estudios han evaluado la carga tumoral (número de nódulos y tamaño) como el principal marcador de mal pronóstico por el riesgo de recidiva del CHC. Sin embargo, en los últimos años se están identificando distintos parámetros basados en la biología molecular del CHC que pueden determinar su comportamiento e influir en el pronóstico del trasplante.

1.2.1.- Factores morfológicos

Aquellos tumores que cumplen los criterios de Milán (nódulo único de ≤ 5 cm de diámetro o 2-3 nódulos de ≤ 3 cm de diámetro, sin invasión macrovascular) son considerados de buen pronóstico (3), puesto que su supervivencia global a los 5 años del trasplante es similar a aquellos pacientes trasplantados sin CHC (70 %) (41) y con una recurrencia inferior al 10-15 %.

Asimismo, el tamaño tumoral se correlaciona con la presencia de invasión vascular y mal pronóstico. Aquellos tumores con menos de 3 cm de diámetro tienen menos probabilidad de presentar invasión microvascular. Por tanto, estos pacientes tienen menos riesgo de recidiva y una mayor supervivencia global y libre de enfermedad (41).

1.2.2.- Estadificación TNM

Los pacientes trasplantados con un TNM estadio III o superior, tienen un riesgo de recurrencia casi 3 veces superior a los pacientes trasplantados con TNM estadios I-II (42).

1.2.3.- Factores anatomopatológicos

a) Grado de diferenciación tumoral

Existen 4 grados de diferenciación tumoral, que tienen correlación directa con la probabilidad de recurrencia del CHC y la supervivencia después del trasplante (43).

Grado de diferenciación	Definición	Supervivencia a 5 años
Bien diferenciado	G1	67 %
Moderadamente diferenciado	G2	
Pobrementemente diferenciado	G3	30,5 %
Indiferenciado	G4	

Tabla 05: Grados de diferenciación tumoral y su correlación con la supervivencia global del CHC

b) Invasión vascular

Se definen 2 tipos de invasión vascular:

- Invasión macrovascular: presencia de infiltración tumoral en vasos de tamaño grande y mediano, visible en pruebas de imagen radiológicas.
- Invasión microvascular: presencia de células tumorales en el interior de vasos de pequeño tamaño. Esto es detectable únicamente con el análisis patológico de las piezas quirúrgicas

La invasión microvascular es un factor independiente de mal pronóstico para el TH por CHC, ya que disminuye la supervivencia global y libre de enfermedad. La presencia de invasión microvascular es más frecuente en tumores grandes (mayores de 3 cm), siendo infrecuente (menos de 3 %) en aquellos tumores con un tamaño inferior a 2 cm. Los tumores múltiples o con bajo grado de diferenciación (pobrementemente diferenciados o indiferenciados) también presentan un mayor riesgo de invasión microvascular (44).

c) Satelitosis

La satelitosis se define como la presencia de nódulos tumorales en las inmediaciones del nódulo principal, a menos de 2 cm de distancia del tumor primario y de tamaño inferior a la mitad del diámetro del nódulo principal. Su presencia es la consecuencia de la diseminación del tumor por invasión de vasos portales, por lo que es un factor de mal pronóstico

relacionado con la disminución de la supervivencia global y con el aumento de recurrencia (45).

1.2.4.- Alfafetoproteína

La AFP es un marcador sérico utilizado habitualmente en el diagnóstico del CHC. Además, está ampliamente aceptada su importancia como factor pronóstico independiente en la recurrencia tumoral después del TH. Así, se ha demostrado que la supervivencia libre de enfermedad en aquellos pacientes con una AFP > 1000 ng/ml es significativamente inferior a pacientes con AFP ≤ 1000 ng/ml (53 % vs. 80 %) (46).

En este sentido, se han desarrollado diversos modelos que incluyen los valores de AFP como factor pronóstico de la supervivencia global y libre de enfermedad después del trasplante hepático. Entre ellos, el llamado Modelo francés (47) es el más comúnmente utilizado. Publicado en 2012, utiliza un sistema de puntuación basado en el diámetro de las lesiones, el número de nódulos tumorales y el valor de la AFP. Aquellos pacientes con una puntuación ≤ 2 tienen una menor probabilidad de recurrencia tumoral (8,8 % vs. 50,6 %) y una mayor supervivencia global (67,8 % vs. 47,5 %) a 5 años frente a pacientes con una puntuación > 2.

En el año 2019, la Sociedad Española de Trasplante Hepático (SETH) reconocía la importancia de la AFP como factor pronóstico después del trasplante hepático en su documento de la 8ª Reunión de

Consenso. En dicho documento se recoge la recomendación de contraindicar formalmente el TH en pacientes con niveles de AFP > 1000 ng/ml (nivel 1A). Asimismo, se recomienda que en pacientes dentro de programas con criterios expandidos y con valores de AFP $\geq$ 400 ng/ml, se realice un tratamiento locorregional sobre el tumor y una reestadificación completa antes de ser considerados candidatos adecuados a TH (17).

No sólo los valores absolutos de AFP se han empleado como factores predictivos de la recurrencia del CHC después del TH, sino también las variaciones en los mismos. Así, el *score* NewYork/California (48) incorpora la variación entre el valor máximo y el valor pretrasplante de AFP durante el tiempo que permanece el paciente en lista de espera. Dicha variación se ha identificado como un factor independiente de buen pronóstico después del TH, incluso en aquellos pacientes con criterios expandidos. Además, este *score* ha sido validado externamente en estudios recientes (48), lo que subraya el papel del comportamiento dinámico de la AFP como factor predictivo de la supervivencia libre de enfermedad postrasplante.

1.2.5.- Respuesta al tratamiento pretrasplante

La respuesta del CHC al tratamiento locorregional pretrasplante se relaciona con la supervivencia global y libre de enfermedad de los pacientes después del TH, de tal manera que dicha respuesta se considera una información fiable y fácilmente aplicable en la práctica

clínica que puede reflejar mejor la biología tumoral y el comportamiento más o menos agresivo del tumor (49,50)

De hecho, habida cuenta de la relación que tiene con la supervivencia global y la recidiva postrasplante, determinadas publicaciones sugieren que la respuesta al tratamiento locorregional pretrasplante valorada con pruebas de imagen podría ser empleada como criterio de selección de pacientes para inclusión en lista de espera de TH (51). Esta relación se evidencia en estudios con pacientes que cumplen los criterios de Milán (52,53), con diferencias estadísticamente significativas de supervivencia global a 5 años después del trasplante entre aquellos pacientes que mostraron respuesta completa o parcial (85,1 % y 63,9 %, respectivamente) y aquellos que mostraron enfermedad estable (51,4 %). Asimismo, otros estudios han demostrado que aquellos pacientes con progresión de enfermedad a pesar de haber recibido tratamientos locorregionales tienen un riesgo aumentado de recurrencia tumoral (HR 5,62, IC 95 % 4,10 – 7,69) (54).

Además, el grado de necrosis tumoral inducido por el tratamiento pretrasplante también se correlaciona con el pronóstico de los pacientes después del trasplante. Se ha demostrado que los pacientes con respuesta patológica completa en los hígados explantados tienen una supervivencia global y libre de enfermedad superior (55); y aquellos pacientes con necrosis tumoral significativa (igual o superior al 60 %) tienen una supervivencia libre de enfermedad mejor que aquellos con peor respuesta al tratamiento locorregional, independientemente de

que se incluyan dentro de los criterios de Milán o hayan requerido *downstaging* (56). En este sentido, el número de tratamientos necesarios para obtener una respuesta del tumor también podría ser un indicador fiable de la biología tumoral, siendo los tumores más agresivos y con más riesgo de recurrencia aquellos que requieren de un mayor número de sesiones para obtener una respuesta completa o parcial (57).

Así pues, aunque no hay estudios prospectivos comparativos que permitan establecer conclusiones más sólidas, la respuesta del CHC al tratamiento locorregional pretrasplante se considera factor de buen pronóstico para la supervivencia global y libre de enfermedad (58).

1.2.6.- Serología del receptor

Existen estudios que apuntan a que los pacientes con CHC y serología positiva para VHC tienen una supervivencia libre de enfermedad inferior a aquellos no infectados (59). Esto se debe a que, en ausencia de tratamiento, la reinfección del injerto hepático era casi universal, asociando además un desarrollo de cirrosis más rápido en el injerto (30 % a los 5 años) con respecto a pacientes no inmunodeprimidos (60).

Sin embargo, la aparición de los nuevos antivirales de acción directa (AAD) ha cambiado radicalmente el panorama de la infección por VHC y, debido a sus buenos resultados, estas terapias se han incorporado rápidamente al tratamiento del VHC tanto en pacientes en

lista de espera de trasplante como en aquellos previamente trasplantados por cirrosis VHC (61).

Paralelamente, distintos estudios han señalado una elevación en la recurrencia de CHC en pacientes tratados con AAD y que ya se habían sometido a resecciones quirúrgicas o tratamientos locorregionales con respuesta completa del CHC (62), con recidivas superiores al 20%, con una mediana de seguimiento de menos de un año después del inicio del tratamiento con AAD. Otros grupos han comunicado resultados de recurrencia similares (alrededor del 20%) en pacientes trasplantados por CHC (63,64). Si estos fármacos asocian un riesgo incrementado de recurrencia de CHC o no, constituye un tema controvertido, puesto que otros grupos han presentado resultados contrarios que no demuestran asociación entre los AAD y riesgo de recidiva de CHC (65). Este tema, así como el mecanismo que explicaría la recurrencia tumoral en pacientes tratados con estos fármacos, continúa siendo objeto de debate e investigación y se necesitan más estudios para extraer conclusiones significativas.

1.2.7.- Factores de biología molecular

Determinados estudios han apuntado que la expresión de ciertos genes podría constituir un factor pronóstico de la recurrencia del CHC después del trasplante. Así, en 2015 un estudio sugirió una posible relación entre la expresión de citoqueratina 19 en el tejido tumoral y el riesgo de recurrencia (66).

Existen también estudios que relacionan una mayor incidencia de invasión microvascular con una sobreexpresión del gen p53 en tejido tumoral, por lo que estos pacientes tendrían asimismo un mayor riesgo de recidiva del CHC después del trasplante (67).

Otros marcadores que se han relacionado con una mayor recurrencia tumoral son el factor de crecimiento del endotelio vascular, los niveles séricos preoperatorios de interleuquina-10 o des-gamma-carboxi protrombina (DCP) (68) y la expresión en tejido tumoral de ácido ribonucleico mensajero de gamma-glutamyl transpeptidasa.

La DCP en concreto, es una variante anormal de la protrombina que se correlaciona con un mal pronóstico (tumores poco diferenciados, invasión vascular, recurrencia tumoral...) en pacientes con CHC, y que se ha incorporado a distintos índices y sistemas pronósticos, como los criterios de Kyoto (68,69).

1.2.8.- Estatus inflamatorio pretrasplante

El estado inflamatorio sistémico del paciente previo al trasplante ha ganado importancia como factor pronóstico en distintos tipos de tumores, incluyendo el CHC. En concreto la *ratio* de neutrófilos / linfocitos (RNL) ha demostrado ser un factor de riesgo para la recurrencia postrasplante del CHC, incluso en pacientes que cumplen los criterios de Milán. Así, un estudio de 2009 demostró que aquellos pacientes con un $RNL \geq 5$ tienen una incidencia mayor de recurrencia del CHC después del trasplante (62 % vs. 14 %, $p < 0,001$), así como

una menor supervivencia global y libre de enfermedad (70). Este mismo parámetro ha sido validado por otros grupos y se ha incluido junto con otros factores con el fin de construir modelos predictivos de recurrencia del CHC después del trasplante hepático (57).

La *ratio* de plaquetas / linfocitos (RPL) también se ha postulado como un factor pronóstico de recurrencia del CHC después del trasplante. En concreto, una $RPL > 125$ se ha vinculado con tumores de mayor tamaño y mayor incidencia de invasión micro y macrovascular, factores todos ellos asociados con mayor riesgo de recidiva del hepatocarcinoma (71).

El mecanismo que explica que la RNL y el recuento linfocitario disminuido se relacionen con un mal pronóstico a largo plazo de CHC no está completamente determinado. Se ha postulado que la depleción linfocitaria provoca una disminución de la producción de citoquinas antitumorales, lo que conllevaría una alteración de la respuesta inmunitaria frente al CHC (72). Por otra parte, la neutrofilia sistémica se ha relacionado con una mayor infiltración de macrófagos y con una mayor producción de citoquinas angiogénicas, favoreciendo por tanto el crecimiento tumoral (73).

1.2.9.- Scores y criterios pronósticos combinados

La publicación de los criterios de Milán en el año 1996 supuso un hito en cuanto al potencial del trasplante hepático como tratamiento capaz de ofrecer una solución curativa en términos oncológicos para

los pacientes con CHC. La adecuada selección de pacientes permitía obtener unos resultados de supervivencia equiparables a indicaciones no oncológicas. Ese éxito ha supuesto que los criterios de Milán sigan siendo el pilar fundamental de las políticas de inclusión en lista de espera de trasplante en la mayoría de los grupos de TH del mundo. Sin embargo, a lo largo de estos 25 años la práctica clínica ha ido evolucionando en paralelo a la comprensión del comportamiento tumoral del CHC, fundamentalmente con la introducción de las terapias locorregionales previas al trasplante y la utilización de biomarcadores (principalmente AFP) que han demostrado su utilidad como herramientas pronósticas. Todo ello ha llevado a que, en el campo de la *transplant oncology*, hayan surgido nuevos criterios para la valoración pronóstica y la toma de decisiones en cuanto a la indicación de TH que no tienen en cuenta factores exclusivamente relacionados con la morfología del tumor (tamaño y número de nódulos). Se trata de criterios y *scores* combinados que tratan de complementar la información quizá excesivamente restringida que aportan los criterios morfológicos. Además, son criterios expandidos en su mayoría, cuyo objetivo es incrementar el número de pacientes que pueden beneficiarse del TH como opción curativa para el CHC, sobre la base de que los datos sobre la morfología del tumor, de forma aislada, no suponen una herramienta lo suficientemente precisa para predecir el comportamiento biológico del CHC.

Modelo francés (AFP-model)

Como hemos comentado anteriormente, este modelo fue publicado en 2012 y fue el primer sistema que incluyó la AFP como criterio pronóstico. El modelo fue creado en base a una cohorte de entrenamiento de 597 pacientes trasplantados por CHC en 16 centros de Francia entre 1988 y 2001, y posteriormente fue validado con un grupo de 474 pacientes trasplantados entre 2003 y 2004 en otros 21 centros del mismo país.

El modelo establece una puntuación que considera el tamaño del nódulo mayor, el número de nódulos y los niveles de AFP. En función de la puntuación obtenida, el modelo clasifica a los pacientes en dos grupos: bajo riesgo (puntuación ≤ 2) y alto riesgo (puntuación > 2). Los pacientes considerados de bajo riesgo tienen una menor probabilidad de recurrencia tumoral (8,8 % vs. 50,6 %) y una mayor supervivencia global a 5 años (67,8 % vs. 47,5 %) frente a los pacientes de alto riesgo. Los resultados de este estudio han sido validados externamente en estudios posteriores (74), demostrando su capacidad de generalización. Poco después de la publicación de este modelo, Francia lo adoptó como criterio de selección de pacientes con CHC para TH.

Metroticket 2.0

Se trata de una actualización de los criterios *Up-to-seven* (*Metroticket*) publicados por el grupo de Milán en el año 2018 (75). Se basa en un modelo de regresión de riesgo competitivo que analizó 1018 pacientes de 3 centros de Italia, posteriormente validados con 341 pacientes de los mismos hospitales. Este estudio demostró que la suma del número total de nódulos tumorales y del tamaño del nódulo mayor, junto con los niveles de AFP, permitían predecir la supervivencia global y libre de enfermedad de los pacientes. Se trata de una herramienta que no ofrece tanto una decisión dicotómica sí/no con respecto a la selección de pacientes para TH, sino que determina un umbral a partir del cual los pacientes pueden tener una supervivencia global a 5 años superior al 70 %. Así, aquellos pacientes que consiguen estos resultados son los que cumplían las siguientes condiciones:

- Suma de nódulo mayor (en cm) + número de nódulos < 7 y AFP < 200 ng/ml
- Suma de nódulo mayor (en cm) + número de nódulos ≤ 5 y AFP 200-400 ng/ml
- Suma de nódulo mayor (en cm) + número de nódulos ≤ 4 y AFP 400-1000 ng/ml

Además, este estudio demostró que el sistema Metroticket 2.0 era superior en términos de capacidad pronóstica a los

criterios de Milán, UCSF, *Up-to-seven* y el modelo francés de AFP.

En el año 2020, un estudio demostró que la incorporación de la respuesta al tratamiento locorregional en base a los criterios mRECIST junto con el sistema Metroticket 2.0 mejora la capacidad predictiva de este sistema (76).

TRAIN score

Este índice pronóstico utiliza la respuesta radiológica al tratamiento locorregional pretrasplante según los criterios mRECIST, los cambios en los niveles de AFP y la RNL como marcador inflamatorio, además del tiempo en lista de espera. Se publicó en el año 2016 (77) sobre una muestra unicéntrica de 179 pacientes y posteriormente fue validado con otra muestra de 110 pacientes de otro centro diferente. El TRAIN score combina los factores anteriormente citados mediante la siguiente fórmula:

$$0,988 \text{ (si mRECIST es PE)} + 0,838 \text{ (si descenso de AFP es } \geq 15,0 \text{ ng/ml/mes)} + 0,452 \text{ (si RNL } \geq 5,0) - 0,03 \times \text{ tiempo en LE (meses)}$$

PE: progresión de enfermedad, LE: lista de espera

Aquellos pacientes con una puntuación en el TRAIN < 1,0 tienen una supervivencia global a 5 años superior a los

pacientes con puntuación $\geq 1,0$ (67,5 % vs. 20,7 %) así como una recurrencia a 5 años inferior (8,9 % vs. 30 %). Además, se comparó la capacidad predictiva de este índice con los criterios de Milán, obteniendo resultados similares. Con este sistema se estimó que se permitiría incrementar hasta en un 24 % el número de pacientes potencialmente trasplantables, en comparación con los criterios de Milán.

HALT-HCC score

Se trata de un índice de puntuación pronóstico continuo más que unos criterios categóricos de inclusión o exclusión. Publicado en 2017 (78), el sistema HALT-HCC (*Hazard Associated with Liver Transplantation for Hepatocellular carcinoma*) utiliza información acerca de la carga tumoral del paciente, los niveles de AFP y la puntuación MELD (*Model for End-stage Liver Disease*) según la siguiente ecuación:

$$(1,27 \times TBS) + (1,85 \times \ln AFP) + (0,26 \times MELD-Na)$$

TBS: Tumor Burden Score (índice de carga tumoral del paciente)

$\ln$ AFP: logaritmo neperiano de la AFP

La puntuación de este índice varía entre 2,40 y 46,42 puntos, y los pacientes del estudio fueron divididos en 4 cuartiles, de tal

forma que la supervivencia global a 5 años empeoraba significativamente a medida que aumentaba la puntuación del HALT-HCC (78,7 % para el primer cuartil, 74,5 % para el segundo, 71,8 % para el tercero y 61,5 % para el último cuartil). Asimismo, un 30 % de pacientes trasplantados fuera de los criterios de Milán con una puntuación HALT-HCC baja tenía un pronóstico similar a pacientes trasplantados dentro de los criterios de Milán; del mismo modo que un 21 % de pacientes trasplantados intraMilán pero con una elevada puntuación HALT-HCC tenía un pronóstico más desfavorable. En un estudio posterior desarrollado con la intención de recalibrar y mejorar este *score*, realizado con 4089 pacientes de 16 centros diferentes, el sistema HALT-HCC demostró una capacidad predictiva de la recurrencia postrasplante superior a otros índices pronósticos como el Metroticket 2.0, así como una adecuada capacidad para predecir factores anatomopatológicos de mal pronóstico en los hígados explantados (invasión microvascular y grado de diferenciación tumoral) (78).

MORAL score

Se trata de un sistema pronóstico que otorga una puntuación en función de 3 factores, conocidos previamente al trasplante: la RNL (≥ 5 , 6 puntos), los niveles máximos de AFP (> 200

ng/ml, 4 puntos) y el diámetro en centímetros del mayor nódulo tumoral (> 3 cm, 3 puntos).

La suma de los 3 ítems corresponde a la puntuación total del sistema MORAL (*Model Of Recurrence After Liver Transplant*). Los pacientes se dividen en 4 grupos según la supervivencia libre de enfermedad a 5 años: bajo riesgo (0-2 puntos, 98,6 %); riesgo medio (3-6 puntos; 69,8 %); alto riesgo (7-10 puntos; 55,8 %). El grupo de riesgo muy alto, correspondiente a los pacientes con una puntuación > 10, tiene una recurrencia al año del trasplante del 80 % y una SLE a un año del 17,9 % (79).

Criterios de Kyoto

Publicados por primera vez en el año 2007 (69), utiliza una combinación de datos morfológicos tumorales con marcadores biológicos, en este caso la DCP. Con este criterio combinado, los pacientes con un número de lesiones inferior a 11 (siempre que ninguna lesión sea mayor de 5 cm) y con unos niveles pretrasplante de DCP < 400 mAU/mL, tienen una recurrencia y una SG a 5 años comparable a los criterios de Milán (5 % y 87 %, respectivamente). Aquellos pacientes que exceden estos criterios, sin embargo, tienen un pronóstico significativamente peor, con una recurrencia y una SG a 5 años de 60,5 % y 34,4 %, respectivamente.

Estos criterios han sido posteriormente validados con un subgrupo de pacientes del mismo centro (80), sin embargo, es importante destacar que tanto en el estudio inicial como en el subgrupo de validación, estos criterios han sido empleados únicamente para trasplante hepático con donantes vivos.

A continuación, se ofrece una tabla resumen sobre los distintos criterios y *scores* pronósticos combinados analizados anteriormente:

Criterio / Score	Información	Punto de corte
<i>AFP-model</i>	Diámetro del nódulo mayor (< 3 cm = 0, 3-6 cm = 1, > 6 cm = 4 puntos) + número de lesiones (1-3 = 0, > 4 = 2 puntos) + AFP (< 100 = 0, 100-1000 = 2, > 1000 = 3 puntos)	Suma de los puntos menor o igual a 2
<i>Metrotick 2.0</i>	Número de nódulos + tamaño del nódulo mayor (cm) < 7 y AFP < 200; o número de nódulos + tamaño del nódulo mayor < 5 y AFP < 400; o número de nódulos + tamaño del nódulo mayor < 4 y AFP < 1000	Cumple criterios (sí / no)
<i>TRAIN</i>	0,988 (si mRECIST es PE) + 0,838 (si descenso de AFP > 15 ng/ml/mes) + 0,452 (si RNL > 5) – (0,03 x meses en LE)	Suma de los puntos menor de 1.
<i>HALT-HCC</i>	(1,27 x TBS) + (1,85 x lnAFP) + (0,26 x MELD-Na)	Score continuo
<i>MORAL</i>	3 (si nódulo más grande > 3 cm) + 6 (si RNL > 5) + 4 (si AFP > 200 ng/ml)	0-2: bajo riesgo 3-6: riesgo intermedio 7-10: riesgo alto 11-13: riesgo muy alto
<i>Criterios de Kyoto</i>	Número total de nódulos < 11, siempre que ninguno sea > 5 cm + DCP < 400 mAU/mL	Cumple criterios (sí / no)

PE: progresión de enfermedad

TBS: *Tumor Burden Score*

lnAFP: logaritmo neperiano de AFP

DCP: Des-gamma-carboxi-protrombina

Tabla 06: Índices y criterios pronósticos combinados

Como hemos visto, en los últimos años se han desarrollado multitud de herramientas, criterios e índices pronósticos con el objetivo de ofrecer la opción del TH a un mayor número de pacientes sin la limitación que pueden suponer los criterios morfológicos, incluyendo otros marcadores predictivos, biológicos y relativos al comportamiento del tumor. En este sentido, con una ingente cantidad de criterios disponibles, la cuestión sigue siendo cuál es el mejor criterio para elegir a los candidatos a TH, ya que a día de hoy no existe ningún estudio cuyo objetivo principal sea la comparación de estos criterios entre sí.

1.3.- Modelos matemáticos de predicción

1.3.1.- Características de un modelo predictivo

Un modelo predictivo es una representación matemática de un evento que permite predecir comportamientos futuros en función de la información disponible en el presente. Si aplicamos un ejemplo al tema que nos ocupa, una vez conocidas las variables que influyen en un resultado concreto (recurrencia precoz del CHC) de un procedimiento (trasplante hepático) se podría realizar una aproximación fiable acerca del resultado que podemos obtener en el siguiente caso, conociendo de antemano los valores de dichas variables relevantes.

Todo modelo predictivo que pretenda utilizarse como prueba diagnóstica debe poder cuantificarse, es decir, debe poder establecerse un umbral o punto de corte que permita establecer su resultado en positivo o negativo. Asimismo, el modelo debe presentar una serie de características básicas (81):

- Validez: grado en que una prueba consigue identificar aquello que desea evaluar.
- Reproducibilidad: capacidad de una prueba para ofrecer los mismos resultados al repetirse en condiciones similares.
- Seguridad: capacidad de una prueba de acertar en la predicción que ha realizado. Está determinada por el valor predictivo de un resultado positivo o negativo.

El análisis de la información que ofrece un modelo predictivo debe hacerse de forma multivariante, considerando todas aquellas variables implicadas en el problema estudiado, con el fin de conseguir que el modelo clasifique de forma adecuada los nuevos casos. De este modo, se consigue generar una prueba o modelo predictivo que pueda aplicarse a las distintas situaciones posibles. Tomamos como ejemplo los escenarios que aparecen en la siguiente matriz de confusión:

		<u>Estado real</u>		TOTAL
		Recidivado (E+)	No Recidivado (E-)	
<u>Resultado test</u>	Positivo (T+)	a (VP)	b (FP)	Test positivo: a+
	Negativo (T-)	c (FN)	d (VN)	Test negativo: c+
TOTAL		Total recidivados: a + c	Total no recidivados: b + d	

T: Test, E: Estado real, VP: Verdadero positivo, VN: Verdadero negativo,

FP: Falso positivo, FN: Falso negativo

Figura 02: Clasificación de las situaciones resultantes tras la aplicación de un modelo predictivo en forma de matriz de confusión

La precisión de un modelo predictivo se define como la capacidad que presenta para modificar los resultados influidos por el azar. Para evaluarla, se utilizan dos medidas:

- Sensibilidad (S) o proporción de verdaderos positivos identificados por la prueba o modelo en el total de casos en que se produce el evento estudiado.

$$\text{Sensibilidad} \equiv S = \frac{a}{a+c}$$

- Especificidad (E) o proporción de verdaderos negativos identificados por la prueba o modelo en el total de casos en que no se produce el evento estudiado.

$$\text{Especificidad} \equiv E = \frac{d}{b+d}$$

Los valores más adecuados de sensibilidad y especificidad de una prueba predictiva dependen de cada situación concreta que se estudie. No obstante, en general se considera que una prueba o modelo presenta una validez aceptable si su sensibilidad y especificidad se encuentran por encima del 80 % (82).

Aunque la sensibilidad y la especificidad se emplean para analizar la calidad del modelo, en la práctica clínica, una buena prueba predictiva para determinar, por ejemplo, la recurrencia de un tumor, es la que ofrece resultados positivos en aquellos casos que recurren y negativos en los que no hay recurrencia tumoral, es decir, si un sujeto ha resultado positivo tras aplicar el modelo, ¿cuál es la probabilidad de que tenga una recidiva tumoral?

O por el contrario, si el sujeto resultó negativo en la prueba ¿cuál es la probabilidad de que no haya una recurrencia?

Estas dos probabilidades posprueba se pueden calcular aplicando el teorema de Bayes, siempre con la condición de que conozcamos la probabilidad de que el sujeto tenga una recurrencia tumoral antes de realizar la prueba (probabilidad preprueba). En este sentido, si llamamos P a la probabilidad preprueba de presentar una recurrencia tumoral, S a la sensibilidad de la prueba y E a su especificidad, tenemos las siguientes fórmulas para calcular la probabilidad posprueba (Pr) de que el sujeto tenga una recurrencia cuando la prueba resultó positiva o de que esté libre de enfermedad cuando resultó negativa:

$$Pr(E+/T+) = \frac{S \cdot P}{S \cdot P + (1 - E) \cdot (1 - P)}$$

$$Pr(E-/T-) = \frac{E \cdot (1 - P)}{E \cdot (1 - P) + (1 - S) \cdot P}$$

$Pr(E+/T+)$ (donde $E+$ es recurrencia tumoral y $T+$ resultado positivo de la prueba) determina la proporción de sujetos que presentan recurrencia tumoral, de entre todos aquellos individuos en los que la prueba resultó positiva. Este valor se conoce como Valor Predictivo Positivo ($VP+$ o VPP):

$$VP+ = \frac{a}{a+b}$$

Pr(E-/T-) (donde E- es ausencia de recurrencia tumoral y T- resultado negativo de la prueba) determina la proporción de sujetos sin recurrencia sobre el total de aquellos individuos en los que la prueba resultó negativa. Este valor recibe el nombre de Valor Predictivo Negativo (VP- o VPN).

$$VP- = \frac{d}{c+d}$$

Para conocer la capacidad discriminatoria de un modelo predictivo, una vez conocidos estos parámetros se utilizan las denominadas curvas ROC (*Receiver Operating Characteristic*), que suponen una medida independiente y global de la utilidad de una prueba diagnóstica (83). Para generar una curva ROC, se representan gráficamente los valores de sensibilidad y especificidad obtenidos con cada uno de los valores observados, según se muestra en la siguiente figura:

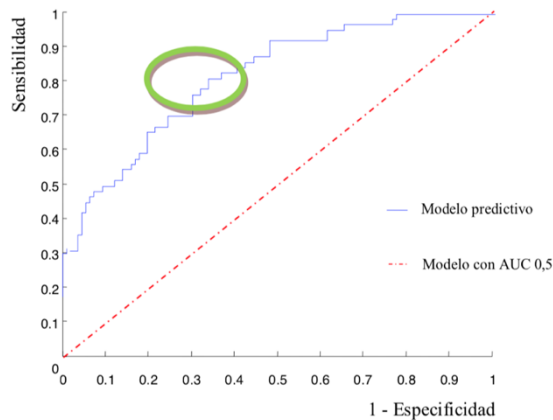


Figura 03: Representación gráfica de una curva ROC

Considerando una situación ideal, un modelo que predijese perfectamente la situación estudiada se representaría gráficamente con una línea que coincidiría al completo con el ángulo correspondiente a la esquina superior izquierda de la gráfica. Por otra parte, una prueba sin capacidad de discriminación en absoluto se representaría gráficamente mediante una línea diagonal de 45° , desde el 0 hasta la esquina superior derecha (situación representada en la figura anterior por la línea discontinua roja). Es decir, cuanto más se aproxime una curva ROC a la esquina superior izquierda de la gráfica, tanto mejor es la capacidad discriminatoria de una prueba. En este sentido, una forma de evaluar de manera global e intuitiva esa capacidad de discriminación (probabilidad de estimar correctamente si un individuo va a presentar recurrencia precoz o no) consiste en calcular el área creada bajo la curva ROC (ABC).

El ABC se interpreta como la probabilidad de que el valor de la prueba para un paciente con recurrencia tumoral sea mayor que el valor de la prueba para un paciente sin recurrencia, y se incrementa con la capacidad predictiva de la prueba. Además, su interpretación es directa, esto es, un resultado de ABC de 0,85 significa que, si se eligen al azar un individuo con recurrencia precoz y otro libre de enfermedad, en el 85 % de las ocasiones, el valor de la prueba en el paciente con recurrencia precoz será mayor que en el que se halla libre de enfermedad.

Como se ha comentado anteriormente, en una situación en la que la prueba diagnóstica no es capaz de diferenciar dos grupos o resultados, dicha prueba tendrá una ABC de 0,5, de tal manera que su curva ROC coincidirá con la diagonal (ascendente de izquierda a derecha) del cuadrado. En el otro extremo, una prueba con una capacidad predictiva perfecta, tendrá una ABC de 1, de tal manera que su curva ROC coincidirá con la esquina superior izquierda del cuadrado de la gráfica.

Teniendo en cuenta su ABC, la capacidad predictiva de una prueba diagnóstica puede establecerse en las siguientes categorías (84):

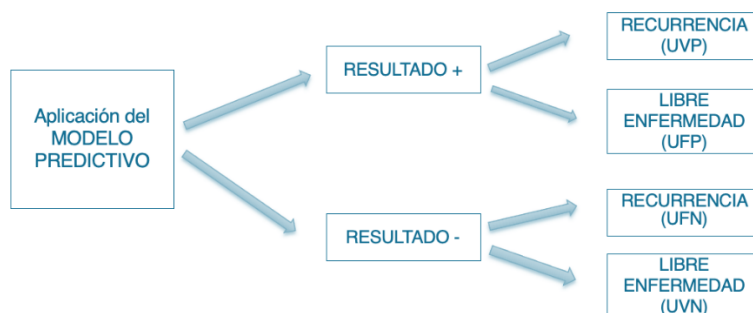
- Excelente: $ABC = 0,90-1$
- Buena: $ABC = 0,80-0,90$
- Regular: $ABC = 0,70-0,80$
- Mala: $ABC = 0,60-0,70$
- Pésima: $ABC = 0,50-0,60$

Con la curva ROC generada es más fácil determinar el punto de

corte (o valor de la prueba) que nos permite diferenciar el límite entre recurrencia y libre de enfermedad. Para ello, disponemos de varias opciones (85):

- El punto de la curva más cercano a la esquina superior izquierda, independientemente de los valores de sensibilidad y especificidad.
- El punto cuyo valor de la suma *sensibilidad + especificidad* sea el más alto.
- El punto en el que sensibilidad y especificidad tengan igual valor.
- El punto con un valor prefijado por el investigador de sensibilidad, especificidad o valor predictivo, según las características de la prueba y los datos que esta analiza, buscando elevar al máximo una característica llamada *función de utilidad*. Este método se basa en la importancia relativa que puede tener para el investigador el hecho de que la prueba diagnóstica haga una predicción verdadera o falsa.

En la siguiente figura se muestra un esquema de las situaciones posibles tras la aplicación de un modelo predictivo y la definición de las diferentes utilidades obtenidas:



UVP: Utilidad Verdadero Positivo / UFP: Utilidad Falso Positivo

UFN: Utilidad Falso Negativo / UVN: Utilidad Verdadero Negativo

Figura 04: Utilidad de un modelo predictivo

Asimismo, en función de la relación entre los valores de utilidad, surgen los conceptos de Coste neto de tratar individuos no enfermos y Beneficio neto de tratar sujetos enfermos:

- Coste neto (C): Representa la diferencia en términos de beneficio entre la Utilidad Verdadero Negativo (UVN) y la Utilidad Falso Positivo (UFP). En nuestro caso, el coste representaría la diferencia entre la utilidad de trasplantar a un receptor que estará vivo y libre de enfermedad en los 2 primeros años después del trasplante (UVN) y la utilidad de no trasplantar a una persona que estaría viva y libre de enfermedad después del trasplante (UFP). En términos matemáticos se representa como:

UVN-UFPP

- Beneficio neto (B): Representa la diferencia en términos de beneficio entre la Utilidad Verdadero Positivo (UVP) y la Utilidad Falso Negativo (UFN). En la situación que nos ocupa, el beneficio determina la diferencia entre la utilidad de no trasplantar a un receptor que tendrá una recurrencia del CHC en los 2 primeros años después del trasplante (UVP) y la utilidad de trasplantar a una persona que recidivaría en esos 2 primeros años tras el procedimiento (UFN). Se expresa como:

UVP-UFN

En términos globales, la utilidad de un modelo predictivo matemático se obtiene del promedio ponderado de las diferentes utilidades:

$$U = p(VP) UVP + p(FP) UFP + p(FN) UFN + p(VN) UVN$$

En este promedio, U es la utilidad esperada, p(VP) la probabilidad de los verdaderos positivos, UVP la utilidad de los verdaderos positivos, p(FP) la probabilidad de los falsos positivos, UFP la utilidad de los falsos positivos, p(FN) la probabilidad de los falsos negativos, UFN la utilidad de los falsos negativos, p(VN) la probabilidad de los verdaderos negativos y UVN la utilidad de los verdaderos negativos.

El valor máximo de utilidad de una prueba diagnóstica o modelo predictivo se relaciona con el coste y el beneficio de la prueba y con la prevalencia del evento analizado, así como con la pendiente de la curva ROC, según la fórmula que sigue:

$$Pendiente = \frac{C}{B} \times \frac{1-P}{P}$$

En esta relación, P (probabilidad preprueba) es la proporción de pacientes con recurrencia de CHC en los dos primeros años después del trasplante sobre el total de la población estudiada. Este valor permite determinar el punto de corte en función de la utilidad que busquemos para nuestra prueba o modelo predictivo, de tal manera que cada valor de pendiente de la curva ROC corresponde a un par de valores de sensibilidad y especificidad obtenidos para un umbral determinado. Dicho umbral de resultado de la prueba supondrá el punto de corte que distinguirá entre recurrencia en los 2 primeros años postrasplante y libre de enfermedad.

En resumen, la capacidad discriminatoria de una prueba o modelo predictivo la determina su ABC y su precisión se obtiene a partir de una pendiente determinada de la curva ROC. Esta pendiente se correlaciona con el valor que se otorga a cada una de las utilidades determinadas, según el problema en concreto que se analice y teniendo en cuenta la importancia clínica que dicho problema tiene. En este sentido, la calidad de un modelo que tiene como objetivo determinar el riesgo de recurrencia del CHC en los 2 primeros años

después de un trasplante hepático dependerá de los distintos valores de utilidad que se asignen para su implementación en la práctica clínica habitual, teniendo en cuenta el coste y el beneficio que se puede derivar de la cada una de las situaciones concretas que pueden darse después de la aplicación de los resultados del modelo o test.

1.3.2.- Modelos gráficos probabilísticos

Los modelos gráficos probabilísticos (MGP) son representaciones de probabilidad entre variables aleatorias relacionadas entre sí mediante un modelo gráfico (dependencias). A través de la representación gráfica, sintetizan las distribuciones de probabilidad que existen en un conjunto de datos. Esa representación o grafo está formada por dos elementos:

- Nodos (también llamados vértices)
- Arcos (también conocidos como flechas o aristas). Así, un arco entre un nodo A y un nodo B, refleja que el valor que toma la variable B depende del valor de A.

Los MGP más comúnmente empleados son los basados en grafos dirigidos acíclicos (GDA), en los que los arcos presentan dirección (flechas) y ningún nodo está conectado a sí mismo a través de dichas flechas, y los basados en grafos no dirigidos.

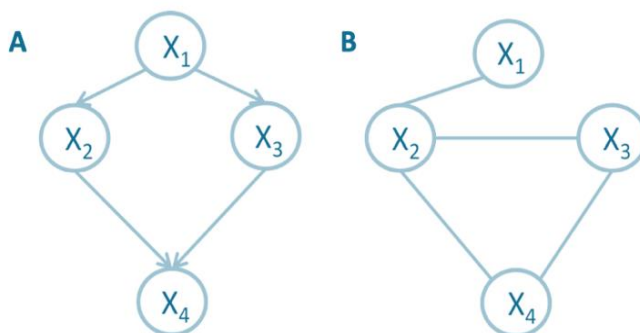


Figura 05: Ejemplos de grafo dirigido acíclico (A) y de grafo no dirigido (B) (86)

En los MGP, los nodos representan variables aleatorias y los arcos las dependencias probabilísticas entre ellas. Existe también un conjunto de términos para definir las relaciones entre los nodos: si hay un arco directo entre A y B, se dice que B es el nodo hijo de A, o bien que éste último es el nodo padre de B, como puede observarse en la siguiente figura:

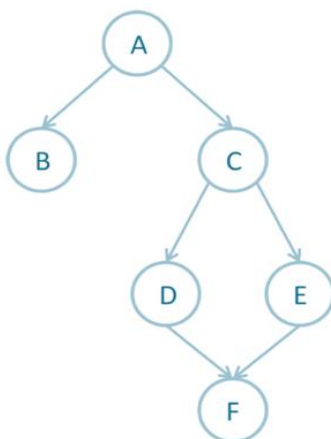


Figura 09: Representación de las dependencias probabilísticas entre variables aleatorias, con nodos “padres” e “hijos” (86)

Asimismo, los MGP presentan distintas características que los convierten en herramientas útiles para la investigación:

- Permiten visualizar de manera simple la estructura de un modelo probabilístico.
- Ciertos cálculos, que podrían resultar complejos cuando se realiza inferencia en modelos sofisticados, pueden ser expresados de manera más sencilla como representaciones gráficas.

En este contexto los GDA proporcionan una manera simple e intuitiva de representar y demostrar la información y el conocimiento disponible sobre relaciones causales entre variables estudiadas (87). El ajuste asociado para estimar con precisión un efecto causal puede deducirse mediante la mera inspección del grafo o bien algorítmicamente, dependiendo de la estructura y complejidad del GDA (88–90). Si bien la precisión del resultado de la estimación depende de cuánto se aproxime el GDA a la realidad, el hecho de que un GDA sea más fácil de generar y de compartir, hace que los supuestos que analiza sean más explícitos y estén más abiertos al escrutinio de la comunidad científica, comparado con otros modelos estadísticos. A pesar de estos beneficios, el uso de grafos dirigidos acíclicos es relativamente poco frecuente en la investigación médica aplicada, probablemente debido al inmovilismo en cuanto a la metodología de investigación (91).

1.3.3.- Redes bayesianas

La Real Academia Española define la inteligencia artificial (IA) como “la disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico” (92). De este modo, las computadoras son capaces de lograr objetivos según la información que procesan, sin estar explícitamente programadas para una acción específica (93).

Impulsada por la creciente disponibilidad de datos y el rápido desarrollo de técnicas de análisis, la IA ha ido implantándose progresivamente en distintas áreas de la medicina, incluyendo el TH. El manejo de los pacientes antes y después del trasplante requiere tomar decisiones en escenarios clínicos complejos. En este contexto la IA, que incluyen los MGP y las redes bayesianas (RB), puede impulsar un cambio de paradigma real, dado que permite analizar y sintetizar grandes cantidades de datos y transformar esa información en directrices o recomendaciones para la toma de decisiones clínicas.

En torno a 1750, el matemático británico Thomas Bayes desarrolló el teorema que lleva su nombre. En la “inferencia bayesiana”, las evidencias u observaciones se emplean para actualizar o inferir la probabilidad de que una hipótesis pueda ser cierta. La creencia inicial en la posibilidad de que ocurra un evento o resultado particular se denomina *probabilidad a priori* (para una hipótesis H , se describe como $P(H)$); mientras que la creencia actualizada a la luz de una nueva evidencia E se denomina *probabilidad a posteriori* (descrito como

$P(H|E)$, que representa la probabilidad condicional de que ocurra H, dado E). Este proceso de inferencia es muy común en la práctica clínica, por ejemplo, cuando profesionales médicos realizan un diagnóstico diferencial en circunstancias en las que múltiples enfermedades pueden presentar síntomas similares. A medida que las pruebas diagnósticas proporcionan nuevos datos, los médicos actualizan la información y las probabilidades acerca de qué patología en concreto afecta al paciente estudiado.

El teorema de Bayes es una ecuación simple para calcular la probabilidad *a posteriori* de una hipótesis H en términos de una probabilidad *a priori* E, en términos de la distribución de probabilidad condicional de la hipótesis H dado E:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|not H)P(not H)}$$

En términos más generales, el teorema de Bayes es de enorme relevancia puesto que vincula la probabilidad de H dado E con la probabilidad de E dado H. Tomando como ejemplo una patología médica, si sabemos la probabilidad de tener fiebre cuando se tiene apendicitis, se podría saber la probabilidad de tener apendicitis si se tiene fiebre. Este sencillo ejemplo destaca la relevancia del teorema en la práctica totalidad de las ramas de la ciencia, puesto que tiene una íntima relación con la comprensión de la probabilidad de aspectos causales dados unos efectos observados.

Las RB están basadas en el teorema de Bayes y usan un grafo para realizar una representación gráfica de distribuciones de probabilidad multivariable y razonamiento eficiente en condiciones de incertidumbre. En síntesis, las RB son un tipo de grafo dirigido acíclico con una parte cualitativa y una parte cuantitativa (94). La parte cualitativa corresponde a la estructura de la RB, constituida por los nodos, que representan variables aleatorias (discretas o continuas) y los arcos dirigidos, que representan las relaciones probabilísticas entre las variables. La parte cuantitativa son los parámetros de la RB, un conjunto de funciones probabilísticas condicionales asociadas a cada nodo, que representan la distribución de probabilidad de un nodo en la RB dados sus nodos padres.

Si bien se asume que las variables que incluye una RB deben ser discretas (tomando un valor determinado dentro de un conjunto finito de posibles valores), una RB puede incluir también variables continuas que adopten un valor dentro de un rango de valores reales (95). Generalmente estas variables continuas suelen adoptar una distribución normal (o de Gauss), de tal forma que se expresan mediante los parámetros habituales, como la media y la varianza. Muchos modelos basados en RB incluyen, por tanto, variables discretas y continuas.

1.3.3.1.- Funcionamiento de las redes bayesianas

Tanto la estructura de una RB como los parámetros que incluye pueden elaborarse mediante 3 métodos diferentes:

- a) Utilizando aprendizaje automatizado si los datos disponibles son suficientes (96).
- b) Construcción manual, utilizando métodos para extraer la información derivada del conocimiento de expertos reflejada en la literatura existente (97).
- c) Una combinación de ambos métodos (98).

Aprendizaje automatizado

En muchos campos de las ciencias de la salud, la información es almacenada en bases de datos. Estas bases de datos contienen información de gran valor sobre las posibles relaciones probabilísticas entre variables, aunque sea de forma implícita. Así, una RB puede ser generada y “aprender” de forma automática a partir de la información contenida en una base de datos. Es decir, se puede construir una RB sin necesidad de acceder al conocimiento de expertos humanos en la materia.

Para ser susceptible de generar el aprendizaje automático de una RB, el conjunto de datos disponibles debe cumplir una serie de condiciones. En primer lugar, dicho conjunto de datos debe haber sido recopilado cuidadosamente. Los sesgos que puedan derivarse de la estrategia de recogida de datos tienen un impacto en la RB resultante, dificultando así el propósito con el que se desarrolla la RB. Además, debe incluir la suficiente cantidad de información (número de casos) como para permitir

una identificación fiable de las relaciones probabilísticas entre las variables analizadas.

Aparte de estos requisitos generales, el conjunto de datos debe tener una serie de propiedades que la mayoría de los algoritmos de aprendizaje asumen implícitamente. Una de ellas es la suposición de que cada caso en el conjunto de datos especifica un valor para cada variable analizada, es decir, que no hay valores perdidos. Desafortunadamente, en la práctica habitual, esta condición no se cumple siempre. Para poder utilizar un conjunto de datos con valores perdidos, estos valores pueden deducirse, por ejemplo, en función de las probabilidades aproximadamente estimadas de estos valores, o con la ayuda de expertos en la materia.

La mayoría de los algoritmos de aprendizaje asumen además que el conjunto de datos se ha generado de forma independiente, es decir, supone que los valores especificados para las variables en un caso no están influenciados de ninguna manera por los valores en casos generados previamente. Finalmente se supone que el proceso de generación de datos no depende del tiempo transcurrido.

El aprendizaje automático de una RB implica asimismo la tarea de aprendizaje de la estructura del grafo, así como el aprendizaje de parámetros, es decir, la estimación de las distribuciones de probabilidad condicionales que se asociarán con el grafo. En muchos algoritmos de aprendizaje, las dos

tareas se llevan a cabo simultáneamente y, como consecuencia, es difícil separarlas como procesos aislados.

Uno de los primeros algoritmos de aprendizaje automático de una RB a partir de datos es el algoritmo *K2* (99). Dado un conjunto de datos D , este algoritmo busca el grafo acíclico que mejor explique la distribución de probabilidades entre las variables, complementado con estimaciones de máxima verosimilitud. Expresado formalmente, el algoritmo busca un grafo G^* que maximice la probabilidad conjunta $Pr(G, D)$ sobre todos los grafos posibles G . Dadas las variables aleatorias en cuestión, el algoritmo construye para cada variable posterior V_i , un conjunto óptimo de nodos padres. Para este objetivo el algoritmo empieza suponiendo que el conjunto de nodos padres está vacío y luego agrega, de forma recurrente, el nodo padre cuya adición aumenta más la probabilidad de la estructura resultante y el conjunto de datos. Asimismo, deja de agregar variables al conjunto de nodos padre tan pronto como la adición no puede aumentar la probabilidad conjunta $Pr(G, D)$. El algoritmo *K2* es un ejemplo de este método de *búsqueda y puntuación*. Este tipo de métodos buscan el mayor espectro posible de grafos acíclicos generando distintos grafos diferentes y comparándolos en función de su capacidad para explicar los datos disponibles.

Otro enfoque de aprendizaje automatizado de una RB se basa en la construcción de la misma sobre un análisis de

dependencia. En esencia, una RB modela un conjunto de relaciones de dependencia (o independencia) condicionadas. Al estudiar el conjunto de datos disponible, se pueden extraer las dependencias e independencias entre variables, mediante pruebas estadísticas, y posteriormente capturarlas e incorporarlas a la estructura del grafo. El algoritmo teórico-informativo de Cheng et al. es un ejemplo de este tipo de enfoque (100).

Construcción manual

Muchas RB desarrolladas para su aplicación en biomedicina se han construido de forma manual (101–103). Este método incluye una serie de fases en cada una de las cuales la información se obtiene consultando a expertos en el campo en cuestión, se analiza la literatura médica relevante que pueda existir y se actualizan datos clínicos de pacientes, si los hubiese. En general, podemos distinguir las siguientes fases:

- 1.- Selección de variables relevantes. Generalmente esta fase se basa en entrevistas con expertos en la materia y un análisis del propósito de la red que se va a construir, por lo que es necesario un conocimiento importante de los procesos fisiopatológicos que se van a estudiar para decidir las posibles variables que pueden tener importancia en el problema analizado (104).

2.- Identificación de las relaciones entre las variables. Una vez seleccionadas las variables que incluye el modelo, la dependencia o independencia entre ellas debe ser expresada en un modelo gráfico. Para este propósito, frecuentemente se emplea, como guía, la noción de causalidad. Durante esta fase es habitual tratar de responder a cuestiones como “¿qué podría causar este efecto?”, o bien, “¿qué manifestaciones tendría esta causa?”. Las conclusiones extraídas se expresan en términos gráficos mediante la dirección de los arcos que conectan los nodos (las variables) (105).

3.- Identificación de restricciones lógicas cualitativas y probabilísticas. En esta fase se descartan determinadas relaciones probabilísticas entre variables, por ilógicas, reduciendo así el número de probabilidades que deben ser incluidas entre las variables (106).

4.- Determinación de probabilidades. En la siguiente fase, las distribuciones probabilísticas para cada variable son introducidas en la RB. Estas distribuciones pueden ser obtenidas a partir de expertos en la materia, si bien hay algoritmos de cálculo prediseñados, que se adaptan para obtener la gran cantidad de probabilidades requeridas en un plazo de tiempo razonable (107). De manera alternativa, las probabilidades pueden ser estimadas en base a datos previamente conocidos.

5.- Análisis de sensibilidad y evaluación. Tras la fase anterior, la RB se encuentra estructuralmente completa. Sin embargo, antes de que la RB se pueda utilizar en la práctica habitual, debe establecerse su calidad y su valor en términos de aplicabilidad clínica mediante un análisis de sensibilidad con datos de pacientes (108). La evaluación de una RB puede realizarse de distintas maneras, como por ejemplo testando la RB para que clasifique un determinado conjunto de pacientes conocidos, o comparando la distribución de probabilidades de la RB construida con otra RB considerada de referencia o bien con otro modelo probabilístico.

Como se ha descrito anteriormente, el desarrollo de una RB es un proceso creativo, de tal forma que las distintas etapas se suceden de forma cíclica. Así pues, en cada repetición de cada fase, se produce un refinamiento de la exactitud de la RB que se está generando (109).

Aprendizaje automatizado vs. construcción manual

La construcción manual de una RB requiere acceso al conocimiento de profesionales expertos en la materia y, en la práctica, conlleva un consumo importante de tiempo. Con la creciente disponibilidad de bases de datos clínicos y biológicos, el aprendizaje automatizado es la alternativa más factible para desarrollar una RB. Como consecuencia, este enfoque está atrayendo un considerable interés dentro de la comunidad

investigadora. Sin embargo, la cuestión sobre cuál es el mejor método para el desarrollo de una RB (construcción manual vs. aprendizaje automatizado) sigue abierta. En muchas áreas de la biomedicina, cabría esperar que el conocimiento humano de los procesos fisiopatológicos fuera más sólido que el conocimiento incorporado en un conjunto de datos de tamaño limitado. Hasta la fecha, sin embargo, existe escasa evidencia que apoye esta afirmación.

Queda asimismo por dilucidar la cuestión de si es eficiente la creación de una red bayesiana de topología más compleja en comparación con el aprendizaje de un clasificador bayesiano simple. Es razonable pensar que cuanto más fiel sea el grafo de una red bayesiana al reflejar las dependencias e independencias incrustadas en los datos, mejor será su rendimiento (110). Sin embargo, diferentes estudios apuntan a que, especialmente cuando se aplican las RB a problemas de clasificación, las RB tipo *naïve* (el tipo de clasificador bayesiano más simple) tienden a superar a las redes más sofisticadas. Esto ha llevado a sugerir que las estructuras de redes más complejas no obtendrían resultados tan favorables. En contraposición, otros trabajos han postulado que las redes tipo *TAN* (*Tree Augmented Network*) (111), que en comparación con las RB *naïve* incorporan dependencias adicionales entre sus variables, a menudo superan a estas RB *naïve*, ya que obtienen un mejor rendimiento al permitir relaciones más complejas entre las variables que componen la red.

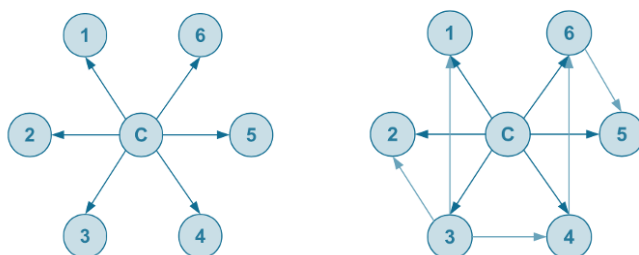


Figura 07: Naïve Bayes (izquierda) y Tree Augmented Naïve (derecha)
(112)

1.3.3.2.- Resolución de problemas mediante redes bayesianas en biomedicina

Las RB están siendo utilizadas de manera cada vez más frecuente con el fin de resolver diferentes tipos de problemas en el ámbito de la investigación biomédica.

a) Diagnóstico

En esencia, establecer un diagnóstico en un paciente en concreto equivale a construir una hipótesis sobre la enfermedad que padece el paciente, basada en un conjunto de observaciones indirectas de las pruebas diagnósticas realizadas. Generalmente, dichas pruebas no sirven para revelar inequívocamente la enfermedad que puede padecer el paciente: las pruebas generalmente tienen tasas de verdadero positivo y tasas de verdadero negativo que no alcanzan el

100 %. Para evitar diagnósticos erróneos a la hora de construir una hipótesis diagnóstica, debe tenerse en cuenta la imprecisión de los resultados obtenidos en las pruebas. Las RB ofrecen una base natural para este tipo de razonamiento en situaciones de incertidumbre (113,114).

Así, las RB se han aplicado para el diagnóstico de la neumonía pediátrica y sus diferentes etiologías (115), los trastornos del espectro autista en función de determinadas características de los pacientes (116) o la presencia de un angor inestable en función de los síntomas que presenta un individuo (117)

b) Pronóstico

El razonamiento pronóstico en medicina equivale a hacer una predicción sobre lo que sucederá en el futuro. Como el conocimiento del futuro es inherentemente incierto, en el razonamiento pronóstico la incertidumbre es aún mayor que en el razonamiento diagnóstico.

Así, las redes bayesianas de pronóstico suelen tener una estructura temporal general, que se representa esquemáticamente en la siguiente figura.



Figura 08: Representación gráfica de un modelo general de red bayesiana pronóstica

El resultado predicho para un paciente específico generalmente se ve influido por la secuencia particular de acciones del tratamiento a realizar, de modo que, formalmente el pronóstico de una enfermedad puede definirse por la siguiente distribución de probabilidad:

$$\text{Pr}(\text{resultado} \mid E, T)$$

Donde E corresponde a los datos disponibles del paciente, incluyendo síntomas, signos y resultados de la prueba, y T se corresponde con una secuencia de acciones terapéuticas realizadas. El resultado de interés puede expresarse mediante una sola variable, por ejemplo, determinando la supervivencia global o la supervivencia libre de enfermedad. Sin embargo, el resultado de interés puede ser más complejo, incluyendo aspectos que relacionen la calidad de vida.

Así, las RB se han implementado con el fin de predecir los resultados de determinados tipos de terapias en pacientes frágiles (118), la supervivencia en algunos tipos de cánceres primarios o metastásicos (119,120), la calidad de vida de los pacientes después de procedimientos de cirugía bariátrica (121), o la probabilidad de infección intraabdominal después de intervenciones quirúrgicas de patología colorrectal (122).

c) Selección del tratamiento

Una RB y sus algoritmos de razonamiento asociados, a menudo se integran en un sistema de apoyo a la decisión que ofrece las herramientas necesarias para seleccionar un tratamiento óptimo dadas ciertas predicciones (103). Un ejemplo es el diagrama de influencia (123), que tiene una estructura general básica simple: como en cualquier RB, el diagrama de influencia incluye variables aleatorias (representadas como nodos ovalados). Asimismo, también incorpora nodos que representan las decisiones tomadas (rectángulos) y los valores de utilidad (U) que constituyen el criterio para la toma de decisiones (representados como rombos). El objetivo del diagrama de influencia es maximizar dicha utilidad.

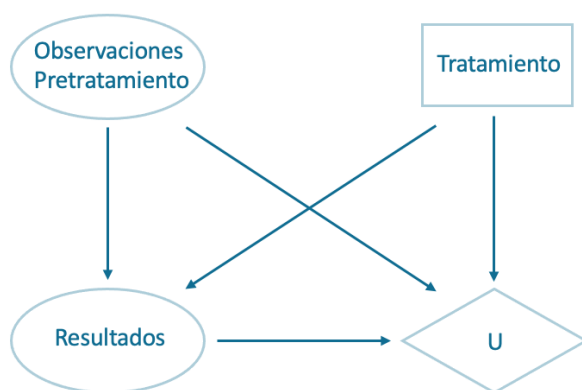


Figura 09: Diagrama de influencia para selección de tratamiento (123)

En este contexto, existen algunos ejemplos de estudios que han aplicado las RB para tratar de establecer estrategias terapéuticas con corticosteroides en escenarios clínicos de sepsis (124) o la selección de tratamiento en patología oncológica mamaria (125).

d) Descubrimiento de interacciones funcionales

Dado que la topología de una red bayesiana puede interpretarse como una representación de las interacciones entre variables, existe un creciente interés en el campo de la inteligencia artificial aplicada a biomedicina para utilizar una RB con el objetivo de desentrañar los mecanismos moleculares que subyacen en el funcionamiento celular. Por ejemplo, en el

contexto de la oncología, las RB pueden tener un gran impacto en el descubrimiento de interacciones previamente desconocidas. En este sentido, se ha analizado mediante RB cómo la expresión de determinados oncogenes puede influir en la agresividad o capacidad metastásica de determinados tipos de tumores (126,127).

Como vemos, las RB han suscitado un interés creciente por parte de la comunidad investigadora en ciencias de la salud en los últimos años. Sin embargo, a pesar de que la cantidad de literatura médica relacionada con RB está aumentando rápidamente, no lo hace de forma equitativa en las distintas ramas de la Medicina. Así, es posible identificar un grupo de patologías o campos de investigación en los que las RB están siendo aplicadas con mayor interés. La patología cardíaca, los procesos oncológicos en general, la psiquiatría y los trastornos psicológicos, así como las afecciones respiratorias aglutinan aproximadamente 2/3 de todos los estudios en los que se emplean RB (128). Estas cifras no deben sorprendernos ya que se trata de enfermedades muy prevalentes y que ocupan las primeras cifras de mortalidad a nivel global, especialmente en países desarrollados.

1.4.- Estado actual del tema

El carcinoma hepatocelular es un tumor con una incidencia creciente en las últimas décadas, no solo en nuestro ámbito sino especialmente en países orientales.

La estrecha relación entre el CHC y la hepatopatía de base limita las opciones terapéuticas del paciente, por lo que, si bien es cierto que un elevado porcentaje de pacientes no son candidatos a tratamiento curativo en el momento del diagnóstico, aquellos que se detectan en estadios precoces pueden beneficiarse de opciones curativas como la ablación, la resección quirúrgica o el trasplante hepático.

La resección quirúrgica precisa de unos criterios estrictos, especialmente en lo que se refiere a la función hepática, la presencia de hipertensión portal y el volumen del remanente hepático. En pacientes con CHC únicos de tamaño igual o inferior a 2 cm, la supervivencia a los 5 años tras la cirugía alcanza resultados del 80-90 % (con una mortalidad del 3 %) que, además, son similares a los obtenidos en pacientes tratados con RF, siendo ésta una técnica que asocia menor morbilidad que la resección quirúrgica (24). No obstante, el desarrollo de técnicas menos invasivas como la laparoscopia o la cirugía robótica han disminuido la morbilidad de la resección, además de que los datos anatomopatológicos ofrecen mejor información para evaluar el riesgo de recidiva.

En el otro extremo, con el trasplante hepático es posible obtener la curación no sólo del tumor, sino también de la hepatopatía crónica

subyacente. El trasplante ofrece unos excelentes resultados de supervivencia a los 5 años, superiores al 60 % cuando se aplican los criterios de Milán, con una mortalidad posoperatoria en torno al 10 %, aunque la consolidación y la estandarización de la técnica han reducido las complicaciones y la morbilidad asociada al trasplante. Comparativamente con la resección quirúrgica, la principal diferencia entre ambas opciones es la recidiva, que se sitúa en torno al 70 % a 5 años para la resección quirúrgica, mientras que el trasplante hepático (con los criterios de Milán) ofrece cifras de recidiva en torno al 10-15 % a 5 años (129).

Los buenos resultados a largo plazo en supervivencia libre de enfermedad han animado a muchos grupos de trasplante a establecer criterios expandidos, siendo menos restrictivos a la hora de incluir pacientes con CHC en lista de espera de trasplante. En este sentido, la SETH recomienda que los criterios de Milán sigan siendo la estrategia de referencia para la selección de pacientes. Sin embargo, reconoce también que una expansión moderada de los criterios de Milán es recomendable, siempre y cuando se mantenga el equilibrio de priorización con respecto al resto de indicaciones de trasplante. Entre todos los criterios de expansión moderada, el sistema *Up-to-seven*, es el que cuenta con mayor evidencia científica y validez externa, por lo que podrían considerarse como de primera elección (17).

Sin embargo, los excelentes resultados a largo plazo del trasplante se enfrentan en muchos países a la realidad del tiempo necesario de

espera y el riesgo de progresión de enfermedad y, por tanto, de salida (*drop-out*) de la lista de espera, especialmente si se amplían los criterios de inclusión, habida cuenta de que el número de injertos disponibles siempre es menor al número de pacientes en lista de espera.

En los últimos años, la aparición de los nuevos AAD ha supuesto un importante cambio en el manejo de la hepatitis crónica por VHC, que era la principal causa de cirrosis en pacientes con CHC en los países occidentales. La curación de muchos de ellos ha supuesto una importante reducción de las listas de espera. Paralelamente, se ha producido un incremento del número de injertos disponibles gracias a la generalización de las técnicas de obtención de órganos procedentes de donantes en asistolia controlada. Todo ello ha llevado a que en España vivamos una situación de relativa bonanza respecto al volumen de donantes ya que, además, somos el país con mayor cantidad de donaciones por número de habitantes.

Sin embargo, es importante destacar que la situación en nuestro medio no es en ningún caso extrapolable a otros países, por lo que la escasez de donantes sigue siendo un problema acuciante para la mayoría de los grupos de trasplante en el mundo. Además, paralelamente se está produciendo un importante crecimiento de la esteatohepatitis no alcohólica como causa de cirrosis y por tanto de CHC, especialmente en los países occidentales, debido a la generalización de hábitos de vida menos saludables.

Por tanto, resulta evidente que nos encontramos en una situación en la que el trasplante hepático es el mejor tratamiento disponible para la curación de aquellos pacientes con CHC en estadios precoces con cirrosis, que puedan beneficiarse de este tratamiento. Sin embargo, debido a la escasez de órganos disponibles, es fundamental realizar una adecuada selección de pacientes con un doble objetivo: identificar aquellos que tengan menor probabilidad de recurrencia y optimizar la utilización de los injertos, garantizando a su vez una mayor supervivencia libre de enfermedad a los pacientes incluidos en lista de espera para trasplante hepático.

Hoy en día existen datos que indican que la respuesta a los tratamientos neoadyuvantes (tanto de *downstaging* como terapias puente) son relevantes y tienen influencia en los resultados de SG y SLE postrasplante. Generalmente, una respuesta satisfactoria a estos tratamientos está relacionada con factores de buen pronóstico como el bajo grado tumoral o la ausencia de satelitosis o invasión microvascular. Sin embargo, estos datos sólo se encuentran disponibles *a posteriori*, una vez estudiado el hígado cirrótico. En este sentido, la respuesta radiológica a estos tratamientos tiene un papel importante en la predicción de la agresividad del tumor, traduciendo en cierta manera la biología molecular y el comportamiento del tumor, por lo que se considera un factor pronóstico relevante y se ha convertido en sí mismo en una herramienta para la selección de pacientes para el trasplante hepático.

Por otra parte, los marcadores biológicos (fundamentalmente la AFP) han pasado de tener un papel meramente diagnóstico o de seguimiento a convertirse en otro factor relevante en el pronóstico de los pacientes, completando la información que aportan los datos morfológicos de las pruebas de imagen pretrasplante. Tanto es así que desde 2018, la SETH recomienda contraindicar formalmente el TH en pacientes con niveles de AFP > 1000 ng/ml (nivel de evidencia 1, grado de recomendación fuerte) así como realizar un tratamiento locorregional sobre el CHC y reestadificar completamente a aquellos pacientes con niveles de AFP ≥ 400 ng/ml incluidos en programas de criterios expandidos (nivel de evidencia 2, grado de recomendación moderado).

Aunque es cierto que los criterios de Milán siguen siendo la piedra angular sobre la que se sustentan las políticas de inclusión en lista de espera de la mayoría de los grupos de trasplante, es evidente también que desde hace más de 10 años se ha generado una creciente evidencia que demuestra no sólo que estos criterios pueden ser excesivamente restrictivos, si no que incluso con criterios expandidos, la información meramente morfológica puede ser insuficiente para predecir el comportamiento oncológico del CHC y, por tanto, el riesgo de recurrencia. En el campo de la *transplant oncology*, parece claro que es fundamental incluir otro tipo de datos (marcadores biológicos, respuesta a la terapia pretrasplante) que completen la información al respecto de la biología tumoral y orientar con mayor precisión el riesgo de recurrencia de los pacientes.

Así pues, en los últimos años se han desarrollado multitud de criterios e índices pronósticos combinados con el objetivo de ofrecer el trasplante hepático a un mayor número de pacientes sin la limitación de los criterios morfológicos, incluyendo marcadores predictivos biológicos y de comportamiento tumoral. En este sentido, no existe unanimidad en torno a la mejor herramienta pronóstica. Si bien es cierto que la mayor parte de grupos de trasplante emplean criterios morfológicos (Milán o expandidos), hoy en día es un hecho comúnmente aceptado que los criterios de trasplante deberían incluir como mínimo no sólo factores morfológicos, sino también factores relacionados con la biología tumoral (en general mediante el uso de marcadores como la AFP). Esta afirmación se refrenda con el hecho de que estos nuevos criterios combinados tienen una capacidad predictiva superior a los criterios morfológicos (tanto de Milán, como expandidos: *Up-to-seven*, UCSF...). Sin embargo, a día de hoy no existe ningún estudio cuyo objetivo principal sea la comparación de estos criterios combinados, con el fin de determinar cuál de las herramientas predictivas disponibles actualmente es la más adecuada.

A continuación, se muestra una tabla comparativa con distintos criterios y scores pronósticos de TH para CHC, incluyendo sus características y los datos que emplean.

	Tipo	Morfología tumoral	Marcadores biológicos	MELD	AP	Respuesta a terapia locorregional	Tiempo en lista de espera
Milán	CAT	X					
UCSF	CAT	X					
<i>Up-to-seven</i>	CAT	X					
Navarra	CAT	X					
Valencia	CAT	X					
Seúl	CAT	X					
Toronto	CAT				X		
AFP-model	CAT	X	AFP				
Metroticket 2.0	CAT	X	AFP				
TRAIN	CAT		AFP – RNL			X	X
HALT-HCC	CONT	X	AFP	X			
MORAL	CAT	X	AFP – RNL				
Kyoto	CAT	X	DCP				

CAT: Categórico

CONT: Continuo

Tabla 07: Resumen de criterios y scores pronósticos de trasplante para CHC

Por otra parte, existen también diferencias en la información que estos índices proporcionan a los especialistas. Algunos de ellos ofrecen una respuesta categórica (sí/no) a la decisión de trasplantar a un paciente. Otros, sin embargo, ofrecen un índice o puntuación continua a partir del cual se toman las decisiones. Cada estrategia tiene sus ventajas y desventajas. Aquellos *scores* que ofrecen una respuesta categórica pueden hacer que las decisiones clínicas sean más sencillas, ya que reducen la cuestión a la dicotomía “trasplantar vs. no trasplantar a un paciente”. Sin embargo, esta estrategia corre el riesgo de ser excesivamente reduccionista, obviando pequeñas diferencias en el riesgo de determinados pacientes que pueden ser significativas y decantar la decisión acerca de trasplantarlo o no. Las herramientas pronósticas que ofrecen una puntuación continua sí pueden ser capaces de detectar estas diferencias, pero son más difíciles de encajar en las estrategias de trasplante, y hacen necesario el establecimiento de un umbral para la toma de decisiones.

En otro orden de cosas, la gran mayoría de estudios disponibles sobre la aplicabilidad de nuevos criterios y herramientas pronósticas del CHC en *transplant oncology*, tienen como objetivo principal analizar el impacto sobre la supervivencia global y/o libre de enfermedad de los pacientes. En este sentido, se obvia la cuestión de que un porcentaje no desdeñable de pacientes presentan recidivas precoces del CHC (dentro de los dos primeros años después del TH). Esta situación es extremadamente ineficiente desde el punto de vista de la utilización de un recurso escaso como son los injertos hepáticos disponibles.

En resumen, el campo de la *transplant oncology* sigue enfrentándose a la tarea de ofrecer una supervivencia prolongada a un número creciente de pacientes, incorporando al esquema de decisiones toda la información disponible acerca de la carga tumoral y del comportamiento biológico del CHC en cada caso concreto. Los avances en las terapias locorregionales han demostrado que estos tratamientos pueden proporcionar una supervivencia aceptable a un gran número de pacientes sin necesidad de recibir un TH. No obstante, los resultados a nivel de recurrencia distan de ser comparables a los que ofrece el TH. Por tanto, es fundamental identificar previamente al trasplante aquellos pacientes que tienen un menor riesgo de recurrencia precoz, ya que permitiría una optimización del uso de los injertos disponibles, al seleccionar a los candidatos idóneos para este tratamiento.

En este sentido los sistemas de inteligencia artificial y *machine learning* como las redes bayesianas pueden ayudar a modelizar una situación clínica de gran complejidad como es el comportamiento tumoral del CHC, en la que influyen múltiples variables que pueden interaccionar de manera distinta entre sí y con la variable objetivo de estudio (recurrencia tumoral postrasplante).

Para ello diseñamos el presente trabajo, en el que utilizamos un modelo de redes bayesianas para predecir de forma más exacta la recurrencia tumoral a los dos años postrasplante (situación menos eficiente del trasplante hepático) de los pacientes candidatos a TH en función de variables conocidas antes de la indicación del trasplante.

Consideramos que este modelo sería de gran utilidad para la toma de decisiones clínicas. Además, el modelo puede crearse con una interfaz informática de acceso *web* abierto que permita la utilización del mismo por los diferentes grupos de trasplante y su aplicación clínica.

Hipótesis de trabajo

2.- Hipótesis de trabajo

Es posible desarrollar un modelo predictivo de recidiva precoz de carcinoma hepatocelular en los 2 primeros años después del trasplante hepático, basado en redes bayesianas, utilizando exclusivamente información disponible antes del trasplante y permitiendo así identificar aquellos pacientes con menor riesgo de recurrencia.

Objetivos del estudio

3.- Objetivos del estudio

Planteamos como objetivo principal del estudio:

- Desarrollar un modelo predictivo basado en redes bayesianas para determinar la recidiva del carcinoma hepatocelular en los dos primeros años postrasplante, utilizando exclusivamente variables preoperatorias del receptor, tumor y donante.

Se definen como objetivos secundarios del estudio:

1. Comparar la calidad predictiva del modelo desarrollado mediante redes bayesianas frente a un modelo estándar de regresión logística.
2. Validar el modelo predictivo elaborado en una serie distinta de casos trasplantados con posterioridad en el mismo centro y comparar los resultados.
3. Elaborar una propuesta para la aplicación en la práctica clínica del modelo diseñado, a través de una interfaz informática accesible.

Material y métodos

4.- Material y métodos

4.1.- Diseño del estudio

Se trata de un estudio retrospectivo, descriptivo, analítico y unicéntrico. El estudio se ha realizado en el Hospital Universitari i Politècnic La Fe de Valencia sobre una muestra de receptores adultos de trasplante hepático por carcinoma hepatocelular, controlados al menos durante 2 años después del trasplante.

El objetivo del estudio es desarrollar un modelo predictivo de recurrencia tumoral en los 2 primeros años postrasplante mediante un sistema de redes bayesianas, empleando solo variables definidas antes del trasplante y conocidas preoperatoriamente.

4.2.- Población del estudio

La población del estudio incluye a todos los pacientes mayores de 18 años receptores de trasplante hepático que cumplían los criterios de inclusión y ninguno de exclusión descritos en el apartado 4.3, trasplantados en el Hospital Universitari i Politècnic La Fe de Valencia entre el 1 de enero de 2004 y el 31 de diciembre de 2018. El número total de pacientes incluidos en el estudio es de 515.

En todos los casos se realizó una técnica estándar de trasplante ortotópico, con hepatectomía *piggy-back* preservando la vena cava.

Para la creación del modelo predictivo basado en RB los 515 pacientes incluidos fueron divididos en 2 grupos, un subconjunto de entrenamiento ($n = 385$) y otro de generalización ($n = 130$), en las proporciones 75 / 25 %. La validación interna del modelo se realizó sobre el subconjunto de generalización, ya que la división se realizó con un criterio temporal, de tal forma que los pacientes del subconjunto de entrenamiento son aquellos trasplantados desde el inicio de la serie.

Asimismo, los integrantes del subconjunto de generalización son individuos trasplantados de forma consecutiva en el mismo centro hasta el final del período de estudio, cumpliendo todos ellos los mismos criterios de inclusión y exclusión, como se puede apreciar en la siguiente tabla:

Desarrollo de los modelos		
	Entrenamiento (01/01/2004 – 01/11/2015)	Generalización (02/11/2015 – 31/12/ 2018)
Redes bayesianas	385	130
Regresión logística	385	130

Tabla 08: División de los pacientes del estudio según grupos para el desarrollo de modelos predictivos

4.3.- Criterios de inclusión y exclusión

A) Los criterios de inclusión en el estudio son:

1. Receptores de trasplante de hígado ortotópico.
2. Mayores de 18 años.
3. Carcinoma hepatocelular confirmado por estudio anatomopatológico postrasplante.

B) Los criterios de exclusión en el estudio son:

1. Pacientes en los que haya una pérdida de seguimiento en los 2 primeros años tras el trasplante hepático.
2. Pacientes sin diagnóstico de CHC antes del trasplante (CHC incidental).

4.4.- Protocolo de tratamiento pretrasplante

Una vez diagnosticado el CHC en pruebas de imagen siguiendo las guías de la EASL (39), los pacientes son evaluados por un comité multidisciplinar (CMD) compuesto por hepatólogos, cirujanos, radiólogos (incluyendo radiólogos intervencionistas), anatomopatólogos y oncólogos. Se analiza la situación general del paciente, comorbilidades, estado de su hepatopatía de base y los factores relacionados con el tumor (tamaño, localización, número de nódulos, afectación vascular, etc.).

En la gran mayoría de casos se indica tratamiento pretrasplante (salvo que la situación de la función hepática lo contraindique), incluso en aquellos pacientes en que se espera un tiempo en lista de espera inferior a 6 meses (como terapia puente). En función de la información disponible se selecciona el tratamiento más adecuado atendiendo a un protocolo que incluye como opciones terapéuticas las siguientes:

1. Ablación percutánea

- Radiofrecuencia
- Microondas
- Inyección percutánea de alcohol

2. Tratamientos intraarteriales

- Quimioembolización transarterial
- Radioembolización transarterial

Cada una de estas terapias se aplica en función de cada escenario clínico concreto:

1.- Ablación percutánea

Como indicaciones generales se consideran:

- Hepatocarcinoma nodular (no infiltrativo)
- Estadios muy precoces o estadios tempranos (BCLC 0 y A)
- Número máximo de lesiones a tratar: 3
- Tamaño máximo para tratar: 3 cm por lesión

En todos los casos se realiza la ablación guiada por ecografía. Se considera técnica de elección la RF, siendo alternativas terapéuticas la ablación por MO y la inyección de alcohol, especialmente en aquellas lesiones de localización menos favorable para la RF.

Asimismo, existen determinados escenarios clínicos en los que es necesario tener en cuenta una serie de consideraciones:

- Estadio BCLC A, lesión única y < 2 cm: en estos casos la supervivencia es equiparable a la resección quirúrgica, si bien esta última ofrece más información sobre factores pronósticos anatomopatológicos. En presencia de factores de mal pronóstico después de la cirugía se considera inclusión en lista de espera de

trasplante *ab initio* (sin esperar a la presencia de recidiva).

- Lesiones únicas 3-5 cm: se realiza tratamiento combinado con tratamiento percutáneo con QETA.
- Lesión sin acceso percutáneo: se valora la posibilidad de realizar una QETA.

Previamente a la realización del tratamiento se realizan en todos los pacientes las siguientes pruebas de imagen:

- Ecografía hepática: con el fin de valorar la viabilidad del acceso percutáneo.
- Técnica de imagen: TC o RM con contraste endovenoso y estudio en fases dinámicas, Se procura que las pruebas de imagen complementarias sean recientes, con una antigüedad no superior a 1-2 meses, en cuyo caso, se actualizan.

En cuanto al seguimiento de los pacientes, también se realizan una serie de pruebas protocolizadas:

- Control TC/RM: se realiza 1 mes después del tratamiento, asociado a analítica completa que incluye AFP y valoración del paciente en consultas externas:

- Si hay respuesta completa: control cada 3 meses con TC/RM y analítica con AFP.
- Si hay respuesta parcial: se valora repetir el tratamiento de ablación.
- Si hay progresión: se valora de nuevo el paciente en comité multidisciplinar, para repetir el tratamiento o ofrecerle una migración terapéutica con cambio de tratamiento (intraarterial / sistémico).

2.- Tratamientos intraarteriales

2.1.- Quimioembolización transarterial

Como indicaciones generales se consideran:

- Pacientes en estadio BCLC B (intermedio)
 - Tumores nodulares hipervascularizados (no infiltrativos).
 - Diámetro máximo de 8 cm en caso de lesión única
 - Función hepática conservada (máximo Child-Pugh B8).
- Migración terapéutica de estadio BCLC A que no pueda tratarse con una técnica curativa (por ejemplo: nódulo < 3 cm sin acceso percutáneo seguro).

- Tratamiento combinado con técnicas de ablación percutánea: lesiones únicas de entre 3 y 5 cm.

Asimismo, como contraindicaciones para esta modalidad terapéutica se consideran:

Absolutas:

- Descompensación función hepática (máximo Child-Pugh B8)
- Afectación tumoral extrahepática
- Afectación masiva bilobar
- Tumores infiltrativos
- Trombosis de vena porta principal
- Flujo portal hepatófugo
- Fístulas arterio-portales no tratables
- Limitaciones de acceso vascular (técnicas, diátesis hemorrágica no corregible)

Relativas:

- Trombosis portal lobar o segmentaria (si se puede asegurar embolización selectiva)
- Insuficiencia renal (creatinina > 2 mg/dL o filtrado glomerular < 30 ml/min), siempre precedido de protocolos de protección renal frente al uso de contrastes yodados.
- Alergia al contraste yodado (el paciente entra previamente en un protocolo de desensibilización)

El seguimiento de los pacientes que reciben un tratamiento con QETA difiere en función de la situación clínica:

a) Tratamiento completo

1º mes tras QETA: TC + analítica con AFP + visita en consultas externas con resultados:

- a) Si existe respuesta completa (según los criterios mRECIST) y analítica normal: Controles cada tres meses con TC/RM hepática + analítica.
- b) Respuesta parcial o enfermedad estable (mRECIST) tras el control del primer mes. Si la analítica es normal (buena tolerancia al tratamiento y mantiene condiciones clínicas adecuadas), se planteará en CMD la realización de una nueva QETA. Asimismo, se puede valorar la posibilidad de realizar ablación percutánea, en función de la carga tumoral presente. Se realiza un nuevo control de imagen con analítica al mes de la segunda sesión. Si tras el segundo tratamiento se obtiene una respuesta completa, el paciente sigue controles cada tres meses durante el primer año (TC/RM + analítica).
- c) Si hay progresión al mes del tratamiento: se considera como tumor de alta agresividad. Se reevalúa el paciente en CMD para plantear migración terapéutica: radioembolización o tratamiento sistémico (sorafenib). Esta misma estrategia terapéutica se aplica si hay

persistencia de enfermedad estable o respuesta parcial tras dos sesiones de QETA.

b) Tratamiento no completo en la primera sesión

Se planifica una nueva sesión de tratamiento en el plazo de quince días, con una nueva analítica completa previa a la nueva sesión.

2.2.- Radioembolización transarterial

Como indicaciones generales de la RETA se consideran:

- Tumor único ≥ 8 cm
- Tumores múltiples > 4 nódulos
- Trombosis portal lobar o segmentaria
- Progresión o recurrencia tras QETA

Asimismo, como contraindicaciones para esta modalidad terapéutica se consideran:

Absolutas:

- Expectativa de vida < 3 meses con ECOG > 2
- Child-Pugh C
- Insuficiencia pulmonar severa

- Ausencia de garantías de seguridad para la infusión del radioisótopo:
 - Shunt hepatopulmonar > 30 Gy en tratamiento único o > 50 Gy en sucesivos retratamientos
 - Depósitos extrahepáticos no corregibles con técnicas intervencionistas
 - Limitaciones de acceso vascular (técnicas, diátesis hemorrágica no corregible...)

Relativas:

- GOT/GPT > 5 veces valor normal
- Bilirrubina total > 2 mg/dl
- Carga tumoral > 50 % del volumen hepático con cifras de albúmina < 3 mg/dl
- Insuficiencia renal (creatinina > 2 mg/dL o filtrado glomerular < 30 ml/min), siempre precedido de protocolos de protección renal frente al uso de contrastes yodados)
- Alergia al contraste yodado (el paciente entra previamente en un protocolo de desensibilización)

En este caso, se trata de un tratamiento dividido en 2 etapas, separadas entre sí 2 o 3 semanas:

- Primera fase: arteriografía inicial para confirmar la ausencia de depósitos extrahepáticos (tracto gastrointestinal y pulmón). Embolización de arterias

hepatófugas e infusión de partículas de macroagregados de albúmina marcadas con Tecnecio 99m, seguido de estudio con gammagrafía y tomografía por emisión de fotón único.

- Segunda fase: arteriografía hepática seguida de la infusión de microesferas cargadas con Itrio-90 junto realización de gammagrafía y tomografía por emisión de positrones (PET).

El seguimiento de los pacientes que reciben un tratamiento con RETA difiere en función de la situación clínica:

	BILIRRUBINA NORMAL	BILIRRUBINA ELEVADA (> 2 mg/dL)
LESIÓN FOCAL	RETA segmentaria (preferido a lobar)	RETA segmentaria (única indicación <i>downstaging</i> THO)
MULTIFOCAL	RETA lobar secuencial	No indicación

Tabla 09: Esquema de tratamiento radioembolización en la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Hospital Universitari i Politècnic La Fe

Seguimiento

- Primer mes postratamiento: TC + analítica completa con AFP
- 3^{er} y 6^o mes postratamiento: TC/RM + analítica completa con AFP. Con estas exploraciones se valora el grado de respuesta tumoral al tratamiento:
 - Respuesta completa: se mantienen controles cada 3 meses.
 - Respuesta parcial o progresión de enfermedad: el paciente se reevalúa en comité multidisciplinar con el fin de valorar la posibilidad de un nuevo tratamiento con RETA o bien una migración de estadio e iniciar tratamiento sistémico (sorafenib).

4.5.- Evaluación de la respuesta tumoral al tratamiento pretrasplante

Una vez realizados los tratamientos correspondientes en función del protocolo anteriormente descrito, se evalúan las pruebas complementarias de seguimiento por parte de radiólogos especialistas en patología hepática. En este sentido la unidad de Cirugía HBP y Trasplante del Hospital La Fe se adhiere a las recomendaciones de las guías clínicas de la EASL para el manejo del CHC, que en el año 2018 establecieron los criterios mRECIST como el estándar para evaluar la respuesta tumoral a los tratamientos locorregionales pretrasplante. Es por ello que el seguimiento se realiza con pruebas de imagen dinámicas (TC y/o RM con contraste en varias fases). Es importante reseñar que se realiza un tratamiento locorregional pretrasplante a todo paciente incluido en lista de espera, siempre que la situación funcional del hígado lo permita. En caso de respuesta completa al tratamiento, el paciente no es retirado de lista de espera de TH, ya que el tratamiento primario establecido es el trasplante.

En aquellos pacientes tratados con intención de *downstaging*, se requieren 6 meses de estabilidad clínica y radiológica después de haber conseguido su entrada en criterios de Milán, que en el centro de estudio son los de referencia para inclusión en lista de espera. Pasado ese período de tiempo es obligatorio reevaluar de nuevo al paciente con TC/RM con el fin de constatar que sigue cumpliendo criterios de trasplante.

Una vez comprobado que no ha habido cambios en la situación clínica, se procede a la inclusión del paciente en lista de espera de TH.

4.6.- Criterios de inclusión en lista de espera de trasplante hepático para pacientes con carcinoma hepatocelular

Como criterio general, en la la unidad de Cirugía HBP y Trasplante del Hospital La Fe se incluyen en lista de espera de TH aquellos pacientes con CHC no tributarios de resección quirúrgica y que cumplen los criterios de Milán. Como criterios expandidos, se utilizan los criterios de Valencia.

En los pacientes en los que se haya realizado previamente resección quirúrgica del CHC, los criterios de inclusión son:

- CHC T1 o T2 con criterios de buen pronóstico (ausencia de satelitosis y de invasión microvascular en el estudio anatomopatológico). Se incluyen en lista de espera únicamente si se produce una recidiva dentro de los criterios de Milán
- CHC T1 o T2 con criterios de mal pronóstico (satelitosis y/o invasión microvascular en el estudio anatomopatológico): control clínico y con pruebas de imagen (TC / RM) y analítica con AFP 6 meses después de la resección. Inclusión en lista de espera con prioridad 3, si no hay evidencia de enfermedad o se encuentran dentro de los criterios de Milán. Inclusión *ab initio*.

- CHC con resección R1 en el estudio anatomopatológico. Se incluye en lista de espera con priorización acorde al estadio anterior a la extirpación tumoral, siempre que éste no sobrepase los criterios morfológicos de Milán o Valencia. En aquellos pacientes que excedan dichos criterios, se individualizará la estrategia terapéutica, con tendencia a indicar el TH, especialmente en los casos con recidivas tardías (más de 2 años después de la resección).

Los pacientes con CHC T1 y MELD inferior a 15, se incluirán en lista de espera siempre que no sea posible realizar un tratamiento con RF o cuando, una vez realizado, no tengan respuesta completa o haya recidiva en el seguimiento.

Como se ha comentado anteriormente, aquellos pacientes tratados con intención de *downstaging*, precisan 6 meses de estabilidad clínica y radiológica después de haber conseguido su entrada en criterios de Milán, además de unos niveles de AFP inferiores a 400 ng/ml, para poder ser incluidos en lista de espera.

Asimismo, unos niveles de AFP superiores a 1000 ng/ml constituyen (desde el año 2016) una contraindicación absoluta para la inclusión en lista de espera.

Por otra parte, se utilizan una serie de criterios clínicos de priorización que permiten trasplantar en primer lugar a aquellos pacientes que más lo necesitan. Los pacientes con CHC, por el riesgo de recidiva, siguen unos criterios de excepción al MELD funcional.

Dentro de la lista de espera, existen 2 tipos de prioridad aplicable a los pacientes con CHC:

- Prioridad 1: MELD 16 desde la inclusión en lista y se añade 1 punto cada mes en lista de espera.
- Prioridad 3: MELD 16 tras 6 meses en lista de espera y se añade 1 punto cada mes a partir del momento de la priorización.

En este sentido, el CHC constituye una patología priorizada dentro de las patologías tributarias de trasplante, con los siguientes criterios de priorización:

- Prioridad 1: Si cumple alguna de las siguientes condiciones
 - Nódulo único entre 3 y 5 cm, o bien 2 – 3 nódulos, ninguno con diámetro superior a 5 cm.
 - Alfafetoproteína > 200 ng/ml.
- En el resto de casos, se asignará prioridad 3.

4.7.- Seguimiento de los pacientes después del trasplante hepático

Una vez el paciente es dado de alta después del trasplante hepático, se realiza un seguimiento por parte de los servicios de Hepatología y Cirugía HBP. El seguimiento por parte de Hepatología es común a todas las indicaciones de trasplante y consiste fundamentalmente en una ecografía Doppler, que se realiza al 2º mes después del procedimiento. Posteriormente, se realiza un seguimiento clínico según su estado general y la función del injerto, asociado a analíticas seriadas que incluyen siempre el valor de AFP.

Con respecto al seguimiento de la recurrencia del CHC, antes del año 2017 era más errático y se realizaba generalmente en consultas externas de la Unidad de Hepatología del Hospital La Fe de Valencia. A partir de la publicación del *score* RETREAT (130) en el año 2017, este sistema pronóstico sirve de guía para el seguimiento postrasplante. Se trata de un *score* pronóstico de recidiva del CHC después del trasplante que tiene en cuenta el valor de la AFP el día del trasplante, el tamaño del tumor viable más grande (cm) asociado al número de tumores viables y la presencia o no de invasión microvascular en el análisis anatomopatológico del CHC en el hígado nativo del receptor. Con estos factores se establece una puntuación que se correlaciona con el riesgo de recidiva a 5 años del CHC. En función de ese riesgo, se realiza un seguimiento más o menos estricto mediante pruebas de imagen.

VARIABLE PREDICTORA	PUNTOS
AFP (día trasplante) (ng / ml)	
0 – 20	0
21 – 99	1
100 – 999	2
≥ 1000	3
Invasión microvascular	2
Diámetro de tumor viable más grande (cm) + número de tumores viables	
0	0
1,1 – 4,9	1
5,0 – 9,9	2
≥ 100	3

Tabla 10: RETREAT score

RETREAT score (puntos)	Recidiva a 5 años	Seguimiento
0	3 %	Sin pruebas de imagen
1	8 %	TC/RM + TC tórax cada 6 meses (durante 2 años)
2	11 %	
3	14 %	
4	29 %	TC/RM + TC tórax cada 6 meses (durante 5 años)
≥ 5	75 %	TC/RM + TC tórax cada 4 meses (durante 2 años) y posteriormente cada 6 meses (hasta el 5º año)

Tabla 11: Riesgo de recidiva de CHC a 5 años según RETREAT score y seguimiento después del TH

4.8.- Definición de las variables del estudio

De las características preoperatorias del receptor y tumor que pueden influir en la recurrencia precoz del CHC se eligieron aquellas que, tras una revisión bibliográfica actualizada, mostraron cierta implicación en dicho resultado y podían ser fácilmente accesibles para un grupo de trabajo especializado en trasplante hepático.

A partir de una base de datos creada con el programa informático Excel® versión 16.60 para Mac, se recopilaron retrospectivamente las variables que se analizan en el estudio descriptivo y que sirvieron de base posteriormente para la elaboración del modelo predictivo basado en RB.

La variable *resultado*, que representa el objetivo principal del estudio es:

- Recurrencia en los 2 primeros años postrasplante: variable dicotómica (sí/no), definida como el diagnóstico de un nuevo nódulo tumoral con características de CHC detectado en una prueba de imagen y/o con confirmación histológica, durante los dos primeros años transcurridos después del trasplante hepático.

Asimismo, se recogió también la presencia de recurrencia sin criterio temporal, esto es, independientemente del tiempo transcurrido desde el trasplante, para poder determinar la recurrencia y supervivencia libre de enfermedad de los pacientes de la serie. Para

caracterizar mejor los patrones de recurrencia, se dividió el tipo de recidiva en:

- Recurrencia hepática aislada única: definida como la aparición de un nódulo hepático único compatible con CHC después del trasplante.
- Recurrencia hepática aislada múltiple: definida como la aparición de más de un nódulo hepático compatible con CHC después del trasplante.
- Recurrencia extrahepática: definida como la aparición en un órgano distinto al hígado de un nódulo compatible con implante de CHC después del trasplante.
- Recurrencia intra-extrahepática: definida como la aparición combinada de uno o varios nódulos hepáticos compatibles con CHC asociada a la aparición en un órgano distinto al hígado de un nódulo compatible con implante de CHC, después del trasplante.

Otros resultados analizados en el estudio son:

- Supervivencia libre de enfermedad: definida en meses transcurridos desde el trasplante hepático hasta el diagnóstico de recurrencia de CHC.

- Supervivencia global: definida en meses transcurridos desde el trasplante hepático hasta el fallecimiento del paciente, finalización del seguimiento o cierre del estudio.

Con el fin de calcular la supervivencia, se determinaron los siguientes datos:

- Fecha del último control antes del cierre del estudio o fecha del *exitus*, en caso de haberse producido.
- Fecha del diagnóstico de la recurrencia del tumor, definido por la primera prueba de imagen en la que se detectó la recidiva.

A continuación, se describen con detalle cada una de las variables incluidas en el estudio:

A) Variables del receptor:

- Edad del receptor en el momento de la realización del trasplante hepático. Variable continua medida en años cumplidos.
- Género del receptor.
- IMC del receptor: variable continua, cociente del peso expresado en kilogramos (kg) y el cuadrado de la talla, expresada en metros (m) y medido, por tanto, en kg/m².
- Síndrome metabólico: variable categórica dicotómica (sí/no). El síndrome metabólico se define como la presencia de diabetes mellitus o resistencia a la insulina, asociada a dos de

los siguientes factores: obesidad, dislipemia, hipertensión o microalbuminuria (131).

- Hepatopatía de base del receptor: variable categórica en función de la etiología de base de la enfermedad hepática del paciente:
 - Hepatopatía crónica por VHC
 - Hepatopatía crónica por VHB
 - Etílica
 - Otras (cirrosis biliar primaria, colangitis esclerosante primaria, esteatohepatitis no alcohólica...)
- Estadio Child del receptor: variable categórica no dicotómica (A, B y C).
- Puntuación MELD y MELD-Na del receptor: variable continua.
- Cumplimiento de los criterios de Milán por parte del receptor en el momento del trasplante: variable categórica dicotómica (sí/no).
- Cumplimiento de los criterios *Up-to-seven* por parte del receptor en el momento del trasplante: variable categórica dicotómica (sí/no).
- *Downstaging* a Milán: Variable dicotómica (sí/no) en función de si el paciente había experimentado una reducción del tamaño tumoral viable mediante terapia locorregional pretrasplante, permitiendo cumplir los criterios de Milán.

- *Downstaging a Up-to-seven*: Variable dicotómica (sí/no) en función de si el paciente había experimentado una reducción del tamaño tumoral viable mediante terapia locorregional pretrasplante, permitiendo cumplir los criterios *Up-to-seven*.
- Tratamiento del CHC previo al trasplante: variable categórica dicotómica (sí/no).
- Tipo de terapia pretrasplante que recibió el paciente: Dividida en distintas variables categóricas dicotómicas (sí/no) en función de los tratamientos previos al trasplante realizados:
 - RF / MO
 - Alcoholización
 - QETA
 - Radioembolización
 - Quimioterapia
- Número de tratamientos previos sobre el tumor antes del trasplante. Variable categórica discreta.
- Tipo de respuesta del tumor a la terapia previa: variable categórica según los criterios mRECIST:
 - Completa
 - Parcial
 - Enfermedad estable
 - Progresión
- Estado postratamiento: variable categórica que resume el estado del paciente teniendo en cuenta el manejo del CHC durante el periodo pretrasplante:

- Sin tratamiento
- Tratamiento puente
- *Dowstaging Milán*
- *Dowstaging Up-to-seven*
- Fracaso *dowstaging*
- Número de lesiones tumorales detectadas previamente al trasplante en la última prueba de imagen realizada previamente al TH. Variable continua discreta.
- Número de lóbulos hepáticos afectados por el tumor. Variable categórica dicotómica.
- Tamaño del nódulo tumoral más grande: medido en centímetros (variable continua) en las pruebas de imagen pretrasplante.
- Tamaño tumoral total: suma del tamaño de todas las lesiones tumorales en las pruebas de imagen pretrasplante, medida en centímetros (variable continua).
- Ratio neutrófilos / linfocitos: variable continua resultado del cociente entre ambos valores en la analítica del estudio pretrasplante.
- Ratio plaquetas / linfocitos: variable continua resultado del cociente entre ambos valores en la analítica del estudio pretrasplante.
- Alfafetoproteína (AFP): variable continua. Valor de la alfafetoproteína sérica en la analítica pretrasplante, medida en ng/ml.

- AFP > 400 ng/ml y AFP > 1000 ng/ml: expresado como variables categóricas dicotómicas (sí/no). Será positiva si el paciente ha tenido un valor de alfafetoproteína mayor de 400 ng/ml o mayor de 1000 ng/ml en cualquier momento del seguimiento pretrasplante. A efectos prácticos, los pacientes del estudio se categorizaron en una variable secundaria, según el valor máximo de la AFP durante el seguimiento pretrasplante en 3 grupos:
 - AFP < 400 ng/ml
 - AFP $\geq$ 400 ng/ml y < 1000 ng/ml
 - AFP $\geq$ 1000 ng/ml
- Tratamiento antiviral de la hepatitis C: Variable categórica dicotómica (sí / no). Expresa si aquellos pacientes que presentan infección por VHC han recibido tratamiento previamente al trasplante.
- Carga viral VHC: Variable continua para los pacientes con VHC. En aquellos pacientes con carga viral indetectable, se consideró como valor 0.
- Tipo de respuesta al tratamiento del VHC: en aquellos pacientes con infección por VHC que recibieron tratamiento, esta variable categórica expresa el tipo de respuesta a la terapia.
 - Ausencia de respuesta
 - Respuesta viral no sostenida
 - Respuesta viral sostenida

- Reaída
- Respuesta parcial
- Tiempo en lista de espera de trasplante: Variable continua (días).
- Paciente fallecido: Estado del paciente en el momento de la finalización del estudio. Variable dicotómica (sí/no).
- Causa de la muerte del receptor: Variable categórica no dicotómica:
 - No fallecido
 - Fallecido por recurrencia tumoral
 - Fallecido por causa distinta a recurrencia del tumor
- Estado del paciente: Situación del paciente en el momento de la finalización del estudio, expresado como variable categórica no dicotómica:
 - Vivo libre de enfermedad
 - Vivo con recurrencia
 - Fallecido libre de enfermedad
 - Fallecido con recurrencia tumoral

B) Variables del donante:

- Tipo donante: Variable categórica no dicotómica.
 - Muerte cerebral
 - Asistolia controlada
 - Donante vivo

4.9.- Métodos estadísticos del estudio

El estudio estadístico se realizó con la ayuda del programa informático RStudio® versión 4.1.2, que se empleó tanto para el análisis descriptivo de la serie, como para la generación de los modelos predictivos de regresión logística y redes bayesianas.

4.9.1.- *Análisis descriptivo de la serie*

- a) Estudio descriptivo de las variables. Se emplearon las siguientes medidas:
- Estudio de distribución de normalidad de las variables continuas mediante el test de Saphiro-Wilks.
 - Estudio descriptivo de las variables continuas que presentan distribución normal: media y desviación típica.
 - Estudio descriptivo de las variables continuas que presentan distribución no normal: mediana y rango intercuartílico.
 - Estudio descriptivo de las variables categóricas (dicotómicas y no dicotómicas): distribución de frecuencias y porcentaje.

- b) Estudio comparativo entre los grupos de entrenamiento y generalización de las variables analizadas.
 - Comparación de variables continuas: prueba t para muestras independientes y U de Mann-Whitney.
 - Comparación de variables cualitativas: prueba de Chi-Cuadrado.
- c) Estudio de supervivencia acumulada (Kaplan-Meier). En ambos grupos (entrenamiento y generalización), se representó mediante una gráfica que incluye la probabilidad de que no suceda la recurrencia postrasplante del CHC o el fallecimiento del receptor en función del tiempo (en años) transcurrido desde la realización del trasplante.

4.9.2.- Estudio analítico univariante

En primer lugar, se llevó a cabo un estudio univariante con el fin de analizar la posible relación entre cada variable estudiada con la recurrencia precoz del CHC. Se emplearon las siguientes pruebas:

- a) Variables continuas con distribución normal: prueba t de Student.
- b) Variables continuas sin distribución normal: prueba U de Mann-Whitney.
- c) Variables categóricas: prueba Chi-Cuadrado o V de Cramer.

Con el análisis univariante, se consideró que existía una relación estadísticamente significativa entre cada variable en concreto y la recurrencia precoz del CHC en aquellos casos en los que, tras la prueba correspondiente, se obtuvo un valor de $p < 0,1$. Esto es, siempre que una variable obtuvo un p-valor inferior a ese umbral, se consideró candidata a predecir la recurrencia precoz del CHC. Se estableció un umbral conservador con el fin de evitar descartar variables con una influencia potencial en los resultados del estudio.

El estudio univariante permite, por un lado, analizar la importancia individual de cada variable en la predicción de los resultados del estudio y por otro, es de utilidad para determinar si las variables que han resultado significativas mediante el estudio univariante son las mismas que aquellas seleccionadas con la regresión logística utilizando el método de retirada secuencial.

4.9.3.- Estudio analítico multivariante

Si consideramos las variables seleccionadas desde una perspectiva clínica, éstas son razonablemente adecuadas para la representación del problema analizado en el presente estudio. Asimismo, matemáticamente son aceptables tanto para el método de regresión logística como para el de redes bayesianas, que son los empleados para desarrollar los modelos predictivos.

El rendimiento de los modelos en cuanto a capacidad predictiva se representó mediante las correspondientes curvas ROC. Para cada una

de ellas se calculó el ABC y el punto óptimo de corte en la pendiente, evaluando los valores de sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y razón de verosimilitud para test positivo y para test negativo.

4.9.3.1.- Diseño del modelo predictivo basado en regresión logística

Según un análisis multivariante mediante el método de RL, se evaluó qué variables influyen en la recurrencia del CHC en los dos primeros años postrasplante. Para ello, se realizó una recodificación de las variables categóricas, transformándolas en variables dumificadas (*dummy*). Se comprobó que el factor de inflación de la varianza (FIV) era inferior a 2 en todos los casos, con lo que se descartaron problemas importantes de colinealidad entre variables.

a) Software utilizado

Para el diseño del modelo predictivo de recurrencia precoz mediante RL, se ha utilizado el lenguaje de programación *R* y la librería *The R Stats Package (stats)*, en su versión 4.1.2.

b) Diseño del modelo de regresión logística

Con los datos correspondientes al grupo de entrenamiento se ha seguido el procedimiento habitual, determinando los parámetros de cada variable del modelo mediante la

estimación de máxima verosimilitud, utilizando el algoritmo de Newton-Raphson y empleando los límites de convergencia que utiliza la librería por defecto.

Para el desarrollo del modelo de RL se ha empleado el método de retirada secuencial basado en la minimización del criterio de información de *Akaike*, incluyendo sólo las variables significativas ($p < 0,05$) y ligeramente significativas ($p < 0,1$), utilizando el estadístico Wilcoxon Mann Whitney para las variables cuantitativas y el Chi-Cuadrado, para las variables cualitativas (considerando las variables nominales y ordinales como categóricas y dumificándolas).

El modelo predictivo diseñado quedó representado por una fórmula que incorpora las variables incluidas en el modelo multivariante. Su capacidad predictiva quedó plasmada mediante la correspondiente curva ROC, de la que se calculó el ABC.

A partir de la relación coste-beneficio para el problema estudiado se determinó el punto de corte óptimo, de tal manera que el modelo fue evaluado mediante una matriz de confusión, clasificando a todos los individuos de la muestra según la concordancia entre los valores observados y los predichos o estimados por el modelo.

Para finalizar, se calcularon la sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo, razón de

verosimilitud para test positivo y razón de verosimilitud para test negativo.

4.9.3.2.- Diseño del modelo predictivo basado en redes bayesianas

Para el desarrollo del modelo predictivo basado en RB se ha empleado una variante de *Tree Augmented Bayes Network* (TAN) que consta de dos etapas. En la primera, se obtiene un grafo que relaciona las variables predictivas más dependientes creando una estructura en árbol en el que, además, todas las variables quedan unidas a la variable dependiente a predecir (*recurrencia precoz*). El grafo se obtiene mediante el algoritmo de Chow-Liu, que aproxima la distribución conjunta de los datos mediante una aproximación de segundo orden. Con el objetivo de simplificar la interpretación y el uso del modelo predictivo se procedió a la discretización de las variables cuantitativas. En una segunda etapa se describe la relación entre las variables del problema mediante tablas de probabilidad condicionada obtenidas a partir de los datos del conjunto de desarrollo.

El grafo junto con las tablas de probabilidad condicionada define la red bayesiana discreta que permite la inferencia de las probabilidades de cualquier evento de interés, así como modificar estas estimaciones a medida que se van conociendo los valores de las variables independientes.

a) *Software* utilizado

Para el diseño del modelo predictivo de recurrencia precoz mediante RB se ha utilizado el lenguaje de programación R y diferentes librerías específicas: *Classification Based on Association Rules* (arulesCBA, versión 1.2.1) para la discretización supervisada de las variables cuantitativas, *Classification and Regression Training* (caret, versión 6.0-90) para la búsqueda de los hiperparámetros óptimos y *Learning Discrete Bayesian Network* (bnclassify, versión 0.4.6) para la obtención el modelo bayesiano final.

b) Diseño del modelo de redes bayesianas

Partiendo de las mismas variables que el modelo de RL ($p < 0,1$), y utilizando únicamente el conjunto de entrenamiento, se procedió a una discretización supervisada de las variables numéricas mediante el algoritmo *Minimum Description Length Principle* (mdlpl), que determina el valor de corte (midiendo la reducción de la entropía) que hay que aplicar a una variable para mejorar la predicción de la recurrencia. Las variables para las que una discretización no suponía una ventaja fueron descartadas para el modelo definitivo de RB.

Con las variables seleccionadas del grupo de entrenamiento se procedió a obtener los dos parámetros (*score*, *smooth*) que definen la red bayesiana, y que maximizan el ABC según una estrategia de validación cruzada. Por la limitación de registros, se prefirió 2 particiones

y 5 repeticiones (5x2 *fold*) frente a la habitual de 10 particiones (10 *fold*).

La puntuación (*score*) que determina el ajuste del grafo a los datos se varió entre 3 posibilidades: logaritmo de la verosimilitud (*loglik*), criterio de información bayesiano y criterio de información de Akaike. El valor del suavizado de la probabilidad (*smooth*) se varió entre 10^{-5} y 10^3 en escala logarítmica.

Finalmente, los parámetros que mayor ABC mostraron en promedio sobre las diferentes particiones (5x2 *fold*) de los datos de desarrollo se utilizaron para definir el modelo TAN.

Al igual que con el modelo de RL, a partir de la relación coste-beneficio para el problema analizado se determinó el punto óptimo de corte. Se evaluó el modelo de RB mediante una matriz de confusión, y se calculó la sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo, razón de verosimilitud para test positivo y razón de verosimilitud para test negativo.

4.9.4.- Función de utilidad

La función de utilidad se empleó para determinar la pendiente de la curva ROC y, por tanto, el umbral que discrimina de la mejor forma posible entre aquellos pacientes que van a presentar una recurrencia precoz del CHC y los que no. Dependiendo de la importancia relativa que tiene ser verdadero o falso positivo o negativo después de la

aplicación del modelo predictivo, la función de utilidad permite determinar el punto de corte óptimo de la curva ROC.

Teniendo en cuenta el elevado número de trasplantes realizados con intención oncológica y en base al objetivo del estudio, con el que se pretende elaborar un modelo predictivo que identifique de la forma más precisa posible aquellos casos en los que se va a producir una recurrencia precoz del CHC, se han establecido los siguientes valores de utilidad (entre 1, que sería la situación de máxima utilidad y -1, que representaría la situación menos beneficiosa) para cada uno de los posibles resultados de la aplicación del modelo predictivo:

1.- Falso positivo: situación en la que se considera que el modelo predice que el paciente sí ha presentado recurrencia precoz cuando realmente el evento no se ha producido. Este error considera que se producirá una recurrencia en los dos primeros años después del TH sin que ello se corresponda con la realidad. Sin embargo, a pesar del error de la prueba, si el test fuera aplicado y el paciente no llega a recibir un trasplante hepático, el injerto se destina a otro receptor, de tal manera que el órgano no se pierde. Sin embargo, si hacemos caso a la prueba, el paciente no se beneficia del TH y, por tanto, este error tiene consecuencias malas para el paciente. Le asignamos un valor de 0,25 porque a pesar de que podría recibir otros tratamientos, el mejor en términos oncológicos es el trasplante.

Utilidad falso positivo: 0,25

2.- Falso negativo: el modelo considera que el paciente no presentará recurrencia precoz cuando realmente ésta sí se producirá. Realmente es un error muy grave y debería evitarse por todos los medios, ya que predecir que el paciente va a estar libre de enfermedad cuando no es así va en contra de la optimización de los injertos disponibles, motivo principal del estudio. Se supone que debería asignarse una importancia considerable a esta situación, con el fin de reducir el número de falsos negativos. Las consecuencias de esta situación son nefastas, dados los malos resultados de supervivencia en pacientes con recurrencia postrasplante del CHC, necesidad de tratamientos para controlar la enfermedad, etc., ya que se estima que no va a haber recidiva en 2 años y efectivamente el órgano se pierde.

Utilidad falso negativo: - 0,9

3.- Verdadero negativo: el modelo considera que el paciente no presentará recurrencia y realmente ésta no se produce durante los dos primeros años después del TH. Este valor tiene bastante importancia desde el punto de vista de la eficiencia si lo consideramos antes de que se produzca el TH, ya que identifica a pacientes sin recurrencia precoz. Por tanto, en esta situación el injerto se destina a un paciente con bajo riesgo de recurrencia precoz y que tiene más probabilidad de obtener una supervivencia libre de enfermedad y global más elevada. Por todo ello se le asigna un valor bastante elevado.

Utilidad verdadero negativo: 0,85

4.- Verdadero positivo: el modelo considera que el paciente sí presentará recurrencia precoz postrasplante y realmente sí se produce dicha recurrencia. Este es el objetivo principal de nuestro estudio: detectar con un modelo basado en variables disponibles previamente al TH la probabilidad de recurrencia del CHC durante los dos primeros años después del TH. Este valor nos permite identificar aquellos pacientes en los que el TH no va a ser eficiente y, en la práctica clínica, intentamos evitar esta situación para no obtener resultados negativos en cuanto a la recurrencia y supervivencia libre de enfermedad. La utilidad concedida a esta situación es máxima, ya que evita que se realice un TH en aquellos pacientes que realmente van a presentar una recurrencia precoz, permitiendo que esos injertos se asignen a otros pacientes con mejores resultados a medio y largo plazo.

Utilidad verdadero positivo: 1

Se determinó la pendiente de la curva ROC en cada punto ajustando una curva paramétrica binormal para la RL y no paramétrica para la RB. Para decidir el punto de corte óptimo se emplearon la probabilidad preprueba y las utilidades, calculando el coste y el beneficio netos.

Se calculó el COSTE NETO (pacientes con verdadera recurrencia en los dos primeros años después del TH), es decir, la diferencia entre utilidad de trasplantar a pacientes con verdadera NO recurrencia

precoz y utilidad de no trasplantar a pacientes con falsa recurrencia precoz (UVN – UFP).

$$\text{COSTE NETO: } 0,85 - (0,25) = 0,6$$

Se calculó el BENEFICIO NETO (aquellos que realmente no van a presentar recurrencia precoz), es decir, la diferencia entre la utilidad de no trasplantar a pacientes con verdadera recurrencia precoz y la utilidad de trasplantar a pacientes con falsa ausencia de recurrencia precoz (UVP – UFN).

$$\text{BENEFICIO NETO: } 1 - (-0,9) = 1,9$$

4.9.5.- Generalización de los modelos desarrollados

La generalización del modelo predictivo elaborado mediante RB se llevó a cabo en base a una muestra diferente a la del grupo de entrenamiento, de 130 pacientes trasplantados en el Hospital Universitari i Politècnic La Fe de Valencia.

Del mismo modo que en el grupo de entrenamiento, se determinó su matriz de confusión, sensibilidad, especificidad, valores predictivos y razón de verosimilitud.

4.9.6.- Comparación de los modelos desarrollados

Se comparó la capacidad predictiva de los modelos de regresión logística y redes bayesianas en base a la medición de la ABC de la curva ROC de cada modelo, según el método de Hanley y McNeil (132), utilizando el software R studio, versión 4.1.2.

$$z = \frac{A_1 - A_2}{\sqrt{SE_1^2 + SE_2^2 - 2\rho SE_1 SE_2}}$$

En esta fórmula, A_1 y SE_1 representan el ABC y su correspondiente error estándar para la curva ROC del modelo de redes bayesianas, mientras que A_2 y SE_2 representan las mismas variables para el modelo de regresión logística. La ecuación se completa determinando el coeficiente “promedio” de correlación entre las ABC (ρ), calculado según el método propuesto por Hanley y McNeil (132). El resultado de la ecuación se conoce como estadístico z , que tiene una distribución normal si las áreas de los modelos son iguales. Por tanto, con un intervalo de confianza del 95 %, se consideró que existen diferencias estadísticamente significativas entre ambos modelos cuando los valores del estadístico z se encuentren fuera del intervalo (-1,96 – 1,96).

Finalmente, la comparación entre ambos modelos predictivos se completó con el análisis de la concordancia de los resultados según el coeficiente de correlación Kappa (K). En función del valor de K se

puede categorizar el grado de concordancia entre los resultados de dos medidas, como se expone en la siguiente tabla:

Valor de K	Grado de concordancia
< 0,20	Muy bajo
0,21 – 0,40	Bajo
0,41 – 0,60	Moderado
0,61 – 0,80	Bueno
0,81 – 1	Muy bueno

Tabla 12: Relación entre el valor del coeficiente K y el grado de concordancia entre 2 medidas

4.10.- Interfaz informática para la aplicación clínica del modelo predictivo

Este estudio tiene un objetivo eminentemente práctico. Teniendo en cuenta este aspecto, se pretende desarrollar una interfaz informática de acceso abierto vía web que permita a los grupos de trasplante utilizar el modelo como una herramienta de ayuda en la toma de decisiones clínicas, determinando el riesgo de recurrencia precoz del CHC en cada caso concreto.

Esta interfaz informática será desarrollada en colaboración con el codirector de esta tesis, el profesor Antonio J. Serrano, miembro del *Intelligent Data Analysis Laboratory* (IDAL) de la Escuela Técnica Superior de Ingeniería de la Universidad de Valencia. Se utilizará el lenguaje de programación *R* y la librería *shiny* (*Web Application Framework*) versión 1.7.4. Posteriormente, será implementada por la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Hospital Universitari i Politècnic La Fe de Valencia.

Resultados

5.- Resultados

5.1.- Análisis descriptivo de la población del estudio

Durante el período de estudio (1 de enero de 2004 a 31 de diciembre de 2018) se realizaron 515 trasplantes con indicación de CHC de forma consecutiva en la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Hospital Universitari i Politècnic La Fe de Valencia. No se excluyó ningún paciente de la serie inicial.

En la siguiente figura se representa el número de trasplantes anuales realizados en el período de estudio.

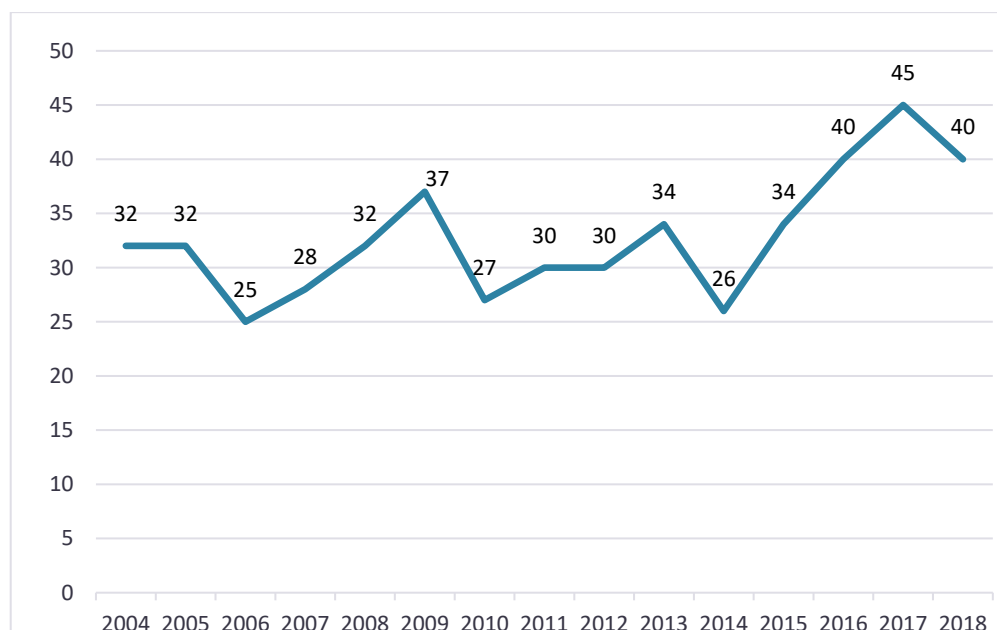


Figura 10: Número de trasplantes realizados anualmente por CHC en la Unidad de Cirugía Hepatobiliopancreática y Trasplante. Hospital Universitari i Politècnic La Fe

Como podemos comprobar, se trata de una cifra relativamente estable en torno a los 30 casos de media al año, con un incremento más o menos progresivo que ha permitido alcanzar cifras de hasta 45 casos de CHC trasplantados en años concretos.

5.1.1.- Análisis descriptivo de las características de los receptores

En las siguientes tablas se exponen los datos referentes a la serie completa de los pacientes incluidos en el estudio. En primer lugar, se exponen las características de base de los receptores, seguidas de las características de la hepatopatía de base de los receptores, así como el manejo de la infección por VHC y la respuesta al tratamiento.

Variable	n = 515
	Mediana (RIC) / Frecuencia (%)
Edad (años)	59 (54 – 63)
Sexo	
• Varón	435 (84,46)
• Mujer	80 (15,53)
IMC (kg / m ²)	27,65 (24,90 – 30,40)
Síndrome Metabólico	
• Sí	115 (22,33)
Ratio N/L	2,19 (1,55 – 3,55)
Ratio P/L	70,79 (50,12 – 107,15)
Tiempo en lista de espera (días)	131 (53 – 221)

N: Neutrófilos

L: Linfocitos

P: Plaquetas

Tabla 13: Estudio descriptivo de las características de base de los pacientes incluidos en el estudio

Variable	n = 515
	Mediana (RIC) / Frecuencia (%)
Causa de hepatopatía	
• VHC	316 (61,36)
• Etílica	119 (23,11)
• VHB	41 (7,96)
• Otras	39 (7,57)
Child	
• A	266 (51,65)
• B	165 (32,05)
• C	84 (16,31)
MELD-UNOS / MELD-Na	11 (8 – 15) / 12 (9 – 17)
Carga viral VHC (UI/mL)	147,91 (0 - 208929)
Tratamiento anti-VHC	
• No	97 (18,83)
• Sí	219 (42,52)
• No procede	199 (38,65)
Respuesta tratamiento VHC	
• No procede	296 (57,48)
• Respuesta viral sostenida	91 (17,67)
• RV no sostenida	22 (4,27)
• Recaida	19 (3,69)
• Respuesta parcial	11 (2,14)
• No respuesta	76 (14,75)

Tabla 14: Estudio descriptivo de las variables relacionadas con la hepatopatía de base de los pacientes incluidos en el estudio

A continuación, se exponen los datos referentes a la carga tumoral y niveles de AFP de los 515 pacientes incluidos en el estudio.

Variable	n = 515
	Mediana (RIC) / Frecuencia (%)
Número de lesiones	
• 1	272 (52,82)
• 2	147 (28,54)
• 3	72 (13,98)
• 4	15 (2,91)
• 5 o más	9 (1,75)
Tamaño máximo tumoral (cm)	2,70 (2,00 – 3,50)
Tamaño tumoral total (cm)	3,50 (2,50 – 5,15)
Lóbulos afectados	
• 1	398 (77,28)
• 2	117 (22,72)
AFP > 400 ng/ml	
• Sí	42 (8,16)
AFP > 1000 ng/ml	
• Sí	25 (4,85)
AFP pretrasplante (ng/ml)	8,91 (4,17 – 14,80)

Tabla 15: Estudio descriptivo de las variables referentes a la carga tumoral de los pacientes incluidos en el estudio

En las dos tablas siguientes se pueden observar los datos referentes al tratamiento pretrasplante del CHC de los pacientes del estudio, la modalidad terapéutica y su respuesta, así como el tipo de donante del que proceden los injertos empleados para el TH.

Variable	n = 515
	Mediana (RIC) / Frecuencia (%)
Terapia previa	
• Sí	391 (75,92)
Nº de tratamientos previos	
• 0	124 (24,08)
• 1	229 (44,47)
• 2	101 (19,61)
• 3	38 (7,38)
• 4	17 (3,30)
• 5 o más	6 (1,16)
Milán	
• Sí	487 (94,56)
Up-to-seven	
• Sí	509 (98,83)
Dowstaging Milán	
• Sí	65 (12,62)
Dowstaging Up-to-seven	
• Sí	24 (4,66)
RF	
• Sí	137 (26,6)
QETA	
• Sí	297 (57,66)

Tabla 16: Estudio descriptivo de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (I)

Variable	n = 515
	Mediana (RIC) / Frecuencia (%)
Alcoholización	
• Sí	17 (3,30)
Radioembolización	
• Sí	7 (1,36)
Respuesta	
• No procede	124 (24,08)
• Completa	193 (37,48)
• Parcial	134 (26,02)
• Estable	30 (5,83)
• Progresión	34 (6,60)
Estatus tratamiento	
• Sin tratamiento	124 (24,08)
• Tratamiento puente	316 (61,36)
• <i>Dowstaging</i> Milán	65 (12,62)
• <i>Downstaging Up-to-seven</i>	5 (0,97)
• Fallo <i>dowstaging</i>	5 (0,97)
Tipo de donante	
• Muerte encefálica	490 (95,16)
• Asistolia	25 (4,84)

Tabla 17: Estudio descriptivo de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (II)

Por último, se puede observar la tabla que incluye los datos descriptivos referentes a la evolución postrasplante de los pacientes del estudio (recurrencia y mortalidad), así como el tipo de recurrencia y causas de mortalidad.

Variable	n = 515
	Mediana (RIC) / Frecuencia (%)
Recurrencia	
• Sí	84 (16,31)
Recurrencia precoz (< 2 años)	
• Sí	54 (10,49)
Fallecido	
• Sí	219 (42,52)
Causa de la muerte	
• No fallecido	296 (57,48)
• Recurrencia	64 (12,43)
• Otra causa	155 (30,09)
Estatus	
• Vivo libre de enfermedad	280 (54,37)
• Vivo con recurrencia	16 (3,11)
• Fallecido libre de enfermedad	151 (29,32)
• Fallecido con recurrencia	68 (13,20)
Localización de la recurrencia	
• Sin recurrencia	431 (83,69)
• Hepática única	11 (2,14)
• Hepática múltiple	11 (2,14)
• Extrahepática	48 (9,32)
• Intra-extrahepática	14 (2,71)

Tabla 18: Estudio descriptivo de las variables referentes a la evolución postrasplante de los pacientes del estudio (recurrencia y mortalidad)

5.1.2.- Análisis comparativo de las características de los receptores del grupo de entrenamiento frente al grupo de generalización

En las próximas tablas se exponen los datos comparativos del análisis descriptivo de los grupos del estudio: entrenamiento (n = 385) frente a generalización (n = 130). En las dos primeras tablas, se puede observar el análisis comparativo de las características de base de los receptores, así como las características de su hepatopatía de base y el manejo de la infección por VHC y la respuesta al tratamiento.

Variable	Mediana (RIC) / Frecuencia (%)		p
	Entrenamiento (n = 385)	Generalización (n = 130)	
Edad (años)	58 (53 – 63)	60 (55 – 63)	0,05
Sexo			
• Varón	330 (85,72)	105 (80,77)	0,207
• Mujer	55 (14,28)	25 (19,23)	
IMC (kg / m ²)	27,44 (24,60 – 30,10)	28,22 (25,54 – 30,83)	0,01
Síndrome Metabólico			
• Sí	82 (21,29)	33 (25,38)	0,332
Ratio N/L	2,09 (1,41 – 3,31)	2,57 (1,91 – 3,80)	0,01
Ratio P/L	66,07 (47,86 – 100)	83,18 (60,26 – 128,82)	< 0,001
Tiempo en lista de espera (días)	147 (73 – 247)	63,5 (21,25 – 161)	< 0,001

N: Neutrófilos

L: Linfocitos

P: Plaquetas

Tabla 19: Estudio descriptivo comparativo de las variables referentes a las características de base del grupo de entrenamiento frente al grupo de generalización

Variable	Mediana (RIC) / Frecuencia (%)		p
	Entrenamiento (n = 385)	Generalización (n = 130)	
Causa de hepatopatía			
• VHC	241 (62,59)	75 (57,69)	0,257
• Etílica	88 (22,86)	31 (23,84)	
• VHB	32 (8,31)	9 (6,92)	
• Otras	24 (6,24)	15 (11,55)	
Child			
• A	183 (47,53)	83 (63,84)	0,004
• B	130 (33,76)	35 (26,92)	
• C	72 (18,71)	12 (9,23)	
MELD-UNOS / MELD-Na	11 (9 – 16) / 13 (10 – 18)	9 (7 – 12) / 10 (8 – 14)	0,01
Carga viral VHC (UI/mL)	388397 (1134991)	0 (0)	< 0,001
Tratamiento Anti-VHC			
• No	90 (23,37)	7 (5,38)	0,0005
• Sí	151 (39,22)	68 (52,31)	
• No procede	144 (37,41)	55 (42,31)	
Respuesta tratamiento VHC			
• No procede	234 (60,78)	62 (47,69)	0,0004
• Respuesta viral sostenida	34 (8,83)	57 (43,85)	
• RV no sostenida	21 (5,45)	1 (0,76)	
• Recaída	14 (3,64)	5 (3,85)	
• Respuesta parcial	11 (2,86)	0 (0,00)	
• No respuesta	71 (18,44)	5 (3,85)	

Tabla 20: Estudio descriptivo comparativo de las variables referentes a la hepatopatía de base del grupo de entrenamiento frente al grupo de generalización

A continuación, se expone el análisis comparativo de las variables referentes a la carga tumoral y niveles de AFP de los grupos de estudio.

Variable	Mediana (RIC) / Frecuencia (%)		p
	Entrenamiento (n = 385)	Generalización (n = 130)	
Número de lesiones			
• 1	209 (54,29)	63 (48,46)	0,2534
• 2	106 (27,53)	41 (31,54)	
• 3	52 (13,51)	20 (15,38)	
• 4	11 (2,86)	4 (3,08)	
• 5 o más	7 (1,81)	2 (1,54)	
Tamaño máximo tumoral (cm)	2,80 (2,10 – 3,50)	2,50 (2,00 – 3,48)	0,09
Tamaño tumoral total (cm)	3,60 (2,60 – 5,00)	3,50 (2,50 – 5,20)	0,43
Lóbulos afectados			
• 1	300 (77,92)	98 (75,38)	0,5471
• 2	85 (22,08)	32 (24,62)	
AFP > 400 ng/ml			
• Sí	33 (8,57)	9 (6,92)	0,711
AFP > 1000 ng/ml			
• Sí	20 (5,19)	5 (3,85)	0,6425
AFP pretrasplante (ng/ml)	10,23 (4,90 – 38,02)	5,62 (3,31 – 15,49)	< 0,001

Tabla 21: Estudio descriptivo comparativo de las variables referentes a la carga tumoral de los pacientes del grupo de entrenamiento frente al grupo de generalización

Asimismo, en las dos próximas tablas se exponen los datos referentes al análisis comparativo entre los grupos de estudio de las variables referentes al tratamiento pretrasplante del CHC, la modalidad terapéutica y su respuesta, así como el tipo de donante.

Variable	Mediana (RIC) / Frecuencia (%)		p
	Entrenamiento (n = 385)	Generalización (n = 130)	
Terapia previa			
• Sí	293 (76,10)	98 (75,38)	0,9057
Nº de tratamientos previos			
• 0	92 (23,90)	32 (24,62)	0,7536
• 1	172 (44,68)	57 (43,85)	
• 2	73 (18,95)	28 (21,54)	
• 3	30 (7,79)	8 (6,15)	
• 4	13 (3,38)	4 (3,08)	
• 5 o más	5 (1,30)	1 (0,76)	
Milán			
• Sí	361 (93,77)	126 (96,92)	0,2614
Up-to-seven			
• Sí	379 (98,44)	130 (100,00)	0,345
Dowstaging Milán			
• Sí	46 (11,95)	19 (14,62)	0,4462
Dowstaging Up-to-seven			
• Sí	16 (4,16)	8 (6,15)	0,3426
RF			
• Sí	99 (25,71)	38 (29,23)	0,4246

Tabla 22: Estudio descriptivo comparativo de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (I)

Variable	Mediana (RIC) / Frecuencia (%)		p
	Entrenamiento (n = 385)	Generalización (n = 130)	
QETA			
• Sí	231 (60,00)	66 (50,77)	0,0807
Alcoholización			
• Sí	16 (4,16)	1 (0,77)	0,0851
Radioembolización			
• Sí	2 (0,52)	5 (3,84)	0,0129
Respuesta			
• No procede	92 (23,90)	32 (24,62)	0,0939
• Completa	134 (34,81)	59 (45,38)	
• Parcial	106 (27,53)	28 (21,54)	
• Estable	27 (7,01)	3 (2,31)	
• Progresión	26 (6,75)	8 (6,15)	
Estado tratamiento			
• Sin tratamiento	92 (23,90)	32 (24,62)	0,7596
• Tratamiento puente	238 (61,82)	78 (60,00)	
• Downstaging Milán	46 (11,95)	19 (14,62)	
• Downstaging Up-to-seven	4 (1,04)	1 (0,76)	
• Fallo downstaging	5 (1,29)	0 (0)	
Tipo de donante			
• Muerte encefálica	385 (100)	105 (80,76)	< 0,001
• Asistolia	0 (0)	25 (19,23)	

Tabla 23: Estudio descriptivo comparativo de las variables referentes al tratamiento pretrasplante del CHC, respuesta y tipo de donante (II)

Finalmente, a continuación, se expone la tabla que incluye los datos referentes al análisis comparativo entre los grupos de la evolución postrasplante (recurrencia y mortalidad), así como el tipo de recurrencia y causas de mortalidad.

Variable	Mediana (RIC) / Frecuencia (%)		p
	Entrenamiento (n = 385)	Generalización (n = 130)	
Recurrencia			
• Sí	70 (18,18)	14 (10,77)	0,0543
Recurrencia precoz (< 2 años)			
• Sí	44 (11,43)	10 (7,69)	0,2514
Fallecido			
• Sí	189 (49,09)	30 (23,08)	< 0,001
Causa de la muerte			
• No fallecido	196 (50,91)	100 (76,92)	0,0005
• Recurrencia	56 (14,54)	8 (6,15)	
• Otra causa	133 (34,55)	22 (16,93)	
Estado			
• Vivo libre de enfermedad	186 (48,31)	94 (72,31)	0,0005
• Vivo con recurrencia	10 (2,60)	6 (4,62)	
• Fallecido libre de enfermedad	129 (33,51)	22 (16,93)	
• Fallecido con recurrencia	60 (15,58)	8 (6,14)	
Localización de la recurrencia			
• Sin recurrencia	315 (81,82)	116 (89,23)	0,3908
• Hepática única	10 (2,60)	1 (0,77)	
• Hepática múltiple	10 (2,60)	1 (0,77)	
• Extrahepática	39 (10,13)	9 (6,92)	
• Intra-extrahepática	11 (2,85)	3 (2,31)	

Tabla 24: Estudio descriptivo comparativo de las variables referentes a la evolución postrasplante (recurrencia y mortalidad)

5.1.3.- Estudio descriptivo de supervivencia y recurrencia del carcinoma hepatocelular en los receptores

La supervivencia global de la serie fue del 85,8 % en el primer año después del TH. A los 3 años postrasplante, la supervivencia global fue del 76,6 % y a los 5 años, del 66,9 %. La mediana de seguimiento fue de 4,9 años (IC 95 %: 4,3 – 5,6).

A continuación, se muestra la curva de supervivencia global acumulada de Kaplan-Meier.

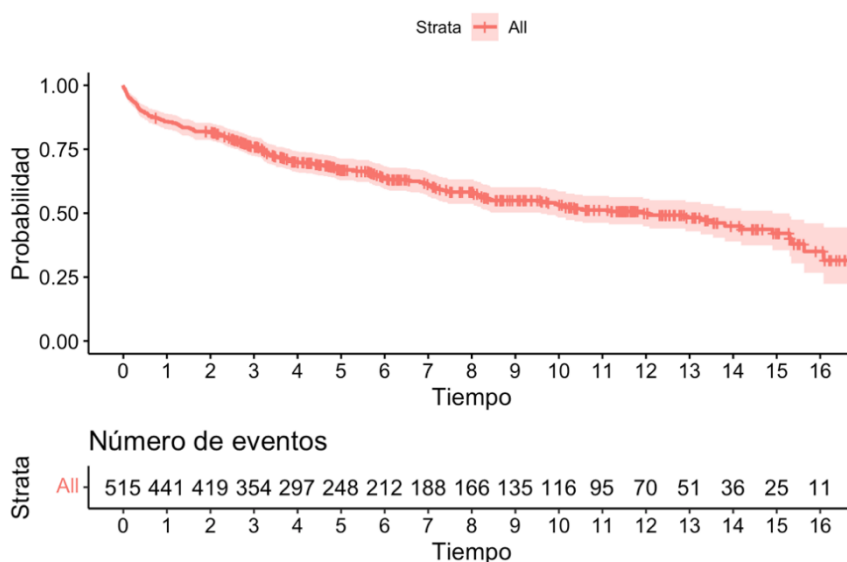


Figura 11: Curva de supervivencia global postrasplante de los pacientes incluidos en el estudio

No se objetivaron diferencias estadísticamente significativas en la supervivencia global entre el grupo de entrenamiento y el grupo de generalización ($p = 0,33$), que fue del 84,4 % y 89,2 % respectivamente al año del TH, del 74,8 % y 78,8 % a los 3 años y del 66,3 % y 66,0 % a los 5 años postrasplante.

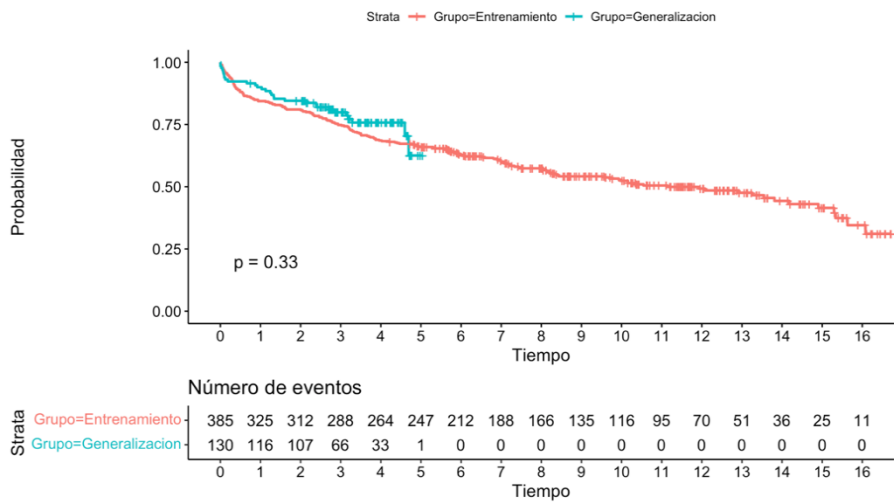


Figura 12: Análisis comparativo de supervivencia global postrasplante de los pacientes del grupo de entrenamiento frente al grupo de generalización

La supervivencia libre de enfermedad de los pacientes del estudio fue del 83,3 % el primer año después del TH, del 71,2 % a los 3 años y del 63,2 % al quinto año postrasplante. (Figura 13)

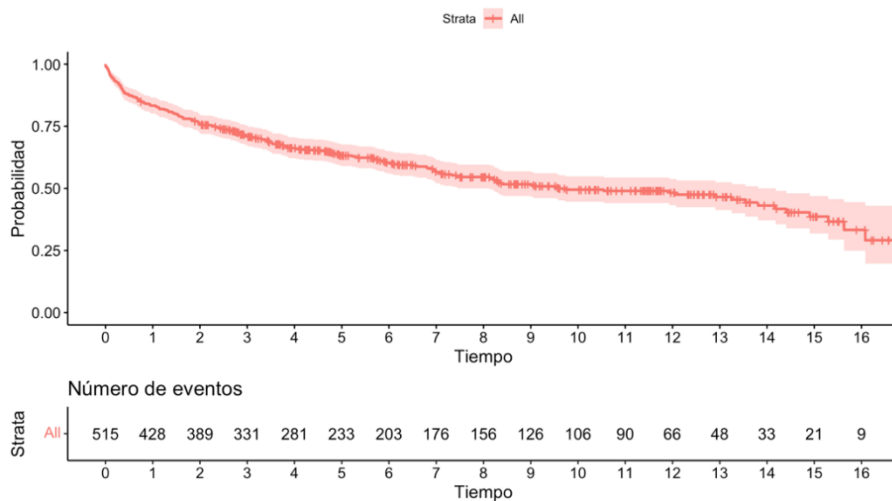


Figura 13: Curva de supervivencia libre de enfermedad postrasplante de los pacientes incluidos en el estudio

Comparando el grupo de entrenamiento y el de generalización, la supervivencia libre de enfermedad en el primer año postrasplante fue del 82,1 % y del 86,9 % respectivamente, del 69,9 % y del 75 % al tercer año postrasplante, y del 62,3 % y 65,4 % a los 5 años del TH. No existieron diferencias estadísticamente significativas en la supervivencia libre de enfermedad entre los grupos de estudio ($p = 0,39$).

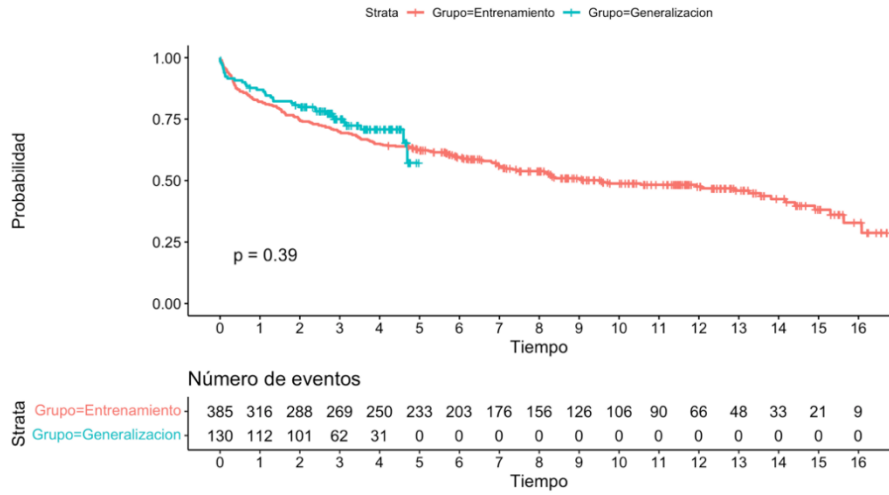


Figura 14: Análisis comparativo de supervivencia libre de enfermedad postrasplante de los pacientes del grupo de entrenamiento frente al grupo de generalización

De los 515 pacientes incluidos en el estudio, 84 presentaron una recurrencia del CHC en algún momento del seguimiento, lo que sitúa la recurrencia global en el 16,3 %. Si comparamos ambos grupos, 70 pacientes presentaron recidiva en el grupo de entrenamiento (18,2 %) y 14 en el grupo de generalización (10,4%) sin diferencias estadísticamente significativas entre grupos ($p = 0,5$).

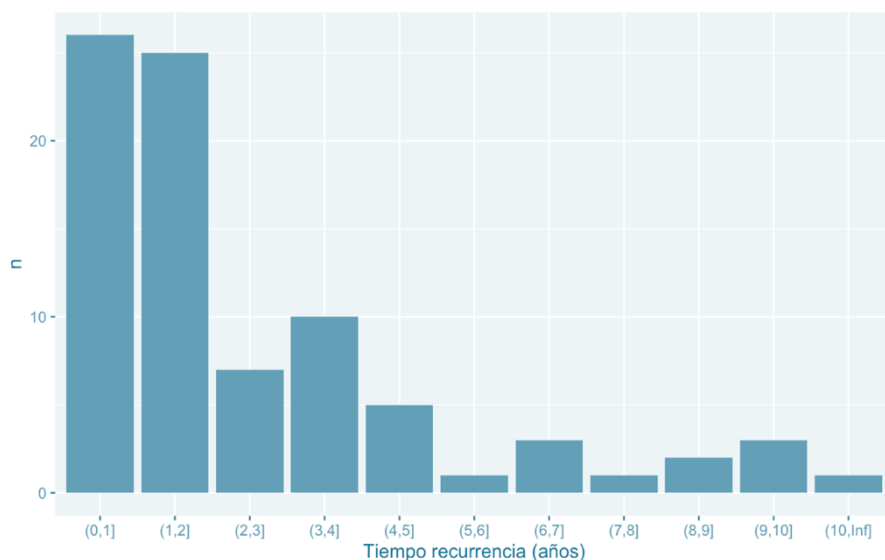


Figura 15: Función de incidencia de recurrencia postrasplante del CHC de los pacientes incluidos en el estudio

La recurrencia precoz (definida como la detección de un nódulo tumoral en pruebas de imagen y/o con confirmación histológica durante los dos primeros años después del TH) se presentó en un total de 54 pacientes (10,5 %). La recurrencia precoz en el grupo de entrenamiento fue del 11,4 % ($n = 44$) y del 7,7 % en el grupo de generalización ($n = 10$), sin diferencias entre grupos ($p = 0,25$).

La presencia de una recurrencia precoz tuvo, en aquellos pacientes que la presentaron, un impacto muy negativo en la supervivencia global, situándose en el 32 % en los 3 primeros años después del trasplante hepático y en el 16,4 % a 5 años, frente a la de

los pacientes sin recurrencia que fue del 79,4 % a los 3 años del TH y del 68,4 % a los 5 años del TH.

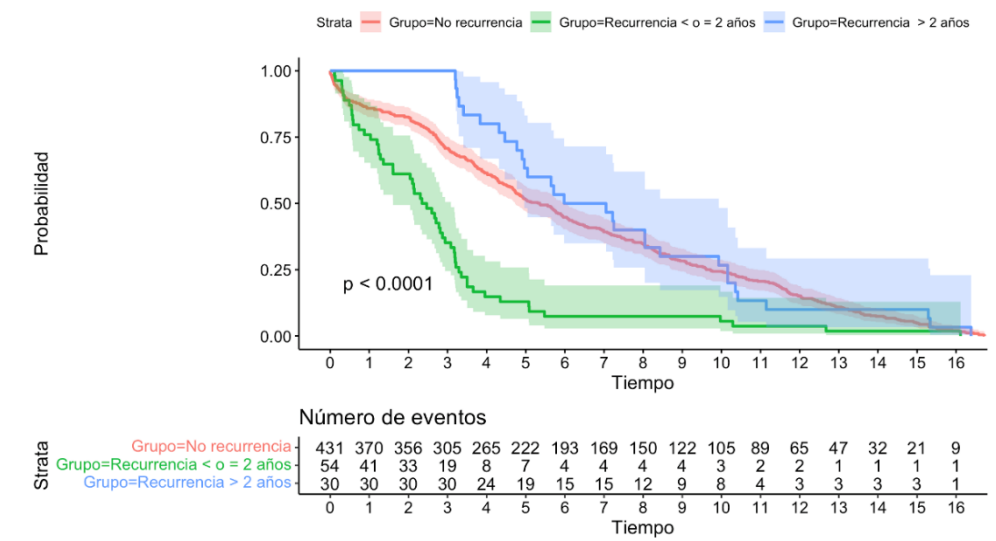


Figura 16: Análisis comparativo de la supervivencia global entre el grupo con recurrencia precoz postrasplante (< 2 años) frente al grupo con recurrencia > 2 años y sin recurrencia

5.2.- Estudio analítico del grupo de entrenamiento

5.2.1.- Estudio univariante de la relación de las variables pretrasplante con la recurrencia precoz postrasplante del carcinoma hepatocelular

Las variables que tras la realización del estudio univariante mostraron una relación estadísticamente significativa ($p < 0,1$) con la recurrencia precoz del CHC después del trasplante hepático se muestran en las siguientes figuras.

A efectos prácticos, se han dividido en cuantitativas y cualitativas. Las primeras se han ordenado en función de su capacidad de discriminación con respecto a la variable objetivo (recurrencia precoz) expresada como ABC de la curva ROC, además de determinarse su significación estadística (valor p) expresada como comparación de medianas entre los dos grupos de estudio (recurrencia precoz / no recurrencia precoz) según el test de Wilcoxon. Del mismo modo, las variables cualitativas también se han ordenado en función de su ABC y se ha expresado su significación estadística según el test V de Cramer.

Así pues, las variables que mostraron relación estadísticamente significativa con la recurrencia precoz postrasplante del CHC en el estudio univariante fueron:

- Respuesta a terapia locorregional pretrasplante
- Valor máximo de AFP durante el seguimiento pretrasplante

- Terapia previa al trasplante sobre el CHC (sí/no)
- Estado después del tratamiento pretrasplante
- *Downstaging* a criterios de Milán (sí/no)
- QETA pretrasplante (sí/no)
- Logaritmo en base 10 del valor de la AFP previa al TH
- Edad del receptor
- Tamaño del nódulo tumoral más grande
- Número de tratamientos sobre el CHC

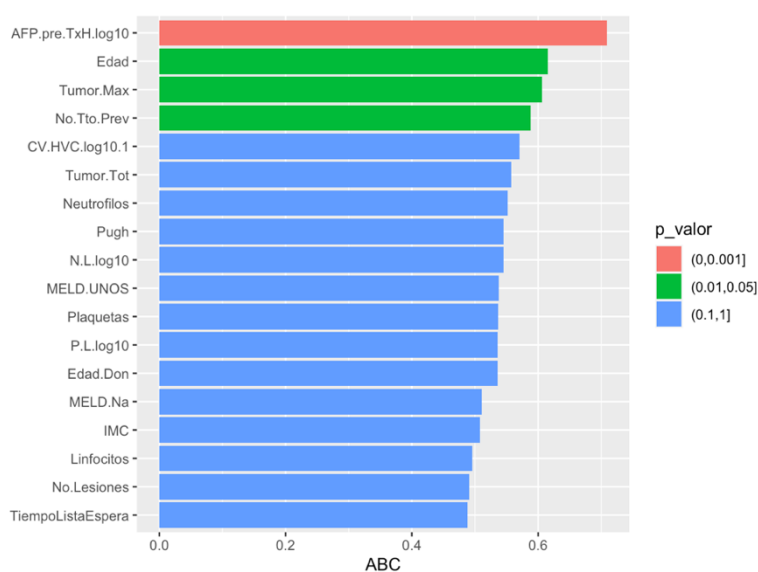


Figura 17: Resultados del análisis univariante de las variables cuantitativas relacionadas con la recurrencia precoz del CHC postrasplante

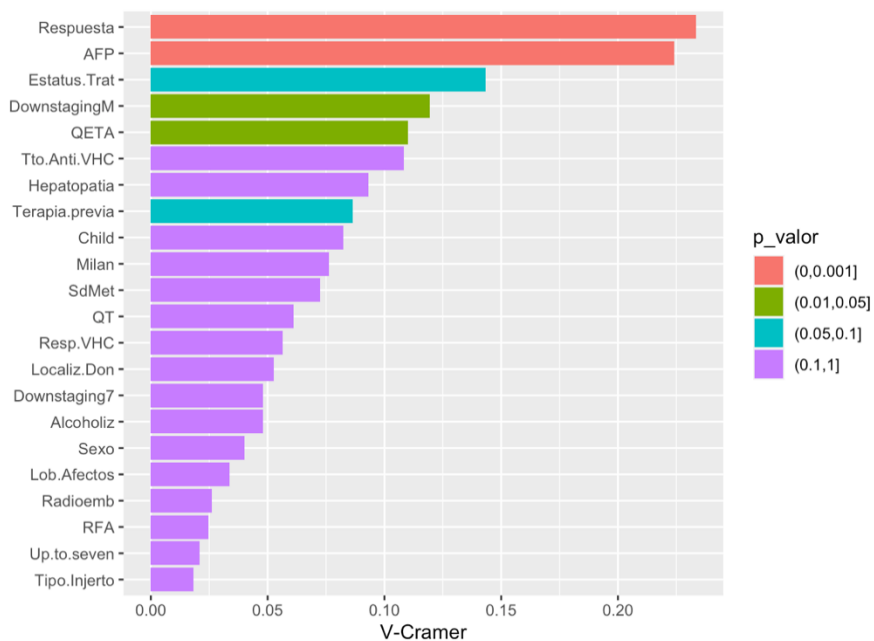


Figura 18: Resultados del análisis univariante de las variables cualitativas relacionadas con la recurrencia precoz del CHC postrasplante

5.2.2.- Estudio multivariante de la relación de las variables pretrasplante con la recurrencia precoz postrasplante del carcinoma hepatocelular

5.2.2.1. – Modelo de regresión logística

Después de 6 pasos se obtuvo el modelo predictivo de regresión logística para estimar la probabilidad de recurrencia precoz postrasplante del CHC. Las variables que incluye este modelo son:

- Edad del receptor
- Logaritmo en base 10 de del valor de la AFP previa al TH
- Tamaño del nódulo tumoral más grande
- Respuesta a terapia pretrasplante

Dichas variables se incorporan al modelo asociadas a una serie de coeficientes según la siguiente fórmula:

$$P(Y=1 | X) = \frac{1}{1 + e^{-z}}$$

donde $z = - 8,084 + (\text{edad} \times (0,055)) + (\text{AFP pretrasplante} \log 10 \times (0,969)) + (\text{tamaño nódulo tumoral} \times (0,340)) + (\text{respuesta a tratamiento pretrasplante} \times (\text{coeficiente según tipo de respuesta}))$

Variable	B	E.T	Wald	gl	Sig.	Exp (B)	IC 95% Inferior	IC 95% Superior
Edad del receptor	0.055	0.029	3.439	1	0.064	1.056	0.997	0.997
Tamaño nódulo mayor	0.340	0.145	5.496	1	0.019	1.405	1.057	1.867
AFP pretrasplante (log 10)	0.969	0.224	18.742	1	< 0,001	2.636	1.700	4.087
Respuesta a tratamiento locorregional			11.838	4	0.019			
- Completa	-0.437	0.626	0.488	1	0.485	0.646	0.189	2.202
- Parcial	0.949	0.526	3.258	1	0.071	2.584	0.922	7.244
- Estable	1.326	0.666	3.961	1	0.047	3.767	1.020	13.911
- Progresión	0.866	0.683	1.606	1	0.205	2.377	0.623	9.073
Constante	-8.084	1.827	19.567	1	< 0,001	< 0,001	-	-

Tabla 25: Relación de variables y constante del modelo de regresión logística en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

En la siguiente figura se ilustra, mediante una curva ROC, la capacidad discriminatoria del modelo de regresión logística para determinar la probabilidad de recurrencia precoz postrasplante del CHC en el grupo de entrenamiento.

El ABC es de 0,791 (IC 95 %: 0,717 – 0,864).

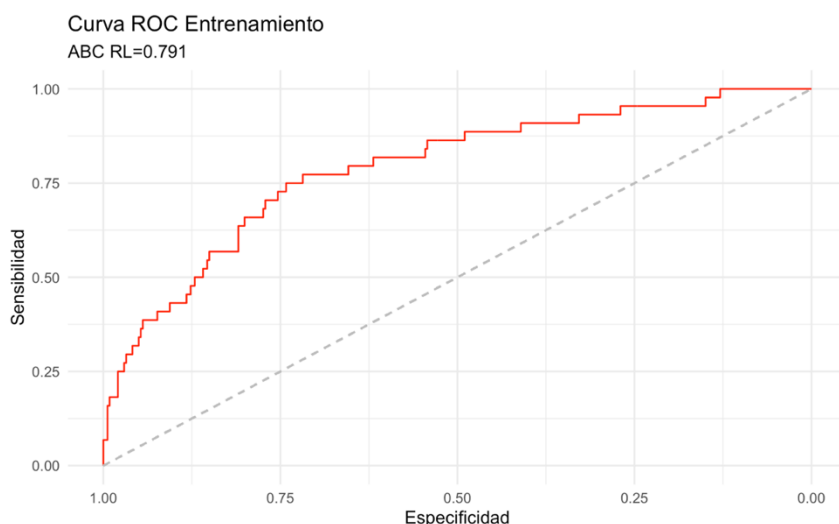


Figura 19: Curva ROC obtenida con el modelo de regresión logística en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

Con los valores de utilidad previamente definidos se obtuvo un umbral para la curva ROC de 0,281 para discriminar entre recurrencia precoz (< 2 años) y no recurrencia precoz (> 2 años).

La matriz de confusión que se obtuvo tras la aplicación del modelo predictivo basado en regresión logística se muestra en la siguiente tabla.

		Realidad	
		Recurrencia	No recurrencia
Predicción	Recurrencia (+)	17	19
	No recurrencia (-)	27	322

Tabla 26: Matriz de confusión del modelo de regresión logística para predecir la recurrencia precoz postrasplante del CHC en el grupo de entrenamiento

Así, el modelo predictivo basado en regresión logística presenta los siguientes parámetros de rendimiento para el grupo de entrenamiento:

- Sensibilidad: 38,64 %
- Especificidad: 94,43 %
- Acierto: 88,10 %
- Valor Predictivo Positivo: 47,22 %
- Valor Predictivo Negativo: 92,26 %
- Razón de Verosimilitud para Test Positivo: 6,93
- Razón de Verosimilitud para Test Negativo: 0,144

5.2.2.2. – Modelo de redes bayesianas

Después de discretizar las variables y procesar los datos según el método descrito previamente, se generó un modelo predictivo de RB tipo TAN (*Tree Augmented Network*) con 5 nodos que representan variables relacionadas con la variable objetivo y que proporcionan la probabilidad de recurrencia del CHC en los dos primeros años después del trasplante hepático. Dichas variables son:

- Respuesta a terapia locorregional pretrasplante
- Estatus después del tratamiento pretrasplante
- *Downstaging* a Criterios de Milán (sí/no)
- Logaritmo en base 10 del valor de la AFP previa al TH, estableciendo un punto de corte en 9 ng/ml.
- Valor máximo de la AFP durante el seguimiento pretrasplante, dividida en 3 intervalos:
 - $AFP < 400$ ng/ml
 - $AFP \geq 400$ ng/ml y < 1000 ng/ml
 - $AFP \geq 1000$ ng/ml

A continuación, se muestra el grafo en el que se basa el modelo predictivo desarrollado:

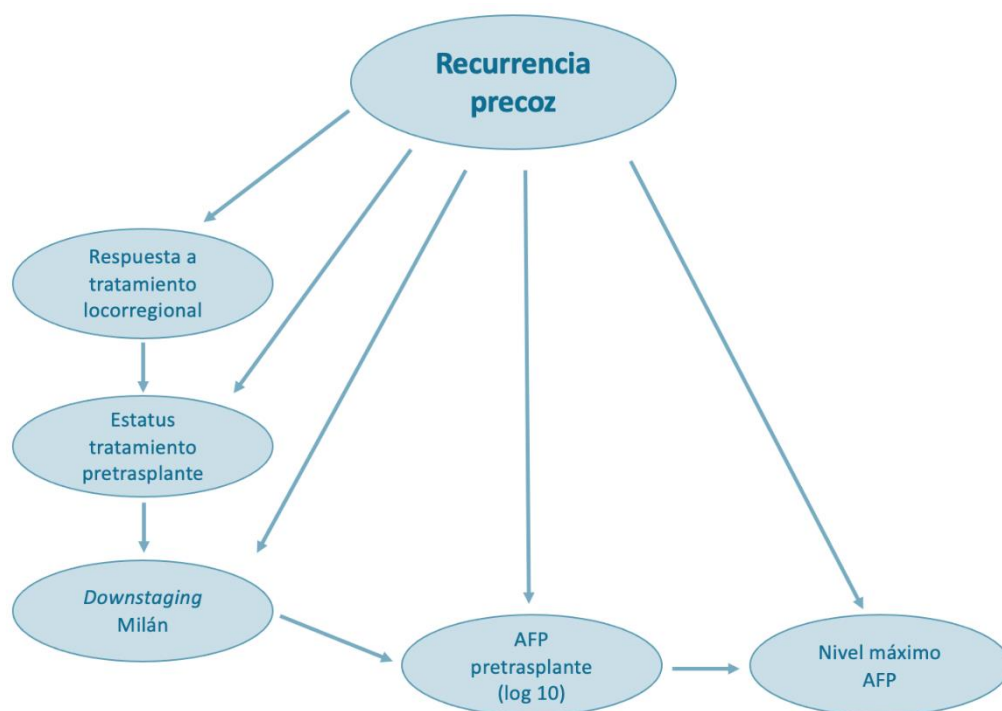


Figura 20: Grafo del modelo de redes bayesianas para predicción de recurrencia precoz postrasplante del CHC

Con los valores de utilidad previamente definidos, se obtuvo una curva ROC con un umbral de 0,229 para discriminar entre recurrencia precoz y no recurrencia precoz.

El ABC de la curva ROC para el modelo de redes bayesianas en el grupo de entrenamiento es de 0,817 (IC 95 %: 0,717 – 0,888).

A continuación, se representa dicha curva.

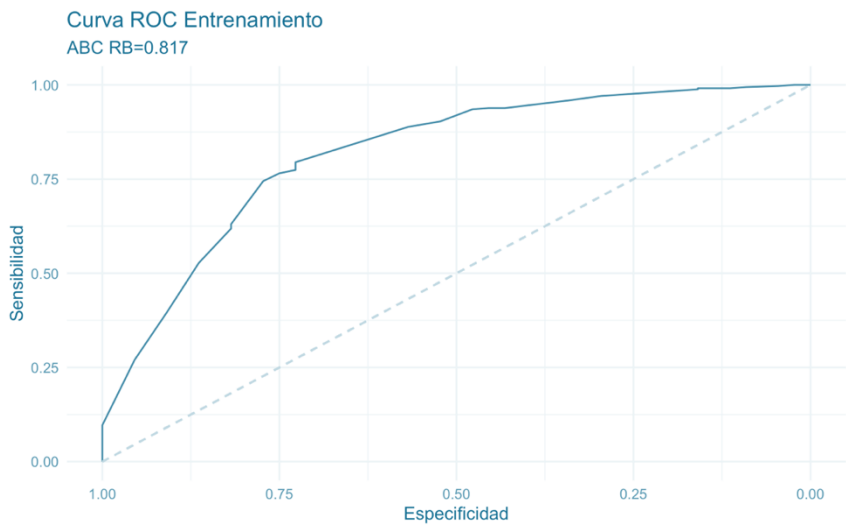


Figura 21: Curva ROC obtenida con el modelo de redes bayesianas en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

La matriz de confusión que se obtuvo tras la aplicación del modelo predictivo basado en redes bayesianas se muestra en la siguiente tabla.

		Realidad	
		Recurrencia	No recurrencia
Predicción	Recurrencia (+)	23	33
	No recurrencia (-)	21	308

Tabla 27: Matriz de confusión del modelo de redes bayesianas para predecir la recurrencia precoz postrasplante del CHC en el grupo de entrenamiento.

Así, el modelo predictivo basado en redes bayesianas presenta los siguientes parámetros de rendimiento para el grupo de entrenamiento:

- Sensibilidad: 52,27 %
- Especificidad: 90,32 %
- Acierto: 86,00 %
- Valor Predictivo Positivo: 41,07 %
- Valor Predictivo Negativo: 93,62 %
- Razón de Verosimilitud para Test Positivo: 5,39
- Razón de Verosimilitud para Test Negativo: 0,53

5.3.- Generalización de los modelos desarrollados

Como se describió en el apartado 4.2, para la generalización de los modelos predictivos se ha utilizado una muestra de 130 receptores de trasplante hepático con indicación de hepatocarcinoma en el mismo centro, desde noviembre de 2015 hasta diciembre de 2018, que cumplieran los criterios de inclusión y ninguno de exclusión y con un seguimiento mínimo de 2 años después del TH. El hecho de que los pacientes del grupo de generalización sean inmediatamente posteriores a los del grupo de entrenamiento permite validar internamente el modelo.

5.3.1.- Modelo de regresión logística

Tras aplicar el modelo de regresión logística al grupo de generalización, se obtuvo la siguiente curva ROC, con una ABC de 0,843 (IC 95 %: 0,727 – 0,958).

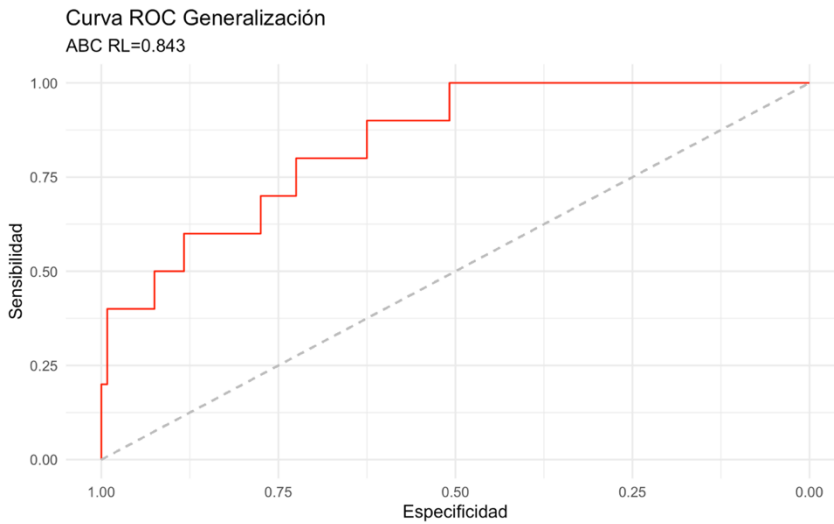


Figura 22: Curva ROC obtenida con el modelo de regresión logística en el grupo de generalización para determinar la recurrencia precoz postrasplante del CHC

Se aplicó el correspondiente umbral de 0,281, de tal manera que se obtuvo la siguiente matriz de confusión del modelo de regresión logística para predecir la recurrencia precoz postrasplante del CHC en el grupo de generalización.

		Realidad	
		Recurrencia	No recurrencia
Predicción	Recurrencia (+)	4	6
	No recurrencia (-)	6	114

Tabla 28: Matriz de confusión del modelo de regresión para predecir la recurrencia precoz postrasplante del CHC en el grupo de generalización

Así, el modelo predictivo basado en regresión logística presenta los siguientes parámetros de rendimiento para el grupo de generalización:

- Sensibilidad: 40,00 %
- Especificidad: 95,00 %
- Acierto: 90,80 %
- Valor Predictivo Positivo: 40,00 %
- Valor Predictivo Negativo: 95,00 %
- Razón de Verosimilitud para Test Positivo: 8,00
- Razón de Verosimilitud para Test Negativo: 0,63

5.3.2.- Modelo de redes bayesianas

Finalmente se aplicó el modelo predictivo de redes bayesianas al grupo de generalización, obteniéndose una curva ROC con una ABC de 0,889 (IC 95 %: 0,792 – 0,985):

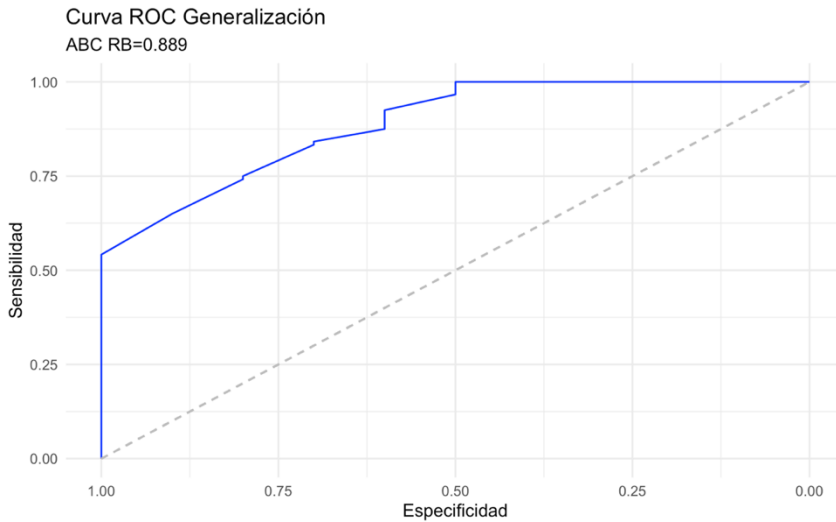


Figura 23: Curva ROC obtenida con el modelo de redes bayesianas en el grupo de generalización para determinar la recurrencia precoz postrasplante del CHC

Después de aplicar el correspondiente umbral de 0,229, se obtuvo la matriz de confusión para predecir la recurrencia precoz postrasplante del CHC en el grupo de generalización.

		Realidad	
		Recurrencia	No recurrencia
Predicción	Recurrencia (+)	5	4
	No recurrencia (-)	5	116

Tabla 29: Matriz de confusión del modelo de redes bayesianas para predecir la recurrencia precoz postrasplante del CHC en el grupo de generalización

El modelo predictivo de redes bayesianas presenta los siguientes parámetros de rendimiento para el grupo de generalización:

- Sensibilidad: 50,00 %
- Especificidad: 96,67 %
- Acierto: 93,1 %
- Valor Predictivo Positivo: 55,56 %
- Valor Predictivo Negativo: 95,87 %
- Razón de Verosimilitud para Test Positivo: 15,15
- Razón de Verosimilitud para Test Negativo: 0,517

5.4.- Comparación de los resultados de los modelos predictivos desarrollados

5.4.1.- Grupo de entrenamiento

Al comparar la capacidad de predicción de la recurrencia precoz postrasplante del CHC de los modelos de regresión logística y de redes bayesianas para el grupo de entrenamiento, se constató que el ABC del modelo de redes bayesianas es superior (0,817) al obtenido por el modelo de regresión logística (0,791). Sin embargo, estas diferencias no alcanzaron significación estadística ($p = 0,2$). A pesar de ello, la curva del modelo de RB se encuentra desplazada más hacia el límite superior izquierdo de la gráfica comparativamente con la curva del modelo de RL.

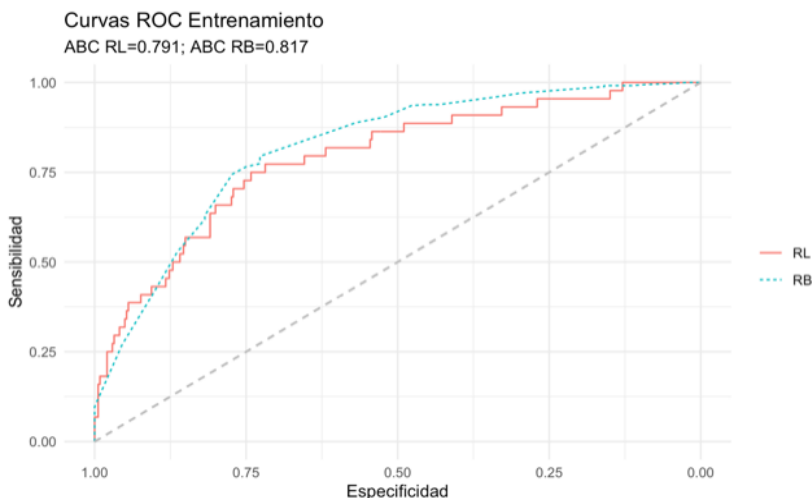


Figura 24: Comparación de las curvas ROC obtenida con los modelos de regresión logística y redes bayesianas en el grupo de entrenamiento para determinar la recurrencia precoz postrasplante del CHC

Asimismo, se analizó el grado de concordancia de los resultados obtenidos por ambos modelos. El coeficiente Kappa fue de 0,485 (IC 95 %: 0,338 – 0,632), por lo que el grado de concordancia entre los modelos puede considerarse como moderado.

5.4.2.- Grupo de generalización

Del mismo modo, también se comparó la capacidad de predicción de la recurrencia precoz postrasplante del CHC de los modelos de regresión logística y de redes bayesianas para el grupo de generalización. El resultado fue una ausencia de significación estadística ($p = 0,4$). Sin embargo, el ABC del modelo de redes bayesianas (0,889) es superior al obtenido por el modelo de regresión logística (0,843) y, de nuevo, la curva del modelo de RB se encuentra claramente desplazada más hacia el límite superior izquierdo de la gráfica comparativamente con la curva del modelo de RL.

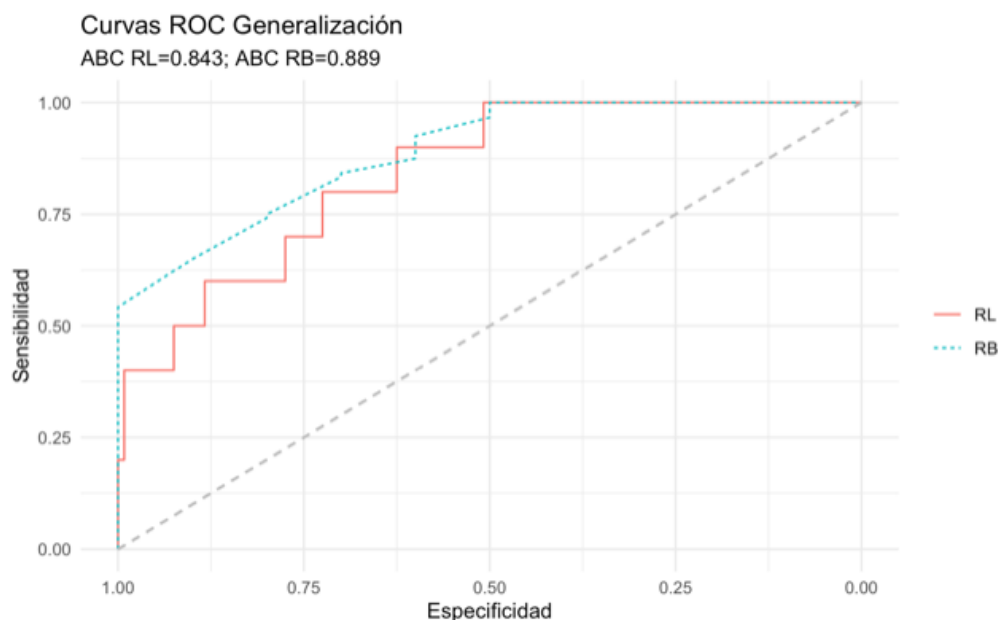


Figura 25: Comparación de las curvas ROC obtenida con los modelos de regresión logística y redes bayesianas en el grupo de generalización para determinar la recurrencia precoz postrasplante del CHC

El coeficiente Kappa para el grado de concordancia fue de 0,613 (IC 95 %: 0,316 – 0,803), por lo que el grado de concordancia entre los modelos para el conjunto de generalización puede considerarse como bueno.

Discusión

6.- Discusión

El CHC es la quinta neoplasia maligna más frecuente y la segunda causa de muerte relacionada con cáncer del mundo (133), representando, además, el 90 % de todos los tumores malignos primarios del hígado. Con las actuales proyecciones, la Organización Mundial de la Salud estima que para el año 2030, aproximadamente un millón de personas fallecerá anualmente a causa del CHC (134). Dado que la gran mayoría de CHC se presentan en pacientes con cirrosis, la hepatopatía de base limita las opciones terapéuticas de las que disponen los pacientes. En este sentido, es fundamental seleccionar el mejor tratamiento para conseguir incrementar la supervivencia con una adecuada calidad de vida.

Entre las opciones curativas se encuentran la ablación, la resección quirúrgica y el trasplante hepático.

La resección quirúrgica sólo puede realizarse en aquellos pacientes sin hipertensión portal, con función hepática adecuada y un volumen de remanente hepático aceptable.

Con una correcta selección de pacientes (tumores de hasta 2 cm de diámetro), la supervivencia global a 5 años supera el 80 %, resultados equiparables a los conseguidos con las técnicas ablativas, que suelen asociar una menor morbilidad.

Sin embargo, tanto la resección como la ablación están gravadas con un elevado porcentaje de recidiva, que alcanza el 70 % a los 5 años.

En este sentido, el TH es la mejor opción terapéutica en términos oncológicos, ya que permite eliminar completamente la enfermedad neoplásica hepática y sustituir el hígado patológico subyacente. Inicialmente, dada la ausencia de criterios de selección, el TH con intención oncológica estaba penalizado con unos malos resultados en términos de SG y SLE. Sin embargo, la descripción publicada en el año 1996 (3) de los criterios de Milán supuso toda una revolución en el campo de la *transplant oncology*, con unas cifras de SG a 5 años de entre 60-80 % y una recurrencia del 15 %. Tanto es así que dichos criterios se han convertido en la piedra angular de la mayoría de estrategias de inclusión en lista de espera en todo el mundo y, desde su publicación y validación, han marcado una era en la que la base de las decisiones clínicas es la evaluación de los candidatos a trasplante oncológico mediante criterios morfológicos preestablecidos.

A pesar de que los criterios de Milán siguen siendo los más utilizados para evaluar la idoneidad de un candidato a TH, su utilidad ha sido objeto de debate, especialmente en los últimos años.

Esto se debe a que los buenos resultados del TH mediante criterios predefinidos han animado a muchos grupos de trasplante a establecer criterios menos restrictivos. Sólo 5 años más tarde aparecieron los primeros criterios expandidos (135), los llamados criterios UCSF, que demostraron en un estudio retrospectivo que una ampliación de los criterios de Milán era posible, con resultados de SG y SLE equivalentes. Con estos resultados, se comenzó a esbozar la duda sobre si quizá los criterios de Milán eran excesivamente restrictivos.

Progresivamente han ido apareciendo multitud de criterios morfológicos expandidos, dentro de los cuales, los llamados criterios *Up-to-seven*, son los que cuentan con mayor evidencia científica y validación externa.

Además, en estos 25 años se han producido numerosos cambios, tanto en la etiología de la enfermedad que motiva el trasplante, como en la propia práctica clínica.

Es importante destacar la progresiva implantación del tratamiento neoadyuvante (sobre todo de las terapias locorregionales, con intención de *downstaging* y como puente al TH) en el manejo de los pacientes con CHC. La constatación de que después de estos tratamientos algunos tumores responden o se reducen mientras que otros tumores similares persisten e incluso progresan, y las implicaciones que esta respuesta tiene sobre la SG y SLE postrasplante, ha llevado a la conclusión de que la morfología tumoral es insuficiente para comprender y predecir el comportamiento biológico del CHC (136,137). De este modo, la respuesta al tratamiento

neoadyuvante (debidamente objetivada mediante los criterios mRECIST) ha pasado a considerarse un factor pronóstico importante y se ha convertido en sí misma en una herramienta para la selección de pacientes para el TH.

Asimismo, se ha producido una creciente comprensión de la implicación de diferentes marcadores biológicos en el comportamiento del CHC. La AFP ha demostrado tener un papel no sólo diagnóstico y de seguimiento, sino también predictivo de los resultados del TH. De ahí que en España, desde 2018 la SETH recomiende contraindicar formalmente el TH en pacientes con cifras de AFP > 1000 ng/ml (17). Otros marcadores, como la DCP o el RNL (que traduce la respuesta inflamatoria) también se han relacionado con la SG y la recurrencia del CHC después del TH.

Por tanto, en el campo de la *transplant oncology* ha quedado demostrado que la morfología tumoral (y, por tanto, los criterios de selección basados exclusivamente en ella) es insuficiente para comprender y predecir el comportamiento biológico del CHC. Este hecho ha conducido al desarrollo de múltiples *scores* y criterios pronósticos combinados que tratan de aunar toda la información disponible con el fin de predecir con mayor exactitud el riesgo de recurrencia y la supervivencia de los receptores de TH con intención oncológica.

Sin embargo, no existe unanimidad acerca de la superioridad de cualquiera estas herramientas pronósticas sobre el resto.

No obstante, y a pesar de todos estos avances, la recurrencia tumoral postrasplante del CHC sigue siendo una causa importante de pérdida del injerto y muerte del receptor, aun con los buenos resultados que ofrecen los criterios de selección (tanto morfológicos como combinados) actualmente disponibles.

Además, existe un determinado subgrupo de pacientes en los que la recurrencia se produce precozmente (en nuestra serie, el 10,5 % de los pacientes en los dos primeros años después del TH). Esta recurrencia precoz supone una situación extremadamente ineficiente desde el punto de vista de la utilización de un recurso escaso como son los órganos disponibles, aun con el incremento relativo que se ha producido en los últimos años en el número de injertos gracias al uso de órganos procedentes de donantes en asistolia controlada.

En este punto concreto se centra el objetivo de nuestro estudio, al considerar que la identificación antes del trasplante de aquellos pacientes con un menor riesgo de recurrencia precoz del CHC, puede mejorar los resultados de SG y SLE, así como el coste-utilidad del TH a medio y largo plazo. De esta manera podemos seleccionar a los candidatos más idóneos para el TH, llevando a cabo una utilización más selectiva de los injertos disponibles, que optimice los resultados.

En este sentido los sistemas de inteligencia artificial y *machine learning*, como las redes bayesianas, resultan útiles para el análisis y

la comprensión de situaciones clínicas complejas, ayudando a predecir el comportamiento tumoral del CHC, en el que influyen múltiples variables que pueden interaccionar de manera distinta entre sí e influir en la recurrencia tumoral precoz postrasplante (variable objetivo del estudio)-

La opción de implantar este tipo de sistemas mediante herramientas informáticas de acceso abierto para la estimación del riesgo de recurrencia precoz abre la posibilidad de participar a todos aquellos investigadores o grupos de trasplante que quieran comparar y validar sus resultados, tanto interna como externamente. El elevado número de pacientes y la experiencia clínica adquirida a lo largo de las últimas décadas en muchos centros españoles pueden ser la base para la realización de estudios consistentes que permitan mejorar estos sistemas predictivos de recurrencia del CHC. En su aplicación clínica, la propia naturaleza de los sistemas de *machine learning* y RB los hace enormemente plásticos, y los modelos generados pueden revisarse mediante reevaluaciones periódicas de su rendimiento, admitiendo la inclusión de nuevas variables, que mejoren la capacidad predictiva de los modelos en el futuro.

6.1.- Estructura del trabajo

Se ha utilizado una serie consecutiva de pacientes trasplantados por CHC, con un seguimiento mínimo de 2 años después del TH.

Se elaboró una base de datos *ad hoc* que incluye únicamente variables disponibles previamente al trasplante. Con ella se han desarrollado modelos predictivos de recurrencia del CHC en los dos primeros años postrasplante mediante dos metodologías distintas: la regresión logística y las redes bayesianas.

La población del estudio incluye aquellos pacientes mayores de 18 años y afectos de CHC, receptores de un TH en el Hospital Universitari i Politècnic La Fe de Valencia desde enero de 2004 hasta diciembre de 2018. En este sentido, se han excluido aquellos pacientes con hepatocarcinomas incidentales, dado que, en este caso, la información relativa a algunas variables pretrasplante no está disponible. Se descartaron también aquellos pacientes con una pérdida del seguimiento en los 2 primeros años después del trasplante hepático, por ser esta situación incompatible con la variable objetivo del estudio.

Los datos empleados para el diseño y elaboración de los modelos y su posterior validación proceden del estudio preoperatorio de los pacientes, que se ha realizado en todos los casos por parte del mismo grupo multidisciplinar, intervenidos por el mismo equipo quirúrgico con una amplia experiencia y sometidos posteriormente a un seguimiento estrecho con recogida de datos según protocolos establecidos en base a vías clínicas. Esta procedencia única de todos los datos, con un

protocolo común de diagnóstico, tratamiento pretrasplante y seguimiento después del mismo, confiere homogeneidad a la información recogida y utilizada para la elaboración de la base de datos. Así pues, a pesar de que la recogida de datos es retrospectiva, el número de datos perdidos es muy bajo (inferior al 1 %), ya que la información procede de una base muy compleja con recogida de datos prospectiva.

La variable objetivo principal fue la recurrencia del CHC en los dos primeros años después del trasplante. Se ha utilizado la recidiva y la SLE en lugar de la SG dado que la intención es que el modelo predictivo se centre en el análisis del comportamiento oncológico del CHC, y en una posible aplicación práctica en el campo de la *trasplant oncology*. En este sentido, en la supervivencia global a largo plazo pueden influir otros factores como la patología de base del paciente, cambios en la inmunosupresión, rechazo, tumores *de novo*, etc.

Por otra parte, se acotó la recurrencia a los dos primeros años después del TH por entender que se trata de la situación menos eficiente en términos de coste-utilidad para el paciente y el procedimiento. Es evidente que, cuando consideramos la supervivencia libre de enfermedad de los receptores de TH para CHC, hay un subgrupo de pacientes que presentan una recidiva precoz (en la presente serie, 10,5 %). En estos casos, el beneficio y la utilidad del TH es nulo o, como poco, dudoso. Además, reducir el número de estas recurrencias precoces identificando previamente al trasplante a los pacientes con mayor riesgo permitiría cumplir un doble objetivo: por un

lado, optimizar el uso de injertos disponibles; por otro, ofrecer a aquellos pacientes de alto riesgo alternativas terapéuticas menos agresivas y más beneficiosas para ellos (como los tratamientos locorregionales o terapias sistémicas) que pueden prolongar su supervivencia (138,139) con una adecuada calidad de vida. Hasta la fecha, ningún modelo o *score* pronóstico (tanto morfológico como combinado) disponible en la literatura se ha centrado en analizar este punto en concreto.

Se realizó una revisión bibliográfica actualizada con el fin de identificar los posibles factores que pueden influir en el riesgo de recurrencia del CHC postrasplante. Tras ello, se seleccionaron un total de 32 variables del receptor y del tumor injertado, de fácil obtención en la práctica clínica habitual y todas ellas disponibles antes del trasplante hepático. Con dichas variables se ha desarrollado un modelo predictivo de redes bayesianas práctico y de fácil aplicabilidad, que permite estimar el riesgo de recurrencia del CHC en los dos primeros años después del TH.

La gran mayoría de las variables recogidas corresponden a datos objetivos procedentes de las historias clínicas de los pacientes incluidos en el estudio: resultados analíticos, pruebas de imagen y variables epidemiológicas. Posiblemente la variable que presenta más problemas por su objetividad e implicación en el modelo definitivo es la respuesta a los tratamientos locorregionales pretrasplante. En este sentido la evaluación y los informes radiológicos de control después de este tipo de terapias en el Hospital Universitario La Fe se han hecho

siempre siguiendo los criterios y las recomendaciones de las guías clínicas vigentes en cada momento. Dado que se trata de una serie amplia que abarca 15 años, estos criterios y por tanto los informes radiológicos disponibles, se han ido modificando y actualizando. Sin embargo, es importante destacar que tanto las terapias locorregionales como la evaluación de las pruebas radiológicas posteriores se han realizado siempre por parte de un grupo concreto y reducido de radiólogos específicamente dedicados al intervencionismo y radiología hepatobiliopancreática, en el seno de comités multidisciplinares, por lo que el grado de subjetividad que cabría esperar se reduce notablemente.

En el mismo sentido, aunque en mucha menor medida, la puntuación Child, incluida entre las variables recogidas en el estudio, contiene datos que también dependen del evaluador, por lo que podría considerarse una variable que incluye cierto grado de subjetividad. No obstante, ésta es limitada ya que, de nuevo, la valoración pretrasplante de los pacientes en nuestro centro se lleva a cabo por un grupo reducido y específicamente dedicado de hepatólogos.

Es importante destacar que determinadas variables relacionadas con el comportamiento del tumor han demostrado tener un impacto importante en el modelo generado y además se encuentran relacionadas entre sí. Es el caso de la respuesta al tratamiento pretrasplante (categorizada según los criterios mRECIST), o el “estado de tratamiento pretrasplante”. No obstante, la base para categorizar este tipo de variables son criterios morfológicos (Milán y *Up-to-seven*)

de uso común para los grupos de trasplante. Asimismo, es evidente que estas variables se encuentran relacionadas entre sí y con otras (como la presencia o no de *downstaging*). Es aquí donde las RB son de gran utilidad para detectar relaciones complejas, múltiples y muchas veces no lineales entre las variables del estudio, mejorando la capacidad predictiva y el rendimiento del modelo generado.

La práctica totalidad de las variables son de fácil obtención en los protocolos de estudio, tratamiento y seguimiento pretrasplante de pacientes con CHC en lista de espera. Además, todas ellas son conocidas con anterioridad al TH, por lo que pueden documentarse en una base de datos, lo que facilita la aplicación del modelo para poder determinar de manera ágil y práctica el riesgo de recurrencia precoz para cada paciente antes de tomar la decisión clínica de proceder o no con el TH o, al menos, de incluir en lista de espera al individuo en cuestión.

Una vez determinadas las variables que debían formar parte del estudio se procedió a completar la base de datos con los datos de la base de trasplante de la Unidad, añadiendo de forma retrospectiva la información específica de las variables del estudio. Se obtuvo una muestra de 515 pacientes receptores de TH en un periodo de 15 años. Para la elaboración de los modelos, esta muestra fue dividida en dos grupos (entrenamiento y generalización) en una proporción 75 / 25 %, disponiendo de 385 pacientes para el grupo de entrenamiento y de 130 pacientes para el grupo de generalización. Asimismo, esta división también sigue un criterio temporal, puesto que los 385 pacientes del

grupo de entrenamiento corresponden a la parte inicial de la serie (enero 2004 a noviembre 2015) mientras que los 130 pacientes del grupo de generalización son pacientes consecutivos e inmediatamente posteriores (noviembre 2015 a diciembre 2018). De esta manera, con el grupo de generalización se cumple también el objetivo de realizar una validación interna mediante un grupo de receptores trasplantados con posterioridad a aquellos con los que se generó el modelo predictivo.

Sin embargo, la división temporal en grupos de entrenamiento y generalización ha podido provocar la aparición de sesgos entre ambos grupos, que pudieran tener influencia en los resultados definitivos del estudio. Existen diferencias estadísticamente significativas entre los grupos, y éstas se dan fundamentalmente en aquellas variables relacionadas con el tratamiento frente al VHC y la respuesta al mismo y la proporción de pacientes tratados mediante radioembolización. En concreto, estas variables no forman parte del modelo definitivo y estas diferencias son atribuibles a la propia división temporal de los datos.

También se han objetivado diferencias en cuanto a la puntuación MELD y a los niveles de AFP pretrasplante, con una tendencia hacia valores ligeramente inferiores en el grupo de generalización. Es posible que, con una distribución aleatoria de los pacientes en los dos grupos de estudio, se pudieran evitar los posibles sesgos en los que se ha podido incurrir al realizar una división con criterio temporal.

El número de trasplantes realizados por CHC en el Hospital Universitari i Politècnic La Fe ha mantenido un ritmo constante, en torno a los 30 procedimientos anuales. Sin embargo, en los últimos años estas cifras han aumentado hasta situarse por encima de los 40 TH al año. Este incremento puede tener distintas causas: los avances en el diagnóstico del CHC por el uso generalizado de la RM, la derivación de pacientes con CHC desde resección quirúrgica a TH por la reducción de la lista de espera y los mejores resultados oncológicos del TH, o la mejoría en la respuesta a los tratamientos locorregionales pretrasplante, que permiten tanto el *downstaging* como el control de la enfermedad mediante terapias puente al TH, evitando la progresión de enfermedad y reduciendo la salida (*drop-out*) de la lista de espera. Asimismo, estas cifras suponen aproximadamente un tercio de todos los trasplantes realizados en el centro, cifras similares a las del resto de centros de trasplante de España (140).

Dado que se trata de una serie amplia y recopilada a lo largo de 15 años y de forma retrospectiva, las distintas variaciones técnicas y terapéuticas en el manejo de los pacientes han podido tener un impacto en los resultados. De esta manera las terapias locorregionales se han instalado como un pilar fundamental en el manejo de los pacientes con CHC potencialmente trasplantables. En este sentido, por ejemplo, la técnica de QETA empleada se ha modificado, cambiando la modalidad convencional por la que emplea DC-Beads. Asimismo, también se ha incorporado la RETA como tratamiento puente y de *downstaging*, favoreciendo el rescate de pacientes y la ampliación del *pool* de receptores que podían beneficiarse de un TH.

Esto ha comportado cambios evolutivos en las características de los receptores a lo largo del tiempo, incrementando la heterogeneidad de la serie (tumores más grandes, mayor número de tratamientos previos al trasplante, etc). Esta variabilidad, y especialmente el hecho de que el grupo de generalización esté formado por los pacientes que corresponden al final de la serie temporal de estudio, otorga una mayor fiabilidad al modelo, corroborando su utilidad y su capacidad predictiva, que puede aplicarse a distintos tipos de receptores y situaciones clínicas cambiantes a pesar de las posibles diferencias.

La supervivencia global de los 515 pacientes incluidos en el estudio fue del 85,8 % al año del trasplante, del 76 % tres años después del TH y del 66,9 % a los 5 años, datos similares a los existentes en la literatura para pacientes trasplantados con criterios de Milán (3) y con criterios expandidos morfológicos (7,141) o combinados (75,76). Asimismo, se trata de resultados similares, aunque ligeramente inferiores, a los del Registro Español de Trasplante Hepático, que muestra supervivencias globales del 89,5 % a 1 año y del 72,3 % a 5 años para la serie de pacientes trasplantados entre 1991 y 2021 (140).

La supervivencia global a 1 y 3 años del grupo de generalización es ligeramente superior al grupo de entrenamiento, probablemente debido a los avances en el manejo perioperatorio de los receptores, las mejoras en los tratamientos pretrasplante del CHC y las terapias antivirales, si bien es cierto que no existen diferencias estadísticamente significativas entre la supervivencia global de ambos

grupos. En este sentido, la implantación del uso de injertos procedentes de donación en asistolia controlada no ha supuesto una merma en los resultados de supervivencia.

En cuanto a la supervivencia libre de enfermedad del conjunto de pacientes de la serie, ésta fue del 83,3 % al año del trasplante, del 71,2 % tres años después del TH y del 63,2 % a los 5 años, en concordancia con los datos publicados en la literatura, sin diferencias significativas entre los dos grupos de estudio. Del mismo modo, los datos del grupo de generalización son ligeramente mejores que los del grupo de entrenamiento: 82,1 % vs. 86,9 % al año del TH, 69,9 % vs. 75 % a los tres años y 62,3 % vs. 65,4 % cinco años después del TH.

En cuanto a las características de los receptores, a pesar de que se trata de un estudio unicéntrico, no existen importantes diferencias comparativamente con los datos procedentes de otras series, tanto a nivel nacional como internacional (140).

La mediana de edad del global de la serie fue de 59 años, siendo de 58 para el grupo de entrenamiento y de 60 para el grupo de generalización, sin diferencias estadísticamente significativas. Asimismo, se observó un claro predominio de pacientes varones en el estudio (84,5 %) y, aunque en el grupo de generalización hay un 5 % menos de varones, éstos siguen predominando en una proporción de 5 a 1.

En ambos grupos existe una clara tendencia al sobrepeso, con una mediana de IMC de 27,44 y 28,22 kg/m², respectivamente. En este

sentido, más del 20 % de los pacientes presentaban un síndrome metabólico en el momento del trasplante hepático (21,3 % vs. 24,5 %), datos que reflejan una presencia de comorbilidades relevante en los receptores con CHC.

En cuanto a la hepatopatía de base, la causa más frecuente fue la infección crónica por VHC (61,4 %), seguida de la ingesta de alcohol (23,1 %), sin diferencias significativas entre los dos grupos. No obstante, se observa una tendencia a una menor presencia de la cirrosis por VHC en el grupo de generalización (57,7 %), con un incremento de hasta casi el doble de otras hepatopatías (6,2 % vs. 11,6 %) debido fundamentalmente al aumento de la esteatohepatitis no alcohólica como indicación para trasplante. Asimismo, también se evidencia una disminución en la gravedad de la hepatopatía de base, con una mayor proporción de pacientes Child A en el grupo de generalización (47,5 % vs. 63,8 %) y un reducido porcentaje de pacientes Child C (18,7 % vs. 9,2 %). Estas diferencias sí alcanzaron significación estadística ($p = 0,004$). Del mismo modo, se objetiva también una tendencia a la disminución de la puntuación MELD funcional con una mediana inferior para el grupo de generalización (13 vs. 10). Esta disminución de pacientes Child C y del MELD se debe fundamentalmente a la disminución del tiempo en lista de espera producido por el aumento del número de donantes. En este sentido, el tiempo en lista de espera fue significativamente menor para el grupo de generalización; 147 vs. 63 días. Asimismo, hay que tener en cuenta que aquellos pacientes que se tratan durante su permanencia en lista de espera (como puente al TH o con intención de *downstaging*) son los

que tienen mejores puntuaciones de Child, mientras que aquellos con puntuaciones más elevadas tienen tendencia a salir de lista de espera por progresión dado que en muchas ocasiones no pueden tratarse.

En este sentido es importante destacar las diferencias detectadas en cuanto al manejo de la infección por VHC entre los grupos de estudio, ya que la proporción de pacientes con VHC no tratados en el grupo de generalización es significativamente menor que en el grupo de entrenamiento (23,4 % vs. 5,4 %, $p = 0,0005$). Además, en el grupo de generalización se observa una mayor proporción de pacientes con respuesta viral sostenida (8,8 % vs. 43,9 %) y una drástica disminución de los pacientes sin respuesta (18,4 % vs. 3,9 %), con diferencias también significativas ($p = 0,0004$), todo ello debido a la implantación del tratamiento con AAD a partir del año 2015.

En cuanto a la carga tumoral, 3 de cada 4 pacientes presentaban afectación de un lóbulo (77,3 %), con presencia de una única lesión en aproximadamente la mitad de los casos, tanto en el grupo de entrenamiento como en el de generalización (54,3 % vs. 48,5 %). Asimismo, la proporción de pacientes con 3 lesiones o menos es superior al 95 % en ambos grupos. En este sentido, una amplia mayoría de los receptores llegaron al TH dentro de los criterios de Milán (94,6 %), cifra que asciende al 98,8 % con los criterios *Up-to-seven*. Esto implica que sólo un grupo residual de 5 pacientes (1 % de la serie) se trasplantaron fuera de los criterios morfológicos *Up-to-seven* (todos ellos en el grupo de entrenamiento). Esto apunta a una excelente selección de los candidatos a trasplante desde el punto de

vista de los criterios estrictamente morfológicos. En ninguna de estas variables se objetivaron diferencias significativas entre el grupo de entrenamiento y el de generalización, ni tampoco en el diámetro del nódulo de mayor tamaño ni en la suma de los tamaños de los nódulos tumorales. En este punto es necesario recordar que, en el presente estudio se han utilizado datos morfológicos procedentes exclusivamente de pruebas de imagen pretrasplante. No se ha realizado una correlación entre la estadificación radiológica y la anatomopatológica, puesto que la intención es elaborar un modelo predictivo basado exclusivamente en variables disponibles previas al trasplante.

En cuanto a los marcadores biológicos, los niveles de AFP se encontraban en una mediana de 8,9 ng/ml (4,2 ng/ml – 14,8 ng/ml) previamente al trasplante, lo que indica un excelente control tumoral desde el punto de vista biológico. Menos del 10 % de los pacientes ($n = 42$) superaron en algún momento el nivel de 400 ng/ml durante el seguimiento pretrasplante, cifra que se reduce a menos del 5 % ($n = 25$) si el umbral lo situamos en 1000 ng/ml. En este sentido, el grupo de generalización mostró unos niveles de AFP pretrasplante significativamente inferiores (10,2 vs. 5,6, $p < 0,001$), si bien dentro de unos márgenes estrechos.

391 pacientes de los 515 recibieron algún tipo de tratamiento antes del TH, lo que supone un 75,9 % de la serie: 76,1 % en el grupo de entrenamiento vs. 75,4 % en el grupo de generalización, sin diferencias entre grupos ($p = 0,75$). De ellos, 229 (44,5 % de la serie) recibió una

única sesión de tratamiento. En una amplia mayoría de los receptores (61,4 %) este tratamiento sirvió como puente al trasplante hepático, mientras que se consiguió un *downstaging* efectivo a criterios de Milán o *Up-to-seven* en un 13,6 % de pacientes (12,9 % vs. 15,4 %), de tal manera que un 24,1 % de los pacientes de la serie recibieron un TH sin tratamiento locorregional previo y un escaso 0,97 % de los mismos llegó al mismo con un *downstaging* inefectivo según criterios morfológicos (n = 5). Todos ellos pertenecen al grupo de entrenamiento (aunque no hay diferencias significativas entre grupos). Estos datos, combinados con lo referido en el párrafo anterior, reflejan el estricto control periódico (analítico y radiológico) al que se someten los pacientes en lista de espera para TH a través del comité multidisciplinar. Así pues, es cierto que determinados pacientes en los que la respuesta al tratamiento es dudosa por pruebas de imagen pueden haber sido receptores de un TH excediendo criterios exclusivamente morfológicos. La presencia de todos ellos en el grupo de entrenamiento probablemente se deba a la mejora en las pruebas de imagen y su capacidad para detectar de nódulos tumorales cada vez más pequeños y caracterizar su comportamiento (lesiones secundarias al tratamiento vs. persistencia tumoral) de los tratamientos locorregionales, sesgo atribuible de nuevo a la división temporal de los grupos de estudio.

Asimismo, entre los pacientes tratados, más del 90 % de los mismos requirieron ≤ 3 sesiones de tratamiento antes del trasplante, sin diferencias entre grupos. La respuesta al tratamiento fue mayoritariamente buena, con predominio de la respuesta completa,

que se produjo en 193 de los 391 pacientes tratados previamente al trasplante (49,4 %). Aunque no se detectaron diferencias significativas entre grupos, sí se observó claramente una mayor proporción de respuestas completas en el grupo de generalización (34,8 % vs. 45,4 %), probablemente debido a las mejoras en las técnicas de tratamiento locorregional, a la introducción de la modalidad de QETA con DC-Beads y a la implantación de la radioembolización, terapias cada vez más selectivas sobre los nódulos tumorales y, además, más efectivas. La progresión tumoral a pesar del tratamiento sólo se produjo en un 6,6 % de los pacientes del estudio (6,7 % grupo de entrenamiento vs. 6,2 % grupo de generalización, $p = 0,093$), aunque como se ha comentado anteriormente, la gran mayoría de ellos se encontraban dentro de criterios morfológicos antes del trasplante.

De entre las modalidades de tratamiento locorregional, la más empleada fue la QETA, ya que más de la mitad (57,7 %) de los pacientes del estudio ($n = 297$) recibieron al menos una sesión, seguida de la radiofrecuencia ($n = 137$, 26,6 %). En cuanto a las diferencias entre grupos, destaca el aumento en la utilización de la RETA, dada su utilidad para el rescate de pacientes con tumores de gran tamaño: $n = 2$ (0,5 %) en el grupo de entrenamiento vs. $n = 5$ (3,8 %) en el grupo de generalización.

Los datos de recurrencia global de la serie ($n = 84$, 16,3 %) también son equiparables a los publicados por otros grupos de trasplante. En este caso, la recurrencia en el grupo de entrenamiento fue superior (18,2 % vs. 10,8 %), probablemente por el mayor tiempo de

seguimiento de este grupo, aunque las diferencias tampoco alcanzaron significación estadística. Se observa un claro predominio de la recurrencia extrahepática, que se produce en más de la mitad de los casos que recidivan.

En cuanto a la variable objetivo del trabajo, la recurrencia precoz de la serie se sitúa en el 10,5 %, siendo ésta ligeramente superior en el grupo de entrenamiento (11,4 % vs. 7,7 %), de nuevo sin diferencias significativas. Si bien es cierto que no existe mucha literatura analizando en concreto la recurrencia precoz postrasplante del CHC, los estudios disponibles la sitúan en torno al 10 % (142) en los 2 primeros años, con un importante impacto sobre la supervivencia a largo plazo de este subgrupo de pacientes. Este hecho queda demostrado en nuestro estudio, en el que la supervivencia global de los pacientes con recurrencia precoz se sitúa en el 32 % a los 3 años del TH, y en el 16,4 % a los 5 años del trasplante.

En el presente estudio, no se han tenido en cuenta los datos procedentes del análisis anatomopatológico de los hígados explantados, por tratarse de un dato sólo disponible después del trasplante hepático. El modelo creado pretende predecir la recurrencia precoz con variables sólo disponibles previamente al trasplante, anticipándonos a la posible recidiva. Sin embargo, es de suponer que aquellos pacientes con mejor respuesta radiológica y, por tanto, menor riesgo de recurrencia precoz y mejor supervivencia global, habrán presentado una mejor respuesta patológica, en consonancia con lo publicado ya por distintos grupos de trasplante, tanto en estudios

unicéntricos como el de UCLA en el año 2015 publicado por Agopian et al. (143), como en estudios multicéntricos (49,55). Estos estudios objetivan supervivencias globales del 84 % y el 75 % a 3 y 5 años postrasplante para pacientes con respuesta patológica completa, así como una recurrencia del 5,2 % a 5 años postrasplante, frente al 16,4 % de recurrencia y 68 % de supervivencia global a los 5 años postrasplante en pacientes sin respuesta patológica completa. En este sentido, se ha apuntado que un número elevado de tratamientos locorregionales (3 o más) previos al trasplante se correlaciona con una mayor dificultad para obtener una respuesta patológica completa y, por tanto, con un peor pronóstico del CHC después del TH (144). En nuestro estudio, el número de tratamientos pretrasplante no forma parte de las variables predictoras de la recurrencia precoz del CHC del modelo predictivo de RB.

6.2.- Estudio estadístico y modelos predictivos

Tomando como base una serie de casos amplia (teniendo en cuenta que se trata de un estudio unicéntrico) a lo largo de un período de 15 años, se elaboraron los modelos predictivos basados en regresión logística y redes bayesianas con un objetivo eminentemente práctico, enfocado a su aplicabilidad en la toma de decisiones en el marco de la *transplant oncology*.

Es importante destacar que los modelos se basan exclusivamente en variables del receptor y del tumor injertado (y su comportamiento frente a tratamientos pretrasplante). Esto permite identificar aquellos pacientes con mayor riesgo de recurrencia precoz de CHC y, por tanto, de fracaso del TH con indicación oncológica, y así lograr una selección más adecuada de los candidatos a TH optimizando el uso de los injertos disponibles, mejorando los resultados de supervivencia libre de enfermedad y global. Además, dado que no tiene en cuenta variables del proceso de donación (tipo y características del donante, tiempos de isquemia...), el modelo puede emplearse para decidir si un paciente se incluye en lista de espera en base a su riesgo de recurrencia y no en función del donante, facilitando mucho la estrategia de toma de decisiones.

La aplicabilidad práctica consistiría en no ofrecer el TH a aquellos receptores cuya probabilidad de recurrencia antes de 2 años se encontrase por encima de un umbral prefijado. Para determinar ese umbral, es importante destacar que aquellos receptores cuyos tumores injertados tienen un peor comportamiento biológico (traducido en su

respuesta al tratamiento pretrasplante y sus niveles de AFP) tienen una mayor probabilidad de recurrencia precoz y, por tanto, existe un elevado riesgo de que el TH sea fútil. En este sentido y con vistas a su aplicación clínica, cabe recordar que el modelo predictivo se ha diseñado para que ofrezca un resultado discriminante según el umbral fijado previamente, aunque también establezca probabilidades de recurrencia precoz del CHC para cada caso concreto, en base a los datos con los que ha sido generado.

Así pues, teniendo en cuenta los datos disponibles en la literatura, la experiencia clínica y el objetivo del estudio, se determinó la importancia de los posibles resultados que se pueden obtener tras la aplicación de la prueba (valores de utilidad del modelo). Dado que se pretende evitar las consecuencias negativas que conlleva una recurrencia precoz del CHC, el principal objetivo es conseguir una elevada especificidad en el modelo. Por tanto, se asignó el máximo valor de utilidad (valor de 1) a los verdaderos positivos, esto es, aquella situación en la que se cumple la predicción del modelo de una recurrencia del CHC en los dos primeros años después del TH. Desde este punto de vista, se pretende que el modelo realice una predicción segura de un escenario clínico concreto (realización de TH en receptores con alto riesgo de riesgo de recidiva precoz) con unas consecuencias negativas trascendentales.

Por otra parte, a la utilidad de los falsos negativos se le ha asignado un valor muy negativo (- 0,9). En esta situación, el modelo predice erróneamente que no se va a producir una recurrencia precoz

del CHC, mientras que en la situación real se produce una recidiva en los dos primeros años después del TH, con las consiguientes consecuencias negativas que este escenario clínico supone en cuanto a disminución de la supervivencia del receptor, necesidad de tratamientos para controlar la enfermedad en un paciente inmunosuprimido, y una utilización subóptima de un injerto hepático.

La utilidad del verdadero negativo también ha recibido un valor relativamente alto (0,85), puesto que en esta situación el modelo predice adecuadamente una ausencia de recidiva precoz del CHC, que es efectivamente la situación que se pretende evitar, destinando así el injerto a un paciente con bajo riesgo de recidiva y que tiene más probabilidad de obtener una supervivencia libre de enfermedad y global más elevada.

Finalmente, a la utilidad falso positivo se le ha otorgado un valor más bajo (0,25). En esta situación el modelo falla en su predicción de recurrencia, y en este escenario el receptor no recibe un trasplante como tratamiento para el CHC. Sin embargo, si se aplica igualmente el modelo como ayuda a la toma de decisiones, el injerto habría sido destinado a otro paciente en lista de espera, por lo que efectivamente, dicho injerto no se pierde, aunque no favorece al receptor en cuestión.

A continuación, mostramos una tabla resumen de los valores de utilidad otorgados en cada escenario clínico:

	Modelo	Realidad	Función de utilidad
VP	Predice recidiva precoz	Sí hay recidiva precoz	1
FN	Predice no recidiva precoz	Sí hay recidiva precoz	-0.9
VN	Predice no recidiva precoz	No hay recidiva precoz	0,85
FP	Predice recidiva precoz	No hay recidiva precoz	0,25

Tabla 30: valores de utilidad otorgados a los diferentes escenarios clínicos según la predicción del modelo

6.2.1.- Estudio univariante

Antes de elaborar los modelos predictivos, se llevó a cabo un análisis univariante, con la intención de determinar la relación de cada variable con la variable objetivo del estudio. En este análisis se consideró una relación estadísticamente significativa aquella con un valor $p < 0,1$ tras la aplicación de la prueba estadística correspondiente en cada caso. Este umbral, más alto de lo habitual, se utilizó con el objetivo de evitar la pérdida de variables que, por sí mismas o en combinación con otras, pudieran tener un impacto sobre el resultado final.

Tras el estudio univariante, las variables que mostraron una relación estadísticamente significativa con la recurrencia precoz postrasplante del CHC fueron:

- Edad del receptor

- Tamaño del nódulo tumoral más grande
- Terapia previa al trasplante sobre el CHC (sí/no)
- QETA pretrasplante (sí/no)
- Número de tratamientos sobre el CHC previos al trasplante
- Respuesta a terapia pretrasplante
- *Downstaging* a criterios de Milán (sí/no)
- Estado después del tratamiento pretrasplante
- AFP máxima en seguimiento pretrasplante:
 - AFP < 400 ng/ml
 - AFP $\geq$ 400 nm/mL y < 1000 ng/ml
 - AFP $\geq$ 1000 nm/mL
- Valor de la AFP pretrasplante (logaritmo en base 10)

En primer lugar, tanto la edad del receptor, como el tamaño del nódulo más grande detectado previamente al trasplante obtuvieron significación estadística en el estudio univariante.

Asimismo, de forma global, el análisis univariante identifica como significativas variables que en la literatura se relacionan con el comportamiento biológico del tumor y, por tanto, probablemente con el riesgo de recurrencia precoz del CHC después del trasplante. Así, el hecho de haber recibido un tratamiento pretrasplante y la respuesta a este tratamiento neoadyuvante, independientemente de su objetivo (*downstaging* o puente), es un marcador de supervivencia

postrasplante y su contribución a la disminución del riesgo de recurrencia está bien establecido en la literatura. De este modo, esta variable se ha incorporado a las estrategias de toma de decisión sobre inclusión y priorización en lista de espera, dado que se relaciona con factores anatomopatológicos de mal pronóstico (20). En 2016 el TRAIN score (77) demostró una selección más adecuada de pacientes con respecto a los criterios de Milán (incrementando en más de un 20 % los pacientes potencialmente trasplantables), así como una mejor capacidad predictiva de la recurrencia postrasplante y supervivencia relacionada con el CHC, comparando ambos sistemas mediante curvas ROC, con una metodología similar a la desarrollada en este estudio. Por su parte, también se ha demostrado que sistemas predictivos bien establecidos como el *Metroticket* mejoran su capacidad predictiva incorporando la respuesta a los tratamientos previos en base a criterios mRECIST (76).

En este sentido, variables que complementan la información en torno a la respuesta tumoral al tratamiento pretrasplante, como el hecho de haber conseguido un *downstaging* efectivo a criterios morfológicos de Milán, o el número de sesiones de terapia pretrasplante recibidos, también tienen una relación estadísticamente significativa con la variable objetivo en el estudio univariante. Probablemente, la presencia de la QETA como variable significativa en el análisis univariante tenga que ver con que se trata de la modalidad terapéutica más empleada entre los pacientes del estudio (57,7 %).

Con el fin de integrar toda la información disponible con respecto al comportamiento tumoral después del tratamiento, se elaboró una variable combinada (“Estado después del tratamiento pretrasplante”) que identifica los pacientes con respecto a la terapia aplicada sobre ellos durante el seguimiento pretrasplante y el resultado de la misma:

- No tratamiento
- Tratamiento puente al TH
- *Downstaging* a criterios de Milán
- *Downstaging* a criterios *Up-to-seven*
- Fallo de *dowstaging*

Por tanto, al condensar la información combinada de variables otras descritas anteriormente, es lógico que esta variable se relacione también con la variable objetivo del estudio.

Por su parte, la AFP también se relaciona en el estudio univariante con la recurrencia precoz postrasplante.

Los valores pico alcanzados durante la evolución previa al trasplante y los niveles pretrasplante han mostrado ambos significación estadística, refrendando los datos disponibles en la literatura al respecto del valor pronóstico de este marcador biológico, tanto en términos absolutos o valores fijos (47,74) como en su comportamiento dinámico después de tratamientos pretrasplante (145).

6.2.2.- Modelo de regresión logística

La regresión logística (RL) es un método de análisis estadístico ampliamente utilizado en el campo de la investigación biomédica para tratar de establecer relaciones probabilísticas (causales) entre variables independientes y una variable dependiente (variable objetivo) categórica dicotómica. Este modelo es de gran utilidad en medicina, puesto que las variables resultado suelen ser binarias y expresarse como vivo/muerto, sano/enfermo o, como en este estudio, recurrencia precoz/no recurrencia precoz.

Desde un punto de vista meramente matemático, la RL trata de determinar probabilidades reduciendo al máximo el número de variables incluidas en la fórmula definitiva, de tal manera que sólo se incorporan aquellas que alcanzan una significación estadística prefijada. En este sentido, la RL corre el riesgo de ser excesivamente reduccionista, especialmente a la hora de analizar problemas complejos, puesto que tenderá a desechar variables que, aunque no alcanzan una relación estadísticamente significativa, pueden tener importancia desde un punto de vista clínico.

Es evidente que predecir el comportamiento biológico del CHC y la recurrencia precoz postrasplante del mismo es un problema complejo, en el que influyen múltiples variables que pueden tener una relación no lineal entre ellas. Si lo que se pretende es ayudar a la toma de decisiones clínicas en el ámbito de la *transplant oncology* e indicar un trasplante hepático según el riesgo de recurrencia, es posible que la RL ofrezca una información limitada.

Por contra, los modelos predictivos basados en inteligencia artificial y *machine learning*, como las redes bayesianas, pueden ser de utilidad para situaciones complejas, en las que no existen reglas sistemáticas que definan completamente el problema.

A pesar de sus limitaciones, la RL se ha utilizado como referencia para comparar el modelo desarrollado mediante redes bayesianas por ser el *gold standard* en el campo de la estadística biomédica. Si bien es cierto que existe un creciente interés por los modelos de inteligencia artificial aplicados al TH, como ocurre con las publicaciones del grupo de Javier Briceño (146–148) en relación al emparejamiento donante-receptor, las publicaciones referentes a la *transplant oncology* aún son escasas (149).

Para la elaboración del modelo de RL se empleó la metodología de retirada secuencial de variables: sobre el conjunto de 385 pacientes del grupo de entrenamiento se analizaron las 32 variables recogidas para cada individuo. Desde este “modelo máximo”, se fueron retirando aquellas variables cuya presencia no mejoraba el rendimiento del modelo en cuanto a su capacidad predictiva (expresada como el ABC de la curva ROC).

Finalmente, las variables incluidas en la fórmula definitiva del modelo predictivo de RL fueron:

- Edad del receptor
- Tamaño del nódulo tumoral más grande
- Respuesta a terapia pretrasplante

- Valor de la AFP pretrasplante (logaritmo en base 10)

En primer lugar, la edad del receptor también se incluyó como variable significativa en el análisis multivariante de RL. Tradicionalmente, es un factor que no se ha tenido en cuenta a la hora de valorar la recurrencia y supervivencia postrasplante. Un estudio reciente publicado en 2021 por Goldberg et al., en base a una cohorte de 6502 pacientes del registro norteamericano de TH (*Organ Procurement and Transplantation Network / United Network for Organ Sharing*), considera la edad como factor pronóstico, estableciendo un *score* mediante el método de *random forest* basado en variables relacionadas con el CHC (carga tumoral total en el momento de inclusión en lista, carga tumoral total pretrasplante, AFP en el momento de inclusión en lista, AFP pretrasplante) y otras no relacionadas con el CHC (edad, función renal, INR, bilirrubinemia...), para determinar la supervivencia relacionada con el CHC (150).

Anteriormente hemos comentado la relación de la respuesta a los tratamientos locorregionales con la supervivencia global postrasplante. Del mismo modo, esta respuesta, que traduce el comportamiento biológico del CHC, tiene una correlación bien documentada en la literatura con la recurrencia postrasplante y la supervivencia libre de enfermedad, ya que una buena respuesta radiológica según los criterios mRECIST suele correlacionarse con factores anatomopatológicos de buen pronóstico (por ejemplo, ausencia de invasión microvascular) (144), así como una mayor probabilidad de respuesta patológica completa (55). Todos ellos son factores

relacionados con una mayor SLE y menor riesgo de recurrencia, incluso en pacientes inicialmente fuera de criterios de Milán, como demuestra el trabajo de Kornberg et al. (151), en el que pacientes con respuesta completa inicialmente fuera de criterios de Milán presentaron resultados similares de supervivencia global a los 5 años postrasplante (80%) a los pacientes que sí cumplían los criterios de Milán.

Del mismo modo, los niveles de AFP pretrasplante (expresados de forma logarítmica) también demostraron una relación significativa con la recurrencia precoz en el modelo de RL, en consonancia con los datos disponibles en la literatura. A pesar de que los criterios morfológicos siguen siendo la base de las políticas de indicación de TH en la mayoría de centros de todo el mundo, en los últimos 10 años existe una creciente evidencia de la mejora de la capacidad predictiva de los sistemas y *scores* pronósticos con la incorporación de la AFP: desde la publicación del modelo francés en el año 2012 y su posterior validación (47,74), pasando por el Metroticket y su evolución 2.0 que incorpora la respuesta a los tratamientos locorreregionales (76,141) o el sistema TRAIN (77). Una reciente revisión sistemática publicada por J. Lerut identificaba en la literatura más de 20 *scores* y criterios de selección para TH que incorporaban la AFP a los criterios morfológicos.

Sin embargo, en el modelo predictivo de RL tras el estudio multivariante, no alcanzaron significación estadística los niveles máximos de AFP durante el seguimiento pretrasplante, por lo que, de

alguna manera, este modelo no ha tenido en cuenta el comportamiento dinámico de la AFP previamente al TH, factor que se ha relacionado con el comportamiento biológico del CHC y el riesgo de recurrencia postrasplante y SLE (48,77,145,152,153).

El modelo de RL sí tuvo en cuenta los factores morfológicos y su relación con la recurrencia precoz, al identificar el tamaño del nódulo más grande como variable estadísticamente significativa en el análisis multivariante. El diámetro del nódulo dominante ha sido incluido en distintos *scores* pronósticos, tanto morfológicos puros como combinados y, en general, con un límite a partir del cual se desestima el TH por un excesivo riesgo de recidiva. Entre los criterios morfológicos expandidos, el diámetro del nódulo mayor se ha combinado tradicionalmente con el número de tumores para establecer un límite a partir del cual se rechaza el trasplante. Ejemplos de ello son los criterios *Up-to-seven*, o los criterios de Seúl publicados por Choi en 2012 (hasta 7 nódulos con 7 cm de diámetro del tumor más grande) (154), si bien es cierto que estos últimos sólo se aplican en trasplante de donante vivo.

Más frecuente ha sido tratar de expandir las indicaciones combinando el diámetro del nódulo mayor con algún tipo de marcador biológico, preferentemente la AFP o la DCP. El grupo de Choi publicó en 2013 (155) unos criterios combinados con un límite de diámetro del tumor más grande ≤ 5 cm y un límite de AFP ≤ 100 ng/ml. Aunque se aplicaban a trasplante de donante vivo, los resultados de SLE fueron similares a los criterios de Milán. De forma análoga, otros grupos han

ido publicando su experiencia con criterios combinados expandidos, siempre con resultados similares a los de Milán, con el fin de intentar ofrecer el TH con intención curativa a un mayor número de pacientes (aunque la mayoría de ellos son series unicéntricas con criterios no validados externamente). Ejemplos de ello son:

- Wan (Shanghai): tumor máximo ≤ 10 cm + AFP ≤ 400 n/mL (156).
- Taketomi (Kyushu): tumor máximo ≤ 5 cm + DCP ≤ 300 mAU/mL (157).
- Halazun (*New York – California score*): distintos diámetros con distintos valores entre AFP máxima y AFP pretrasplante (48)

Resulta llamativo que determinadas variables que alcanzan significación estadística en el estudio univariante, no hayan sido incluidas en el modelo de RL después del análisis multivariante. Así, el pico máximo de AFP durante el seguimiento pretrasplante (400 ng/ml o 1000 ng/ml) es un ejemplo de ello. Unos valores elevados de AFP previos al trasplante se han correlacionado con malos resultados en cuanto a recurrencia y SLE, pero en la literatura se ha demostrado que la variación en sus niveles, especialmente en el contexto de terapias locorregionales pretrasplante, también tienen un peso considerable a la hora de determinar el pronóstico de los pacientes. Ello adquiere mayor importancia si tenemos en cuenta las características de nuestra serie de estudio, donde aproximadamente el 75 % de los individuos han recibido algún tipo de tratamiento previo al TH. En este sentido, otra variable que tampoco ha superado el corte del análisis

multivariante es el número de tratamientos previos al TH que han recibido los pacientes y que, como hemos visto, ha sido relacionada en la literatura como factor pronóstico de recurrencia y SLE.

Pese a estas limitaciones (atribuibles a la propia naturaleza metodológica) el modelo predictivo de regresión logística obtuvo un ABC de 0,791 para el grupo de entrenamiento (aceptable), que asciende hasta 0,843 para el grupo de generalización (buena). Estos resultados desiguales, junto con el hecho de que el modelo haya omitido determinadas variables o relaciones entre ellas que puedan ser de importancia en la práctica clínica, plantea dudas de una posible aplicación clínica del modelo de RL.

Teniendo en cuenta que los valores de utilidad asignados en este estudio están orientados a evitar una recurrencia precoz y, por tanto, un TH fútil desde el punto de vista oncológico, los modelos de RL obtuvieron valores bajos de sensibilidad (entrenamiento 38,6 % - generalización 40 %) en comparación con su especificidad (entrenamiento 94,4 % - generalización 95 %). Igualmente, los valores predictivos positivos se sitúan por debajo del 50 % en ambos grupos, mientras que los valores predictivos negativos son del 92,3 % y 95 %, respectivamente. De este modo se consigue que el modelo no falle al predecir una ausencia de recurrencia precoz, por lo que conseguimos realizar un trasplante en un paciente con bajo riesgo de recidiva.

6.2.3.- Modelo de redes bayesianas

Las RB son modelos gráficos probabilísticos que permiten, debido a su funcionamiento basado en el Teorema de Bayes, establecer la probabilidad de que un evento ocurra dados ciertos efectos observados. Por tanto, su aplicación en ciencias de la salud puede ser de mucha ayuda en la predicción de eventos y en la toma de decisiones clínicas en entornos complejos. Comparativamente con la RL, este modelo de RB se ha incorporado recientemente a los métodos estadísticos en investigación biomédica, especialmente en investigación clínica aplicada, si bien es cierto que existe un creciente interés debido a sus potenciales usos para la resolución de problemas complejos. Además, el hecho de representar gráficamente las distribuciones de probabilidad entre variables mediante nodos y arcos hace que su interpretación resulte muy intuitiva, permitiendo resumir con cierta sencillez problemas complejos o en los que existen múltiples relaciones causales entre variables.

Por otra parte, se trata de un tipo de modelo predictivo de inteligencia artificial (*machine learning*) que no requiere un tamaño muestral importante para ser desarrollado. Su versatilidad hace que este método pueda ser aplicado tanto a conjuntos relativamente pequeños como a grandes bases de datos (por ejemplo, registros nacionales o internacionales de TH). Además, su gran plasticidad hace que puedan ir incorporándose nuevos registros a la base sobre la que se construye el modelo, de tal manera que la estructura del grafo puede volver a “aprender sobre la marcha” y determinar nuevas

relaciones probabilísticas, incrementando así su rendimiento predictivo.

Por otra parte, las RB tienen la ventaja de poder incluir de manera sencilla y matemáticamente coherente el conocimiento previo acerca del problema analizado, combinándolo con los hallazgos de la investigación en curso y potenciando la capacidad predictiva del modelo. Finalmente, una vez generado el modelo predictivo, éste contiene una distribución de probabilidad para cada combinación de valores de variables, por lo que es capaz de generar respuestas rápidas cuando se introducen valores para obtener una determinada información, que puede ser necesaria para la toma de decisiones.

Por todo ello, las RB son modelos que están orientados a su aplicación en entornos de incertidumbre o en aquellas situaciones complejas en las que múltiples variables pueden tener un peso específico y estar relacionadas de forma sutil o no bien determinada entre sí y con la variable objetivo. Así pues, son una herramienta que puede ser de gran ayuda en la toma de decisiones clínicas en un campo tan actual y dinámico como la *transplant oncology*.

En este sentido, nuestro modelo toma como base unas variables conocidas previamente e interconectadas entre sí mediante distribuciones de probabilidad. Después de una fase de entrenamiento, genera una salida, que en el caso de este estudio supone determinar la probabilidad de que ocurra una recurrencia del CHC en los dos primeros años después del TH.

Las RB presentan, no obstante, algunas desventajas. El principal inconveniente reside en que este tipo de modelo predictivo tiene dificultades a la hora de razonar con variables continuas, de uso común en investigación clínica. La solución pasa por discretizar este tipo de variables, convirtiéndolas en categóricas. Se corre el riesgo, sin embargo, de perder información sobre la distribución de la variable original al discretizarla, perdiendo por tanto capacidad predictiva ya que, además, la forma de discretizar las variables (número de intervalos, puntos de corte, etc.) puede tener impacto en el modelo resultante y su rendimiento. Este inconveniente puede compensarse complementando la información de las variables discretizadas con conocimiento previo o de expertos en el problema analizado. Sin embargo, este método puede asimismo plantear otros problemas, puesto que ese conocimiento previo puede emplearse para ayudar a elaborar la estructura del grafo, para determinar las distribuciones de probabilidad entre las variables de este, o para ambos propósitos. Por tanto, el modelo final y su capacidad predictiva puede ser distinto en función de la influencia externa que se aplique en su elaboración.

En este sentido, el modelo de RB utilizó un método de *machine learning* automático y autónomo, y se elaboró sobre las mismas variables con las que se generó el modelo de RL, con la única salvedad de la discretización de las variables continuas.

Al analizar la estructura del grafo, observamos que existe un nodo “padre” que es la variable objetivo (recurrencia precoz) y 5 nodos “hijos” relacionados con dicha variable, que son:

- Respuesta a terapia pretrasplante
- Estado después del tratamiento pretrasplante
- *Downstaging* a criterios de Milán (sí/no)
- AFP máxima en seguimiento pretrasplante:
 - AFP < 400 ng/ml
 - AFP $\geq$ 400 nm/mL y < 1000 ng/ml
 - AFP $\geq$ 1000 nm/mL
- Valor de la AFP pretrasplante (logaritmo en base 10), estableciendo un punto de corte en 9 ng/ml.

Como se puede comprobar, todas ellas son variables que hacen referencia al comportamiento biológico del tumor, tanto en relación al tratamiento previo al trasplante como a la evolución de un marcador biológico cuya relación con la recurrencia postrasplante está ampliamente demostrada en la literatura. En este sentido, el modelo es comparable a otros publicados. No existen variables relacionadas con las características del receptor ni con el donante, por lo que consideramos que el modelo utiliza correctamente la información disponible para caracterizar de una forma más precisa la recurrencia precoz del CHC, seleccionando las variables de forma adecuada.

De entre todas las variables independientes, aquella que adquiere mayor peso en la estructura del grafo es la respuesta al tratamiento pretrasplante, de tal manera que el resto de las variables también tienen una relación probabilística con ella. Esto es lógico,

especialmente en el caso del estado después del tratamiento pretrasplante, puesto que ambas variables comparten parte de la información y describen dos caras de una misma moneda, como es el comportamiento del tumor previamente al trasplante frente a los posibles tratamientos que se hayan podido aplicar.

De la misma manera, la presencia de una respuesta completa (o como mínimo parcial) es condición *sine qua non* para que se produzca un *downstaging*, por lo que también resulta lógico que ambas variables estén relacionadas.

Por otra parte, el modelo incluye también la relación de la AFP con la respuesta al tratamiento locorregional, aunando las dos armas de las que disponemos para evaluar el comportamiento del CHC: el diagnóstico por imagen radiológico y los marcadores serológicos. En este sentido, el modelo ofrece una información completa, puesto que incluye no sólo el valor pretrasplante de la AFP, sino también el valor máximo que ha presentado este marcador durante el seguimiento pretrasplante del paciente (discretizado en 3 intervalos: ≤ 400 ng/ml, 401-999 nd/dL, ≥ 1000 ng/ml). De esta forma, el modelo sintetiza la evolución de los niveles de AFP durante el seguimiento pretrasplante, de la misma forma que hacen otros modelos, como el anteriormente mencionado de Goldberg et al. (150) o los modelos publicados por Lai et al. (TRAIN *score* y EurHeCaLT) (77,153), si bien es cierto que en estos casos, se incluye una curva específica de descenso de AFP y no límites o valores concretos.

Llama la atención que el tamaño tumoral no figure entre las variables incluidas en el modelo de RB. Esto podría deberse a que en el centro en el que se ha realizado el estudio, se aplican unos criterios de selección estrictos para la inclusión en lista de espera de los pacientes.

Desde el punto de vista radiológico, en nuestra serie, sólo un 5,44 % de los pacientes se trasplantaron fuera de los criterios de Milán (con pruebas de imagen) y menos de un 2 % se trasplantaron teóricamente fuera de los criterios *Up-to-seven*, también con datos radiológicos (todos ellos en el grupo de entrenamiento y ninguno en el grupo de generalización). Por tanto, se trata de una serie con escasas diferencias en cuanto a la carga tumoral (tamaño de los nódulos, tamaño tumoral total, número de lesiones, etc.) lo que explicaría que ninguna de estas variables haya logrado ser incluida en el modelo definitivo de RB, siendo éste un potencial sesgo de selección.

La asunción de que el tamaño tumoral en base a criterios morfológicos, por sí misma, no es una contraindicación para el TH, y que los criterios morfológicos por sí solos proporcionan una información insuficiente para predecir la SLE en el CHC es una idea extendida en la literatura referida a la *transplant oncology*. El estudio de Kornberg et al. (151) hipotetiza que una respuesta patológica adecuada, independientemente del tamaño, y valorada previamente al trasplante con PET-TC, se correlaciona con excelentes resultados de SLE. Por su parte, el *TRAIN score* (77) únicamente tiene en cuenta la respuesta tumoral a los tratamientos locorregionales pretrasplante

mediante los criterios mRECIST y su reflejo en la curva de AFP y el ratio neutrófilos/linfocitos. Otros criterios de trasplante (como los del grupo de Toronto) van más allá, y no ponen ningún límite en cuanto a la carga tumoral, siempre y cuando no se trate de un tumor pobremente diferenciado y de pacientes con síntomas sistémicos (15).

Las ratios neutrófilos/linfocitos y neutrófilos/plaquetas tampoco fueron incluidas entre las variables del modelo de RB. Esto puede deberse a que, de forma distinta a la AFP, no se han recogido datos sobre su evolución, sino simplemente un valor fijo, por lo que la información que incluye la variable podría ser insuficiente para predecir el comportamiento del CHC y su recurrencia. Del mismo modo, tampoco se han incluido variables relacionadas con la etiología de la hepatopatía, ni con la gravedad de la misma (Child-Pugh o MELD), si bien es cierto que son escasos los criterios o *scores* pronósticos que tienen en cuenta estos factores desde el punto de vista de la recurrencia y SLE, estando más relacionados con la supervivencia global y mortalidad específica relacionada con el CHC.

En cuanto a su capacidad predictiva, el ABC obtenida para el grupo de entrenamiento en el modelo de RB supera el 0,8 (0,817), siendo considerada buena, por lo que estaría justificada su utilización en la práctica clínica. Además, el modelo posee una elevada especificidad (90,3 %), lo que minimiza la presencia de falsos positivos. El altísimo valor predictivo negativo del modelo (93,6 %), evita que se produzca una recurrencia precoz, en pacientes en los que el modelo predice que

no van a presentarla y, por tanto, optimizando así el uso de los injertos implantados.

Por otra parte, el modelo de RB también tiene una sensibilidad superior al de la RL para el grupo de entrenamiento (52,3 % vs. 38,6 %), por lo que se reduce el número de falsos negativos y, por tanto, el número de recurrencias precoces, de injertos perdidos y de trasplantes considerados fútiles.

6.3.- Comparación de los modelos desarrollados

6.3.1.- Comparación de las variables incluidas en los modelos

De entre todas las variables analizadas en el estudio, los modelos de regresión logística y redes bayesianas comparten dos de ellas: la respuesta a las terapias pretrasplante y los niveles de AFP pretrasplante. Se trata de dos parámetros bien conocidos en el campo de la *transplant oncology* que traducen, en cierta manera, el comportamiento biológico del tumor y que aportan información previamente al TH, que resulta de utilidad para predecir una posible recurrencia o una mayor o menor agresividad del CHC. Además, cumplen la condición de ser variables fácilmente disponibles en la práctica clínica antes de establecer la indicación de TH, por lo que pueden ser de ayuda en la toma de decisiones sobre si proceder con el trasplante en cada caso concreto.

Comparativamente, el modelo de RL ha incluido el tamaño del nódulo tumoral mayor entre las variables significativas, sin tener en cuenta la suma total de los diámetros de los nódulos (tamaño tumoral total), ni tampoco el número de lesiones. Por ello, quizá la información que aporta el modelo acerca de la carga tumoral sea incompleta. No obstante, hay una evidencia creciente al respecto de que la información meramente morfológica es insuficiente para valorar la agresividad del CHC. En este sentido, el modelo de RB es una prueba de ello, puesto que ninguna de las variables acerca de la carga tumoral

ha logrado ser incluida como relevante. Esto puede deberse a dos factores:

- En primer lugar, es posible que exista un sesgo de selección. Nuestro estudio analiza únicamente pacientes ya trasplantados en base a criterios morfológicos. Sólo un pequeño porcentaje de pacientes se trasplantaron con dudas acerca de su carga tumoral, por lo que esta homogeneidad en la muestra puede haber influido en que el tamaño de los nódulos o su número no hayan alcanzado significación estadística.

- Por otra parte, el modelo de RB incide en la idea de que el tamaño tumoral puede no ser un criterio a considerar a la hora de realizar un TH, siempre y cuando el paciente haya recibido un tratamiento pretrasplante, haya presentado una respuesta adecuada en base a criterios mRECIST y se acompañe de unos marcadores biológicos (fundamentalmente AFP) en rangos aceptables. En este sentido, quizá sea el momento de abandonar definitivamente la morfología como el pilar central de la *trasplant oncology* y poner el comportamiento biológico del tumor en el centro de la toma de decisiones.

Teniendo en cuenta estas consideraciones, el modelo de RB incluye 5 variables que, en conjunto, aportan más información acerca del comportamiento biológico del tumor. Al considerar tanto la respuesta a los tratamientos pretrasplante como el estado del paciente después de los mismos (tratamiento puente, *dowstaging*, indicación de trasplante fuera de criterios morfológicos...) aporta una información dinámica acerca del comportamiento del CHC a lo largo de su

evolución previa al TH. Esta información se complementa con los valores de AFP, por lo que el modelo de RB permite, de alguna manera, saber si el comportamiento radiológico del tumor tiene un reflejo en los marcadores biológicos (AFP) y si ésta presenta una evolución aceptable y se mantiene o no dentro de rangos considerados seguros desde el punto de vista del riesgo de recurrencia.

6.3.2.- Comparación de los resultados obtenidos en el grupo de entrenamiento

A nivel metodológico, se ha escogido el ABC de las curvas ROC como medida de rendimiento de la capacidad predictiva de los modelos desarrollados, así como la sensibilidad, especificidad y valores predictivos positivo y negativo. Se trata de parámetros simples y de amplia utilización en investigación clínica, lo que permite extrapolar los resultados y compararlos, no solo entre nuestros modelos sino con otros que hayan sido publicados. En origen, las curvas ROC se emplearon con el fin de analizar la potencia de pruebas diagnósticas. Sin embargo, hoy en día su uso se ha extendido a la evaluación del rendimiento diagnóstico de modelos matemáticos (77,122).

Si bien es cierto que las curvas ROC permiten comparar el resultado de distintos tipos de modelos predictivos, es importante destacar que, desde el punto de vista metodológico, los modelos desarrollados utilizan herramientas diferentes para generar una predicción. Así, la RL parte de todas las variables disponibles para

buscar el modelo más sencillo. Esto es, en base a unas pruebas estadísticas que aplica a las variables independientes, determina si cada una de ellas alcanza un nivel estadístico suficiente, buscando la combinación adecuada para que, con el menor número de variables posible, se resuelva de la mejor forma el problema planteado.

Por su parte, las RB pretenden encontrar la mejor relación entre variables independientes y la variable objetivo. Es decir, no trata de eliminar variables que puedan considerarse “poco útiles” porque no superen un umbral, sino que analiza las distribuciones de probabilidad entre las variables consideradas y escoge la combinación de variables cuya distribución resuelva mejor el problema, por lo que es capaz de incluir variables con una relación más sutil entre sí y con la variable objetivo, si esto aumenta la capacidad predictiva del modelo.

Por tanto, aunque se trata de modelos metodológicamente diferentes, el hecho de presentar los resultados con un mismo lenguaje permite la comparación de su rendimiento predictivo, lo que supone una ventaja considerable.

Para el grupo de entrenamiento, el modelo de RL obtuvo una ABC de 0,791 (IC 95 %: 0,717 – 0,864), mientras que el modelo de RB obtuvo una ABC de 0,817 (IC 95 %: 0,717 – 0,888), por lo que su capacidad predictiva en el grupo de entrenamiento es superior y puede considerarse como buena y clínicamente aceptable ($> 0,8$), si bien es cierto que las diferencias entre ambas ABC no alcanzaron significación estadística ($p = 0,2$), por lo que es posible que las diferencias entre ambos se deban al azar. No obstante, la concordancia entre ambos

modelos para este grupo de pacientes se considera como moderada, con un coeficiente kappa de 0,485 (IC 95 %: 0,338 – 0,632). Sin embargo, es posible que esto se deba también a que el tamaño muestral sea insuficiente para que las diferencias consigan una significación estadística.

6.4.- Aplicación de los modelos desarrollados en el grupo de generalización. Comparación de resultados

Tras la creación de los modelos predictivos de RL y RB a partir del grupo de entrenamiento, que incluía el 75 % de todos los pacientes del estudio ($n = 385$), estos modelos fueron aplicados al grupo de generalización, que incluye al 25 % restante ($n = 130$) y que son pacientes receptores de TH consecutivos e inmediatamente posteriores en el tiempo, que cumplían los mismos criterios de inclusión y exclusión que los pacientes del grupo de entrenamiento y que fueron manejados clínicamente de forma similar.

Se trata de grupos comparables, especialmente desde el punto de vista de las variables referidas a la carga tumoral de los pacientes puesto que no se objetivaron diferencias estadísticamente significativas en cuanto al tamaño tumoral (suma de los diámetros de los nódulos y tamaño del nódulo mayor), número de nódulos, o cumplimiento de criterios de Milán y *Up-to-seven*. Tampoco se objetivaron diferencias en cuanto a la proporción de pacientes que recibieron tratamiento locorregional previamente al trasplante, ni en su respuesta. En general el tipo de terapia previa fue similar, con predominio de la QETA como técnica preferente, seguido de la RF. Sin embargo, en el grupo de generalización se observó un incremento significativo de pacientes tratados mediante RETA (3,8 % vs. 0,5 %) debido a la progresiva implantación de esta modalidad de tratamiento y sus buenos resultados. En cuanto a la respuesta a estos tratamientos, no se identificaron diferencias significativas entre grupos

(con un predominio de las respuestas completas o parciales), ni en la proporción de pacientes que experimentaron un *dowstaging* a criterios de Milán o *Up-to-seven*.

Ahora bien, sí se evidenciaron ciertas diferencias significativas en las características de los receptores, con una mayor proporción de pacientes en estadio Child A en el grupo de generalización (63,8 % vs. 47,5 %) y una puntuación MELD ligeramente inferior en el grupo de generalización (11 grupo entrenamiento vs. 9 grupo de generalización, $p = 0,01$). Esto probablemente se deba a las mejoras en el manejo pretrasplante de los pacientes gracias a la experiencia adquirida a lo largo del tiempo, consiguiendo que los pacientes lleguen al trasplante con un mejor control de su hepatopatía de base.

La otra gran diferencia significativa la encontramos en el manejo de la infección por VHC. La implantación de los antivirales de acción directa ha redundado en una mayor proporción de pacientes tratados en el grupo de generalización (52,3 % vs. 39,2 %) y, sobre todo, una mayor proporción de pacientes con respuesta viral sostenida (43,8 % vs. 8,8 %). Es conocido que la generalización del uso de AAD, también ha tenido un impacto considerable en la reducción de las listas de espera, y los pacientes de nuestro estudio no son ajenos a ello, con una reducción significativa del tiempo en lista en el grupo de generalización (63,5 días) frente al tiempo en lista de espera en el grupo de entrenamiento (147 días).

La implantación del uso de injertos procedentes de donantes en asistolia controlada también se objetiva en los datos ya que la

proporción de estos donantes alcanza casi el 20 % en el grupo de generalización, frente a 0 donantes en asistolia en el grupo de entrenamiento ($p < 0,001$).

Es evidente que todas estas diferencias se deben a los cambios que ha experimentado la práctica clínica en el manejo del CHC a lo largo de los 15 años que abarca el período de estudio (2004 – 2018) y que estas diferencias suponen un cierto sesgo de selección, que podría haberse evitado mediante la distribución aleatoria de los pacientes en los grupos de estudio. Sin embargo, nos decantamos por una división temporal ya que el grupo de generalización permite una validación interna del modelo, al utilizar una proporción de casos adecuada (25 %) y ser pacientes trasplantados con posterioridad a los del grupo de entrenamiento, cumpliendo así uno de los objetivos secundarios del estudio.

Por otra parte, el modelo de RB obtiene una ABC de 0,889 para el grupo de generalización, por lo que su capacidad predictiva puede considerarse como buena, justificando además su aplicación en la práctica clínica. Es importante destacar que el modelo se aplica a los casos del grupo de generalización sin que los “conozca” previamente, por lo que conseguir una ABC tan elevada permite deducir que se trata de un modelo sólido y confiable. Es decir, su capacidad predictiva no sólo no disminuye, sino que se incrementa (ABC entrenamiento = 0,817) a pesar de evaluarse utilizando casos con hasta 15 años de diferencia, por lo que resiste los posibles cambios que se hayan producido en la práctica clínica.

Por su parte, el modelo de RL también consigue una ABC buena (0,842) aunque inferior a la conseguida por el modelo de RB, si bien las diferencias entre ambos grupos no alcanzaron significación estadística ($p = 0,4$), probablemente debido a que el número de pacientes ($n = 130$) es insuficiente para conseguir diferencias significativas.

6.5.- Comparación con otros modelos predictivos de recurrencia del carcinoma hepatocelular

El TH representa la mejor alternativa terapéutica con intención curativa para el CHC. Desde su publicación en 1996, los criterios de Milán siguen siendo el *gold standard* para la inclusión de pacientes en lista de espera de TH. Sin embargo, a lo largo de los últimos 30 años, la práctica clínica ha evolucionado y el conocimiento acerca del comportamiento del CHC se ha ampliado. En el campo de la *transplant oncology*, se han ido desarrollando multitud de nuevos criterios y sistemas pronósticos que tratan de ofrecer el TH a un mayor número de pacientes dado que obtienen resultados comparables en cuanto a recurrencia. Estos nuevos criterios expandidos se han creado bajo la premisa de que los criterios de Milán son excesivamente restrictivos, y además la morfología tumoral es insuficiente y poco exacta para predecir la recurrencia y la supervivencia libre de enfermedad.

Es en este marco donde se sitúa el presente estudio, en el que se ha desarrollado un modelo de inteligencia artificial basado en RB con el fin de predecir con exactitud la posibilidad de que se produzca una recurrencia precoz del CHC en los dos primeros años postrasplante. Desde este punto de vista, el modelo creado difiere de otros muchos publicados en la literatura, ya que no se centra en evaluar la supervivencia libre de enfermedad, sino que determina la posibilidad de una recurrencia precoz, con el fin de evitar que se realice un TH que se considera fútil, dado el impacto que la recidiva en los dos primeros años postrasplante tiene en la supervivencia de los pacientes

y los costes que de ello se derivan. Es importante destacar que, hasta el momento, no existen modelos o criterios publicados que se centren en este objetivo concreto, ya que la mayoría de ellos apuntan a la supervivencia libre de enfermedad o mortalidad relacionada con el CHC. Esto dificulta la comparación del modelo del estudio con otros disponibles.

El modelo desarrollado tiene una finalidad eminentemente práctica, que pretende optimizar un recurso escaso como son los injertos disponibles, a pesar de la situación de relativa bonanza que vivimos en España gracias al desarrollo y consolidación del uso de injertos de donantes en asistolia controlada y de otras estrategias para obtener donantes efectivos.

El objetivo principal es crear un modelo predictivo que se convierta en una herramienta fiable que ayude en la toma de decisiones sobre la inclusión del paciente en lista de espera, por lo que únicamente utiliza información de variables fácilmente disponibles en la práctica clínica antes del trasplante. Desde este punto de vista, el modelo sólo puede incluir una cantidad limitada de información, pero es capaz de caracterizar de forma adecuada la agresividad del tumor para cada caso concreto, permitiendo estimar la probabilidad de recurrencia precoz y orientar la decisión sobre realizar un TH o no. Es ahí donde consideramos que nuestro modelo (mediante *machine learning* e inteligencia artificial) es capaz de utilizar toda la información disponible previamente al trasplante y que la literatura ha demostrado como relevante en cuanto al comportamiento biológico del tumor.

En este sentido, el modelo creado supera la información de los tradicionales criterios morfológicos (3,7), ya que en principio no tiene en cuenta el tamaño tumoral como variable relevante. Esto puede deberse a que un porcentaje de los pacientes del estudio están trasplantados en base a los criterios morfológicos de Valencia (13). Sin embargo, el modelo incluye la respuesta a los tratamientos locorregionales pretrasplante y es capaz de determinar si un paciente se ha trasplantado fuera de criterios morfológicos, habiendo recibido tratamiento previo o no (*dowsntaging* inefectivo), por lo que complementa adecuadamente la información acerca de la morfología tumoral. También se puede considerar que, al no incluir el tamaño tumoral, el modelo se alinea con las ideas de ciertos centros que obvian la morfología tumoral a la hora de incluir a pacientes para TH, como los grupos de Toronto (15) o Padua (158), si bien es cierto que en estos casos se requiere una biopsia del tumor previa al trasplante. En el modelo creado, dicha información no está disponible, pero se incluye información que traduce en cierta manera la agresividad del tumor (*surrogates of tumor biology*), por lo que se puede considerar que se supera el tamaño tumoral como restricción al TH, siempre que la respuesta a terapias locorregionales sea la adecuada y haya unos niveles de AFP aceptables. Estas dos variables, fundamentales en el campo de la *transplant oncology* para predecir la recurrencia postrasplante, forman parte del modelo predictivo, de tal manera que lo convierten en una herramienta completa que permite comprender mejor el comportamiento del CHC.

En este sentido nuestro modelo se alinea con otros modelos o criterios combinados que tratan de aunar toda la información disponible en la selección de pacientes para TH, como el propio Metroticket 2.0 (75,76), el MORAL *score* (159), el LiTES-HCC (150) o los trabajos publicados por Q. Lai que incluyen la respuesta a los tratamientos locorregionales en los sistemas EurHeCaLT (153) y TRAIN (77). Como hemos comentado anteriormente, ninguno de estos trabajos tiene como objetivo determinar la recurrencia precoz del CHC, pero sí tratan de predecir la supervivencia libre de enfermedad, la recurrencia o la mortalidad relacionada con el CHC, y evalúan su capacidad predictiva con métodos similares a los empleados en nuestro estudio.

Así, por ejemplo, el MORAL *score* obtiene un *estadístico c* (equiparable al ABC de una curva ROC) de 0,870 para la predicción de la recurrencia postrasplante del CHC, resultados similares a los obtenidos por nuestro modelo de RB, tanto para el grupo de entrenamiento como de generalización. Por su parte, el Metroticket 2.0 también evalúa mediante el *estadístico c* su capacidad predictiva para la mortalidad relacionada con el CHC en un conjunto de entrenamiento y en un conjunto de validación interna (de manera análoga a nuestro estudio) con unos resultados de 0,780 para el grupo de entrenamiento y de 0,733 para el grupo de validación interna. El sistema EurHeCaLT analiza la supervivencia global y libre de enfermedad integrando la curva de AFP y la respuesta a los tratamientos locorregionales pretrasplante en base a los criterios mRECIST, con una ABC inferior (0,66) y unos valores de sensibilidad inferiores a los obtenidos por nuestro modelo de RB (que oscilan entre el 17 % y el 30 %) y unos

valores de especificidad similares a los de nuestro estudio (por encima del 90 %). El sistema Lites-HCC publicado por Goldberg et. al en 2021 analiza la supervivencia global, con una ABC de entre 0,59 y 0,63 para distintas supervivencias postrasplante (1, 3, 5 y 10 años postrasplante).

De todos los *scores* y criterios pronósticos publicados, quizá el TRAIN *score* es el que más se asemeja metodológicamente al presente estudio, puesto que cuenta con un conjunto de entrenamiento y otro de validación, y analiza mediante curvas ROC y sensibilidad/especificidad la capacidad predictiva para determinar mortalidad postrasplante relacionada con CHC y la invasión microvascular (como factor de riesgo de recurrencia postrasplante). Con un número de pacientes inferior al utilizado en este estudio (179 para el grupo de entrenamiento y 110 para el grupo de validación), el sistema TRAIN obtiene una ABC de 0,791 para la mortalidad relacionada con CHC y de 0,682 para la probabilidad de invasión microvascular, con unos valores de sensibilidad inferiores a los obtenidos por el modelo de RB (26,1 %) y unos valores de especificidad similares a los de nuestro estudio (91,8 %).

Asimismo, es importante destacar que muchos de los mencionados estudios comparan su capacidad predictiva frente a los criterios de Milán, calculando la cantidad de pacientes fuera de Milán que obtendrían buenos resultados de haberse trasplantado en base a los criterios propuestos para cada estudio. En este sentido, es de suponer que, al no tener en cuenta criterios morfológicos sino

respuesta al tratamiento locorregional, nuestro modelo predictivo permitiría igualmente rescatar a pacientes para trasplante inicialmente fuera de Milán, si bien esta hipótesis debe ser refrendada mediante la validación externa del modelo. Así, el desarrollo de una interfaz informática de acceso abierto permitirá que el modelo sea una herramienta accesible a cualquier investigador o grupo de trasplante que quiera comparar y validar resultados, dado que también puede utilizarse para realizar simulaciones de riesgo de recurrencia precoz postrasplante para casos concretos.

Por otra parte, los sistemas basados en inteligencia artificial y *machine learning* ofrecen potenciales beneficios en cuanto a capacidad predictiva, ya que superan a los sistemas basados en bioestadística tradicional y experiencia clínica. Además, las RB permiten la interacción con el conocimiento clínico adquirido, obteniendo resultados sinérgicos. No obstante, es importante destacar que cada modelo predictivo debe ser utilizado para el problema concreto para el que fue diseñado y que su capacidad de aplicación, precisión y exactitud se fundamenta en la homogeneidad de la base de datos a partir de la cual fue creado. Sistemas predictivos teóricamente eficaces pueden obtener resultados desiguales cuando se aplican a conjuntos de datos diferentes, con datos incompletos o inconsistentes, como ha demostrado el grupo de J. Briceño en sus recientes trabajos acerca del emparejamiento donante-receptor con distintos modelos de *machine learning* (148,160).

En este sentido, nuestro modelo se ha desarrollado sobre una base de datos con una proporción de datos perdidos inferior al 1 %, por lo que este problema queda resuelto.

Por tanto, a pesar de los inevitables sesgos que incluye este estudio, el modelo desarrollado mediante RB integra información completa (morfológica y biológica) acerca del comportamiento del CHC, con variables fácilmente disponibles previamente al TH, que permite realizar una estimación precisa del riesgo de recurrencia precoz del CHC. Esta capacidad predictiva ha sido objetivada mediante el ABC de la curva ROC y parámetros de sensibilidad y especificidad. En base a ello, los resultados obtenidos son superiores a los de otros modelos conocidos y aplicados en el campo de la *transplant oncology*, por lo que consideramos que se trata de una herramienta útil cuyo uso estaría justificado en la toma de decisiones en la práctica clínica habitual.

6.6.- Aplicación clínica del modelo de redes bayesianas. Predicción de la recurrencia precoz postrasplante del carcinoma hepatocelular

Los resultados de nuestro estudio muestran un modelo de RB capaz de predecir adecuadamente el riesgo de recidiva precoz postrasplante del CHC, con una ABC superior a 0,8 tanto para el grupo de entrenamiento como para el grupo de generalización. En base a ello, se plantea una aplicación clínica del modelo que oriente en la toma de decisiones clínicas sobre el TH con indicación oncológica para el CHC, con el fin de evitar procedimientos fútiles y optimizar el uso de injertos disponibles.

El modelo se implementa mediante una aplicación web de acceso abierto, en aquellos casos que cumplan los mismos criterios de inclusión con los que fue elaborado, es decir, pacientes mayores de 18 años afectos de carcinoma hepatocelular, potenciales receptores de un trasplante de hígado.

Recopilada la información y cumplimentados los campos de la aplicación web, esta ofrece un resultado teniendo en cuenta las utilidades asignadas al proceso de creación del modelo. Dicho resultado se compara con el umbral del modelo (0,229), lo que permite categorizar cada caso concreto en 2 posibles grupos:

- Cuando el valor obtenido es inferior al umbral, el modelo considera que dicho caso tiene un riesgo bajo de recurrencia en los dos primeros años después del TH, por lo que el trasplante puede

realizarse con garantías.

- Si el valor obtenido es superior al umbral considerado, el modelo estima que este caso tiene un riesgo elevado de recurrencia precoz postrasplante, por lo que sería conveniente considerar otras alternativas terapéuticas para evitar un procedimiento fútil.

Así, es posible plantear el uso del modelo predictivo de RB según un protocolo clínico:

1.- Evaluación de cada caso concreto por parte de de cada grupo de trasplante según su práctica clínica habitual y consideración como potencial receptor.

2.- Recopilación de los datos requeridos para cumplimentar el formulario de la aplicación web.

3.- Introducción de los datos y cumplimentación del formulario en la aplicación web de acceso abierto. Obtención del resultado predictivo de riesgo de recurrencia precoz postrasplante del CHC.

4.- Los resultados obtenidos con la aplicación del modelo ayudan a la toma de decisión sobre la conveniencia de realizar el trasplante, con una elevada capacidad predictiva, teniendo también en cuenta que la decisión es multifactorial.

Del mismo modo que el modelo puede aplicarse para la toma de decisiones clínicas concretas, gracias a su facilidad de acceso mediante la aplicación web, también puede emplearse para evaluar el riesgo de recurrencia en TH (con el fin de implementar un seguimiento

intensivo en casos concretos de riesgo elevado de recidiva precoz) o bien realizar simulaciones de riesgo para situaciones clínicas específicas.

Predicción de recurrencia precoz postrasplante del CHC

Estatus del Tratamiento

☒ No tratamiento

☐ Tratamiento puente

☐ Downstaging M

☐ Downstaging 7

☐ Fallo Downstaging

Respuesta al tratamiento

☒ No procede

☐ Completa

☐ Parcial

☐ Estable

☐ Progresión

Downstaging Milán

☒ No

☐ Sí

AFP máximo (ng/mL)

☒ AFP $\geq$ 1000

☐ 400 $\leq$ AFP<1000

☐ AFP<400

AFP pre-trasplante (ng/mL)

☒ AFP-pre-TxH $\leq$ 9

☐ AFP-pre-TxH>9

Riesgo BAJO de recurrencia del CHC en los dos primeros años después del tra

Con una probabilidad del 0.000361%

Se recomienda proceder con el Trasplante Hepático.

Información modelo Bayesiano de predicción:

Sensibilidad : 50.00%

Especificidad : 96.68%

Valor predictivo positivo : 55.56%

Valor predictivo negativo : 95.87%

Explicación de la predicción:

Break Down profile

caret

Figura 26: Formulario para la inclusión de datos en la aplicación web de acceso abierto: <http://oceano.uv.es:3838/RBRecidiva>

Por tanto, la utilización del modelo resulta intuitiva, sencilla y versátil, con distintos tipos de aplicaciones. Sin embargo, es necesario hacer una serie de consideraciones:

- El modelo predictivo constituye una herramienta útil para su aplicación en la práctica clínica dentro del campo de la *transplant oncology*, puesto que cumple las condiciones de incorporar información fiable y fácilmente accesible acerca de las características

morfológicas y biológicas del CHC, todo ello disponible previamente al TH.

- La aplicación web de acceso abierto permite la validación externa del modelo, con la inclusión de datos de pacientes procedentes de aquellos otros grupos de TH que estén abiertos a incluir pacientes.

- Nuestro modelo no pretende realizar una estimación de la supervivencia libre de enfermedad postrasplante, sino que se centra en determinar el riesgo de recurrencia precoz, evitando por tanto trasplantes fútiles y optimizando el uso de los injertos disponibles. Esto también permite ofrecer a los pacientes con riesgo elevado alternativas terapéuticas menos agresivas que permitan prolongar su supervivencia con una adecuada calidad de vida.

- El modelo se ha generado mediante una base retrospectiva de pacientes trasplantados en base a unos criterios morfológicos (Valencia) y analizados frente a los criterios de Milán y *Up-to-seven*. A pesar de que el modelo incluye información complementaria, es importante considerar los resultados dentro del marco de unos criterios morfológicos establecidos.

- El modelo no incluye información acerca de otros marcadores o variables que han demostrado en la literatura su influencia en la recurrencia y SLE del CHC después del TH. Sin embargo, al tratarse de un modelo basado en *machine learning*, su plasticidad intrínseca hace que pueda reevaluarse mediante la introducción o combinación de nuevas variables, que favorezcan un incremento en su capacidad

predictiva.

- Este modelo de nueva creación no pretende desestimar criterios de trasplante ya establecidos, sino mejorarlos e integrarlos dentro de una herramienta estadística más completa, que permita predecir mejor un hecho complejo como es el comportamiento del CHC, y optimizar el procedimiento del trasplante hepático.

Conclusiones

7.- Conclusiones

Según los resultados obtenidos y respondiendo a los objetivos planteados al inicio de este estudio, podemos establecer las siguientes conclusiones:

1.- A partir de un conjunto de 5 variables del receptor y tumor conocidas previamente al TH, de fácil acceso y recopiladas en un conjunto de 385 pacientes, se ha generado un modelo predictivo de recurrencia precoz de carcinoma hepatocelular postrasplante basado en redes bayesianas. Las variables que forman parte del modelo predictivo son:

- Respuesta a terapia pretrasplante
- Estado después del tratamiento pretrasplante
- *Downstaging* a criterios de Milán (sí/no)
- AFP máxima en seguimiento pretrasplante:
- Valor de la AFP pretrasplante (logaritmo en base 10), estableciendo un punto de corte en 9 ng/ml

Este modelo permite predecir el riesgo de recurrencia del CHC en los dos primeros años después del TH con resultados válidos (área bajo la curva de 0,817 para el grupo de entrenamiento).

- 2.- El modelo de redes bayesianas ha sido comparado con un modelo de regresión logística, que ha obtenido un área bajo la curva inferior (0,791 para el grupo de entrenamiento), si bien estas diferencias no alcanzaron significación estadística ($p = 0,2$).
- 3.- Estos resultados han sido validados internamente con un grupo de generalización formado por 130 pacientes trasplantados con posterioridad en el mismo centro, obteniendo mejores resultados en cuanto a capacidad predictiva (área bajo la curva del modelo de redes bayesianas = 0,889). Estos resultados son superiores a los obtenidos por el modelo de regresión logística para el grupo de generalización (0,843), aunque las diferencias no alcanzaron significación estadística ($p = 0,4$).
- 4.- Dado que los resultados de rendimiento del modelo de redes bayesianas para predecir la recurrencia precoz postrasplante del CHC avalan su posible uso en la práctica clínica como ayuda a la toma de decisiones, se ha elaborado un protocolo de aplicación que se llevará a cabo de forma prospectiva, en la Unidad de Cirugía Hepatobiliopancreática y Trasplante del Hospital Universitari i Politècnic La Fe de Valencia, mediante una aplicación web de acceso abierto.

Bibliografía

8.- Bibliografía

1. Starzl TE, Demetris AJ. Liver transplantation: A 31-year perspective part III. *Curr Probl Surg*. 1990 Apr 1;27(4):187–240.
2. Pichlmayr R, Weimann A, Oldhafer KJ, Schlitt HJ, Tusch G, Raab R. Appraisal of transplantation for malignant tumours of the liver with special reference to early stage hepatocellular carcinoma. *Eur J Surg Oncol*. 1998;24(1):60–7.
3. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, Montalto F, Ammatuna M, Morabito A, Gennari L. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med*. 1996 Mar 14;334(11):693-9.
4. Sapisochin G, Bruix J. Liver transplantation for hepatocellular carcinoma: Outcomes and novel surgical approaches. *Nat Rev Gastroenterol Hepatol*. 2017;14(4):203–17.
5. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018;391(10127):1301–14.
6. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med*. 2019 Apr 11;380(15):1450-1462.
7. Yao FY, Xiao L, Bass NM, Kerlan R, Ascher NL, Roberts JP. Liver transplantation for hepatocellular carcinoma: Validation of the UCSF-expanded criteria based on preoperative imaging. *Am J Transplant*. 2007;7(11):2587–96.

8. Decaens T, Roudot-Thoraval F, Hadni-Bresson S, Meyer C, Gugenheim J, Durand F, et al. Impact of UCSF criteria according to pre- and post-OLT tumor features: Analysis of 479 patients listed for HCC with a short waiting time. *Liver Transplant*. 2006;12(12):1761–9.
9. Leung JY, Zhu AX, Gordon FD, Pratt DS, Mithoefer A, Garrigan K, et al. Liver transplantation outcomes for early-stage hepatocellular carcinoma: Results of a multicenter study. *Liver Transplant*. 2004 Nov 1;10(11):1343–54.
10. Vibert E, Schwartz M, Olthoff KM. Advances in resection and transplantation for hepatocellular carcinoma. *J Hepatol*. 2020 Feb 1;72(2):262–76.
11. Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *Lancet Oncol*. 2009;10(1):35–43.
12. Herrero JI, Sangro B, Pardo F, Quiroga J, Iñarrairaegui M, Rotellar F, et al. Liver transplantation in patients with hepatocellular carcinoma across Milan criteria. *Liver Transplant*. 2008 Mar;14(3):272–8.
13. Silva M, Moya A, Berenguer M, Sanjuan F, López-Andujar R, Pareja E, et al. Expanded criteria for liver transplantation in patients with cirrhosis and hepatocellular carcinoma. *Liver Transplant*. 2008 Oct;14(10):1449–60.
14. Lee SG, Hwang S, Moon DB, Ahn CS, Kim KH, Sung KB, Ko GY, Park KM, Ha TY, Song GW. Expanded indication criteria of living donor liver transplantation for hepatocellular carcinoma at one large-volume center. *Liver Transpl*. 2008 Jul;14(7):935–45.

-
15. Dubay D, Sandroussi C, Sandhu L, Cleary S, Guba M, Cattral MS, et al. Liver transplantation for advanced hepatocellular carcinoma using poor tumor differentiation on biopsy as an exclusion criterion. *Ann Surg*. 2011 Jan;253(1):166–72.
 16. Ferrer-Fàbrega J, Forner A, Liccioni A, Miquel R, Molina V, Navasa M, et al. Prospective validation of *ab initio* liver transplantation in hepatocellular carcinoma upon detection of risk factors for recurrence after resection. *Hepatology*. 2016 Mar 1;63(3):839–49.
 17. Grupo de Trabajo de la SETH sobre ampliación de criterios de Trasplante Hepático. 8a Reunión de Consenso de la SET - Expansión de indicaciones en Trasplante Hepático. 2018.
 18. Vibert E, Azoulay D, Hoti E, Iacopinelli S, Samuel D, Salloum C, et al. Progression of alphafetoprotein before liver transplantation for hepatocellular carcinoma in cirrhotic patients: A critical factor. *Am J Transplant*. 2010;10(1):129–37.
 19. DuBay DA, Sandroussi C, Kachura JR, et al. Radiofrequency ablation of hepatocellular carcinoma as a bridge to liver transplantation. *HPB (Oxford)*. 2011;13(1):24-32.
 20. Yao FY, Mehta N, Flemming J, Dodge J, Hameed B, Fix O, et al. Downstaging of hepatocellular cancer before liver transplant: Long-term outcome compared to tumors within Milan criteria. *Hepatology*. 2015;61(6):1968–77.
 21. Cucchetti A, Zanello M, Cescon M, Ercolani G, Del Gaudio M, Ravaioli M, et al. Improved diagnostic imaging and interventional therapies prolong survival after resection for hepatocellular carcinoma in cirrhosis: The university of Bologna experience over 10 years. *Ann Surg*

- Oncol. 2011;18(6):1630–7.
22. Kim HD, Shim JH, Kim GA, Shin YM, Yu E, Lee SG, et al. Optimal methods for measuring eligibility for liver transplant in hepatocellular carcinoma patients undergoing transarterial chemoembolization. *J Hepatol.* 2015;62(5):1076–84.
 23. Finkenstedt A, Vikoler A, Portenkirchner M, Mülleder K, Maglione M, Margreiter C, et al. Excellent post-transplant survival in patients with intermediate stage hepatocellular carcinoma responding to neoadjuvant therapy. *Liver Int.* 2016;36(5):688–95.
 24. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, Kelley RK, Galle PR, Mazzaferro V, Salem R, Sangro B, Singal AG, Vogel A, Fuster J, Ayuso C, Bruix J. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol.* 2022 Mar;76(3):681-693.
 25. Lu DSK, Yu NC, Raman SS, Limanond P, Lassman C, Murray K, et al. Radiofrequency ablation of hepatocellular carcinoma: treatment success as defined by histologic examination of the explanted liver. *Radiology.* 2005 Mar;234(3):954–60.
 26. Mazzaferro V, Battiston C, Perrone S, Pulvirenti A, Regalia E, Romito R, et al. Radiofrequency ablation of small hepatocellular carcinoma in cirrhotic patients awaiting liver transplantation: A prospective study. *Ann Surg.* 2004;240(5):900–9.
 27. Pompili M, Mirante VG, Rondinara G, Fassati LR, Piscaglia F, Agnes S, et al. Percutaneous ablation procedures in cirrhotic patients with hepatocellular carcinoma submitted to liver transplantation: Assessment of efficacy at explant analysis and of safety for tumor

- recurrence. *Liver Transpl.* 2005 Sep;11(9):1117–26.
28. Zanús G, Boetto R, Gringeri E, Vitale A, D'Amico F, Carraro A, et al. Microwave thermal ablation for hepatocarcinoma: six liver transplantation cases. *Transplant Proc.* 2011 May;43(4):1091–4.
 29. Alba E, Valls C, Dominguez J, Martinez L, Escalante E, Lladó L, et al. Transcatheter Arterial Chemoembolization in Patients with Hepatocellular Carcinoma on the Waiting List for Orthotopic Liver Transplantation. *Am J Roentgenol.* 2008 May 23;190(5):1341–8.
 30. De Luna W, Sze DY, Ahmed A, Ha BY, Ayoub W, Keeffe EB, et al. Transarterial chemoinfusion for hepatocellular carcinoma as downstaging therapy and a bridge toward liver transplantation. *Am J Transplant.* 2009 May;9(5):1158–68.
 31. Kulik L, Heimbach JK, Zaiem F, Almasri J, Prokop LJ, Wang Z, et al. Therapies for patients with hepatocellular carcinoma awaiting liver transplantation: A systematic review and meta-analysis. *Hepatology.* 2018 Jan;67(1):381–400.
 32. Gabr A, Abouchaleh N, Ali R, Vouche M, Atassi R, Memon K, et al. Comparative study of post-transplant outcomes in hepatocellular carcinoma patients treated with chemoembolization or radioembolization. *Eur J Radiol.* 2017;93:100–6.
 33. Lewandowski RJ, Kulik LM, Riaz A, Senthilnathan S, Mulcahy MF, Ryu RK, et al. A comparative analysis of transarterial downstaging for hepatocellular carcinoma: Chemoembolization versus radioembolization. *Am J Transplant.* 2009;9(8):1920–8.
 34. Salem R, Johnson GE, Kim E, Riaz A, Bishay V, Boucher E, et al. Yttrium-90 Radioembolization for the Treatment of Solitary,

- Unresectable HCC: The LEGACY Study. *Hepatology*. 2021 Nov 1;74(5):2342–52.
35. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*. 2000 Feb 2;92(3):205-16.
 36. Forner A, Ayuso C, Varela M, Rimola J, Hessheimer AJ, De Lope CR, et al. Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: Are response evaluation criteria in solid tumors reliable? *Cancer*. 2009;115(3):616–23.
 37. Bruix J, Sherman M, Llovet JM, Beaugrand M, Lencioni R, Burroughs AK, et al. Clinical management of hepatocellular carcinoma. Conclusions of the barcelona-2000 EASL conference. *J Hepatol*. 2001;35(3):421–30.
 38. Lencioni R, Llovet JM. Modified RECIST (mRECIST) Assessment for Hepatocellular Carcinoma. *Semin Liver Dis*. 2010;30(212):52–60.
 39. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol*. 2018 Jul;69(1):182-236.
 40. Lencioni R. New data supporting modified RECIST (mRECIST) for hepatocellular carcinoma. *Clin Cancer Res*. 2013;19(6):1312–4.
 41. Clavien PA, Lesurtel M, Bossuyt PMM, Gores GJ, Langer B, Perrier A. Recommendations for liver transplantation for hepatocellular carcinoma: An international consensus conference report. *Lancet*

-
- Oncol. 2012;13(1):e11–22.
42. Zavaglia C, De Carlis L, Alberti AB, Minola E, Belli LS, Slim AO, et al. Predictors of long-term survival after liver transplantation for hepatocellular carcinoma. *Am J Gastroenterol*. 2005;100(12):2708–16.
 43. Fan J, Yang GS, Fu ZR, Peng ZH, Xia Q, Peng CH, et al. Liver transplantation outcomes in 1,078 hepatocellular carcinoma patients: A multi-center experience in Shanghai, China. *J Cancer Res Clin Oncol*. 2009;135(10):1403–12.
 44. Shah SA, Tan JCC, McGilvray ID, Cattral MS, Levy GA, Greig PD, et al. Does microvascular invasion affect outcomes after liver transplantation for HCC? A histopathological analysis of 155 consecutive explants. *J Gastrointest Surg*. 2007;11(4):464–71.
 45. Fuster J, García-Valdecasas JC, Grande L, Tabet J, Bruix J, Anglada T, et al. Hepatocellular carcinoma and cirrhosis: Results of surgical treatment in a European series. *Ann Surg*. 1996;223(3):297–302.
 46. Hameed B, Mehta N, Sapisochin G, Roberts JP, Yao FY. Alpha-fetoprotein level > 1000 ng/mL as an exclusion criterion for liver transplantation in patients with hepatocellular carcinoma meeting the Milan criteria. *Liver Transpl*. 2014 Aug;20(8):945-51.
 47. Duvoux C, Roudot-Thoraval F, Decaens T, et al. Liver transplantation for hepatocellular carcinoma: a model including α -fetoprotein improves the performance of Milan criteria. *Gastroenterology*. 2012;143(4):986-e15. doi:10.1053/j.gastro.2012.05.052
 48. Halazun KJ, Tabrizian P, Najjar M, Florman S, Schwartz M, Michelassi F, et al. Is it time to abandon the Milan criteria? Results of a bicoastal US collaboration to redefine hepatocellular carcinoma liver

- transplantation selection policies. *Ann Surg.* 2018;268(4):690–9.
49. Agopian VG, Harlander-Locke MP, Ruiz RM, Klintmalm GB, Senguttuvan S, Florman SS, et al. Impact of Pretransplant Bridging Locoregional Therapy for Patients with Hepatocellular Carcinoma Within Milan Criteria Undergoing Liver Transplantation: Analysis of 3601 Patients from the US Multicenter HCC Transplant Consortium. *Ann Surg.* 2017 Sep 1;266(3):525–35.
 50. Gordic S, Corcuera-Solano I, Stueck A, Besa C, Argiriadi P, Guniganti P, et al. Evaluation of HCC response to locoregional therapy: Validation of MRI-based response criteria versus explant pathology. *J Hepatol.* 2017 Dec 1;67(6):1213–21.
 51. Otto G, Herber S, Heise M, Lohse AW, Mönch C, Bittinger F, et al. Response to transarterial chemoembolization as a biological selection criterion for liver transplantation in hepatocellular carcinoma. *Liver Transplant.* 2006 Aug;12(8):1260–7.
 52. Millonig G, Graziadei IW, Freund MC, Jaschke W, Stadlmann S, Ladurner R, et al. Response to preoperative chemoembolization correlates with outcome after liver transplantation in patients with hepatocellular carcinoma. *Liver Transpl.* 2007 Feb;13(2):272–9.
 53. Cucchetti A, Cescon M, Bigonzi E, Piscaglia F, Golfieri R, Ercolani G, et al. Priority of candidates with hepatocellular carcinoma awaiting liver transplantation can be reduced after successful bridge therapy. *Liver Transplant.* 2011 Nov 1;17(11):1344–54.
 54. Lai Q, Vitale A, Iesari S, Finkenstedt A, Mennini G, Onali S, et al. The Intention-to-Treat Effect of Bridging Treatments in the Setting of Milan Criteria—In Patients Waiting for Liver Transplantation. *Liver Transplant.*

- 2019 Jul 1;25(7):1023–33.
55. DiNorcia J, Florman SS, Haydel B, Tabrizian P, Ruiz RM, Klintmalm GB, et al. Pathologic Response to Pretransplant Locoregional Therapy is Predictive of Patient Outcome After Liver Transplantation for Hepatocellular Carcinoma. *Ann Surg.* 2020 Apr;271(4):616–24.
 56. Chan K-M, Yu M-C, Chou H-S, Wu T-J, Lee C-F, Lee W-C. Significance of Tumor Necrosis for Outcome of Patients with Hepatocellular Carcinoma Receiving Locoregional Therapy Prior to Liver Transplantation. *Ann Surg Oncol.* 2011 Sep 17;18(9):2638–46.
 57. Agopian VG, Harlander-Locke M, Zarrinpar A, Kaldas FM, Farmer DG, Yersiz H, et al. A novel prognostic nomogram accurately predicts hepatocellular carcinoma recurrence after liver transplantation: analysis of 865 consecutive liver transplant recipients. *J Am Coll Surg.* 2015 Apr;220(4):416–27.
 58. Pompili M, Francica G, Ponziani FR, Iezzi R, Avolio AW. Bridging and downstaging treatments for hepatocellular carcinoma in patients on the waiting list for liver transplantation. *World J Gastroenterol.* 2013 Nov 21;19(43):7515–30.
 59. Groeschl RT, Hong JC, Christians KK, Turaga KK, Tsai S, Pilgrim CHC, et al. Viral status at the time of liver transplantation for hepatocellular carcinoma: a modern predictor of longterm survival. *HPB (Oxford).* 2013 Oct;15(10):794–802.
 60. Moya A, Berenguer M, Aguilera V, Juan FS, Nicolás D, Pastor M, et al. Hepatocellular carcinoma: Can it be considered a controversial indication for liver transplantation in centers with high rates of hepatitis C? *Liver Transplant.* 2002 Nov;8(11):1020–7.

61. Chen T, Terrault NA. Perspectives on treating hepatitis C infection in the liver transplantation setting. *Curr Opin Organ Transplant*. 2016 Apr;21(2):111–9.
62. Reig M, Mariño Z, Perelló C, Iñarrairaegui M, Ribeiro A, Lens S, et al. Unexpected high rate of early tumor recurrence in patients with HCV-related HCC undergoing interferon-free therapy. *J Hepatol*. 2016 Oct;65(4):719–26.
63. Yang JD, Aql BA, Pungpapong S, Gores GJ, Roberts LR, Leise MD. Direct acting antiviral therapy and tumor recurrence after liver transplantation for hepatitis C-associated hepatocellular carcinoma. *J Hepatol*. 2016 Oct;65(4):859–60.
64. Conti F, Buonfiglioli F, Scuteri A, Crespi C, Bolondi L, Caraceni P, et al. Early occurrence and recurrence of hepatocellular carcinoma in HCV-related cirrhosis treated with direct-acting antivirals. *J Hepatol*. 2016 Oct;65(4):727–33.
65. ANRS collaborative study group on hepatocellular carcinoma. Lack of evidence of an effect of direct-acting antivirals on the recurrence of hepatocellular carcinoma: Data from three ANRS cohorts. *J Hepatol*. 2016;65(4):734–740.
66. Miltiadows O, Sia D, Hoshida Y, Fiel MI, Harrington AN, Thung SN, et al. Progenitor cell markers predict outcome of patients with hepatocellular carcinoma beyond Milan criteria undergoing liver transplantation. *J Hepatol*. 2015 Dec;63(6):1368–77.
67. Qin L-X, Tang Z-Y, Ma Z-C, Wu Z-Q, Zhou X-D, Ye Q-H, et al. P53 immunohistochemical scoring: an independent prognostic marker for patients after hepatocellular carcinoma resection. *World J*

- Gastroenterol. 2002 Jun;8(3):459–63.
68. Fujiki M, Takada Y, Ogura Y, Oike F, Kaido T, Teramukai S, et al. Significance of Des-Gamma-Carboxy Prothrombin in Selection Criteria for Living Donor Liver Transplantation for Hepatocellular Carcinoma. *Am J Transplant.* 2009 Oct 1;9(10):2362–71.
 69. Takada Y, Ito T, Ueda M, Sakamoto S, Haga H, Maetani Y, et al. Living Donor Liver Transplantation for Patients with HCC Exceeding the Milan Criteria: A Proposal of Expanded Criteria. *Dig Dis.* 2007 Oct;25(4):299–302.
 70. Halazun KJ, Hardy MA, Rana AA, Woodland DC, Luyten EJ, Mahadev S, et al. Negative Impact of Neutrophil-Lymphocyte Ratio on Outcome After Liver Transplantation for Hepatocellular Carcinoma. *Ann Surg.* 2009 Jul;250(1):141–51.
 71. Xia W, Ke Q, Wang Y, Wang W, Zhang M, Shen Y, et al. Predictive value of pre-transplant platelet to lymphocyte ratio for hepatocellular carcinoma recurrence after liver transplantation. *World J Surg Oncol.* 2015 Feb 18;13:60.
 72. Ohtani H. Focus on TILs: Prognostic significance of tumor infiltrating lymphocytes in human colorectal cancer. *Cancer Immunol J Acad Cancer Immunol.* 2007 Feb 21;7:4.
 73. Mouchli M, Reddy S, Gerrard M, Boardman L, Rubio M. Usefulness of neutrophil-to-lymphocyte ratio (NLR) as a prognostic predictor after treatment of hepatocellular carcinoma. Review article. *Ann Hepatol.* 2021 May 1;22:100249.
 74. Notarpaolo A, Layese R, Magistri P, Gambato M, Colledan M, Magini G, et al. Validation of the AFP model as a predictor of HCC recurrence

- in patients with viral hepatitis-related cirrhosis who had received a liver transplant for HCC. *J Hepatol.* 2017;66:552–9.
75. Mazzaferro V, Sposito C, Zhou J, Pinna AD, De Carlis L, Fan J, Cescon M, Di Sandro S, Yi-Feng H, Lauterio A, Bongini M, Cucchetti A. Metroticket 2.0 Model for Analysis of Competing Risks of Death After Liver Transplantation for Hepatocellular Carcinoma. *Gastroenterology.* 2018 Jan;154(1):128-139.
76. Cucchetti A, Serenari M, Sposito C, Di Sandro S, Mosconi C, Vicentin I, et al. Including mRECIST in the Metroticket 2.0 criteria improves prediction of hepatocellular carcinoma-related death after liver transplant. *J Hepatol.* 2020;73(2):342–8.
77. Lai Q, Nicolini D, Inostroza M, Iesari S, Goffette P, Agostini A, et al. A novel prognostic index in patients with hepatocellular cancer waiting for liver transplantation Time-Radiological-response-Alpha-fetoprotein-INflammation (TRAIN) score. *Ann Surg.* 2016;264(5):787–96.
78. Sasaki K, Firl DJ, Hashimoto K, Fujiki M, Diago-Uso T, Quintini C, Eghtesad B, Fung JJ, Aucejo FN, Miller CM. Development and validation of the HALT-HCC score to predict mortality in liver transplant recipients with hepatocellular carcinoma: a retrospective cohort analysis. *Lancet Gastroenterol Hepatol.* 2017 Aug;2(8):595-603.
79. Firl DJ, Kimura S, McVey J, Hashimoto K, Yeh H, Miller CM, Markmann JF, Sasaki K, Aucejo FN. Reframing the approach to patients with hepatocellular carcinoma: Longitudinal assessment with hazard associated with liver transplantation for HCC (HALTHCC) improves ablate and wait strategy. *Hepatology.* 2018 Oct;68(4):1448-1458.
80. Kaido T, Ogawa K, Mori A, Fujimoto Y, Ito T, Tomiyama K, et al.

- Usefulness of the Kyoto criteria as expanded selection criteria for liver transplantation for hepatocellular carcinoma. *Surgery*. 2013 Nov 1;154(5):1053–60.
81. Shapiro DE. The interpretation of diagnostic tests. *Stat Methods Med Res*. 1999 Apr 2;8(2):113–34.
82. Swets JA, Pickett RM, Whitehead SF, Getty DJ, Schnur JA, Swets JB, Freeman BA. Assessment of diagnostic technologies. *Science*. 1979 Aug 24;205(4408):753-9.
83. Burgueño MJ, García-Bastos JL, González-Buitrago JM. Las curvas ROC en la evaluación de las pruebas diagnósticas. *Med Clin (Barc)* 1995; 104: 661-670.
84. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988 Jun 3;240(4857):1285–93.
85. Cantor SB, Sun CC, Tortolero-Luna G, Richards-Kortum R, Follen M. A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. *J Clin Epidemiol*. 1999 Sep;52(9):885–92
86. Jordan MI. *Graphical Models*.. 2004 Feb 1;19(1):140–55.
87. Greenland S, Pearl J RJ. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37–48.
88. Knüppel S, Stang A. DAG program: identifying minimal sufficient adjustment sets. *Epidemiology*. 2010 Jan;21(1):159.
89. Breitling LP. dagR: a suite of R functions for directed acyclic graphs. *Epidemiology*. 2010 Jul;21(4):586–7.
90. Textor J, Hardt J, Knüppel S. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology*. 2011 Sep;22(5):745.

91. Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol*. 2016;45(6):1776–86.
92. Real Academia Española: Diccionario de la lengua española, 23^a ed., (versión 23.6 en línea).
93. Aceto G, Persico V, Pescapé A. The role of Information and Communication Technologies in healthcare: taxonomies, perspectives, and challenges. *J Netw Comput Appl*. 2018 Apr 1;107:125–54.
94. Hill J. Bayesian Analysis in Expert Systems: Comment: Graphical Models, Causality and Intervention. *Stat Sci*. 1993;8(3):266–9.
95. Lauritzen S. Propagation of probabilities, means and variances in mixed graphical models. *J Am Stat Assoc*. 1992;87:1098–108.
96. Dranca L, de Abetxuko Ruiz de Mendarozketa L, Goñi A et al. Using Kinect to classify Parkinson's disease stages related to severity of gait impairment. *BMC Bioinformatics*. 2018;19(1):1–15.
97. Philipp P, Beyerer J, Fischer Y, Beyerer J. Expert-based probabilistic modeling of workflows in context of surgical interventions. In: 2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA). IEEE; 2017. p. 1–7.
98. Julia Flores M, Nicholson AE, Brunskill A, Korb KB, Mascaro S. Incorporating expert knowledge when learning Bayesian network structure: a medical case study. *Artif Intell Med*. 2011 Nov;53(3):181–204.
99. Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Mach Learn*. 1992 Oct;9(4):309–47.

100. Cheng J, Bell DA, Liu W. Learning belief networks from data. *Artificial Intelligence*; 137: 43–90
101. Andreassen S, Benn JJ, Hovorka R, Olesen KG, Carson ER. A probabilistic approach to glucose prediction and insulin dose adjustment: description of metabolic model and pilot evaluation study. *Comput Methods Programs Biomed.* 1994 Jan;41(3–4):153–65.
102. Andreassen S, Riekehr C, Kristensen B, Schønheyder HC, Leibovici L. Using probabilistic and decision-theoretic methods in treatment and prognosis modeling. *Artif Intell Med.* 1999 Feb;15(2):121–34.
103. Lucas PFJ, Boot H TB. Computer-based decision-support in the management of primary gastric non- Hodgkin lymphoma. *Meth Inf Med.* 1998;37:206–19.
104. Korver M, Lucas PJ. Converting a rule-based expert system into a belief network. *Med Inform (Lond).* 1993;18(3):219–41.
105. van der Gaag, L.C., Helsper, E.M. (2002). Experiences with Modelling Issues in Building Probabilistic Networks. In: Gómez-Pérez, A., Benjamins, V.R. (eds) *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web. EKAW 2002. Lecture Notes in Computer Science*, vol 2473. Springer, Berlin, Heidelberg p. 21–6.
106. Silja Renooij and Linda C. van der Gaag. 2002. From qualitative to quantitative probabilistic networks. In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence (UAI'02)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 422–429
107. van der Gaag LC, Renooij S, Witteman CLM, Aleman BMP, Taal BG. Probabilities for a probabilistic network: a case study in oesophageal

- cancer. *Artif Intell Med*. 2002 Jun;25(2):123–48.
108. Veerle M. H. Coupé, Linda C. Van Der Gaag, and J. Dik F. Habbema. 2000. Sensitivity analysis: an aid for belief-network quantification. *Knowl. Eng. Rev.* 15, 3, 215–232
109. Eveline M. Helsper and Linda C. van der Gaag. 2002. Building Bayesian networks through ontologies. In *Proceedings of the 15th European Conference on Artificial Intelligence (ECAI'02)*. IOS Press, NLD, 680–684
110. Domingos, P.M., & Pazzani, M.J. (1997). On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. *Machine Learning*, 29, 103-130.
111. Cheng, J., & Greiner, R. (1999). Comparing Bayesian Network Classifiers.
112. Gamez JA, Gómez-Olmedo M, Nielsen TD. Bayesian network structure learning. *Proc First Eur Work Graph Model PGM'02*;117–26.
113. Heckerman DE, Horvitz EJ, Nathwani BN. Toward normative expert systems: Part I. The Pathfinder project. *Methods Inf Med*. 1992 Jun;31(2):90-105.
114. Heckerman DE, Nathwani BN. An evaluation of the diagnostic accuracy of Pathfinder. *Comput Biomed Res*. 1992 Feb;25(1):56–74.
115. Fernandes V, Junior GB, de Paiva AC, Silva AC, Gattass M. Bayesian convolutional neural network estimation for pediatric pneumonia detection and diagnosis. *Comput Methods Programs Biomed*. 2021 Sep 1;208:106259.
116. Briganti G, Williams DR, Mulder J, Linkowski P. Bayesian Network

- Structure and Predictability of Autistic Traits. *Psychol Rep.* 2022 Feb 1;125(1):344–57.
117. Vila-Francés J, Sanchís J, Soria-Olivas E, Serrano AJ, Martínez-Sober M, Bonanad C, et al. Expert system for predicting unstable angina based on Bayesian networks. *Expert Syst Appl.* 2013 Sep 15;40(12):5004–10.
118. da Cunha Leme DE. The use of Bayesian network models to identify factors related to frailty phenotype and health outcomes in middle-aged and older persons. *Arch Gerontol Geriatr.* 2021 Jan-Feb;92:104212.
119. Wang KM, Wang KJ, Makond B. Survivability modelling using Bayesian network for patients with first and secondary primary cancers. *Comput Methods Programs Biomed.* 2020 Nov;196:105686.
120. Zhang Z, Chai H, Wang Y, Pan Z, Yang Y. Cancer survival prognosis with Deep Bayesian Perturbation Cox Network. *Comput Biol Med.* 2022 Feb;141:105012.
121. Małczak P, Mizera M, Lee Y, Pisarska-Adamczyk M, Wysocki M, Bała MM, et al. Quality of Life After Bariatric Surgery-a Systematic Review with Bayesian Network Meta-analysis. *Obes Surg.* 2021 Dec 1;31(12):5213–23.
122. Cagigas Fernández C, Palazuelos C, Cristobal Poch L, Gomez Ruiz M. A probabilistic model for the prediction of intra-abdominal infection after colorectal surgery. *Int J Colorectal Dis.* 2021 Nov 1;36(11):2481–8.
123. Shachter RD. Evaluating Influence Diagrams. *Oper Res.* 1986 Dec 1;34(6):871–82.
124. Zhang S, Chang W, Xie J, Wu Z, Yang Y, Qiu H. The Efficacy, Safety,

- and Optimal Regimen of Corticosteroids in Sepsis: A Bayesian Network Meta-Analysis. *Crit Care Explor.* 2020 Apr [;2(4):e0094.
125. Khademi M, Nedialkov NS. Probabilistic Graphical Models and Deep Belief Networks for Prognosis of Breast Cancer. In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). IEEE; 2015. p. 727–32.
 126. Lee S, Jiang X. Modeling miRNA-mRNA interactions that cause phenotypic abnormality in breast cancer patients. *PLoS One.* 2017;12(8):e0182666.
 127. Liu E, Li J, Kinnebrew GH, Zhang P, Zhang Y, Cheng L, et al. A Fast and Furious Bayesian Network and Its Application of Identifying Colon Cancer to Liver Metastasis Gene Regulatory Networks. *IEEE/ACM Trans Comput Biol Bioinforma.* 2021 Jul 1;18(4):1325–35.
 128. McLachlan S, Dube K, Hitman GA, Fenton NE, Kyrimi E. Bayesian networks in healthcare: Distribution by medical condition. *Artif Intell Med.* 2020 Jul 1;107:101912.
 129. Menahem B, Lubrano J, Duvoux C, Mulliri A, Alves A, Costentin C, et al. Liver transplantation versus liver resection for hepatocellular carcinoma in intention to treat: An attempt to perform an ideal meta-analysis. *Liver Transplant.* 2017 Jun 1;23(6):836–44.
 130. Mehta N, Heimbach J, Harnois DM, Sapisochin G, Dodge JL, Lee D, et al. Validation of a Risk Estimation of Tumor Recurrence After Transplant (RETREAT) score for hepatocellular carcinoma recurrence after liver transplant. *JAMA Oncol.* 2017 Apr 1;3(4):493–500.
 131. Huang PL. A comprehensive definition for metabolic syndrome. *Dis Model Mech.* 2009;2(5–6):231.

-
132. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982 Apr;143(1):29-36.
 133. Akinyemiju T, Abera S, Ahmed M, Alam N, Alemayohu MA, Allen C, et al. The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level. *JAMA Oncol*. 2017 Dec 1;3(12):1683.
 134. Sleeman KE, de Brito M, Etkind S, Nkhoma K, Guo P, Higginson IJ, Gomes B, Harding R. The escalating global burden of serious health-related suffering: projections to 2060 by world regions, age groups, and health conditions. *Lancet Glob Health*. 2019 Jul;7(7):e883-e892.
 135. Yao FY, Ferrell L, Bass NM, Watson JJ, Bacchetti P, Venook A, et al. Liver transplantation for hepatocellular carcinoma: Expansion of the tumor size limits does not adversely impact survival. *Hepatology*. 2001;33(6):1394–403.
 136. Morris PD, Laurence JM, Yeo D, Crawford M, Strasser SI, McCaughan GW, et al. Can response to locoregional therapy help predict longterm survival after liver transplantation for hepatocellular carcinoma? A systematic review. *Liver Transplant*. 2017 Mar 1;23(3):375–85.
 137. Sheth RA, Patel MS, Koottappillil B, Shah JA, Oklu R, Mueller P, et al. Role of Locoregional Therapy and Predictors for Dropout in Patients with Hepatocellular Carcinoma Listed for Liver Transplantation. *J Vasc Interv Radiol*. 2015 Dec 1;26(12):1761–8.
 138. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, Kim TY, Choo SP, Trojan J, Welling TH Rd, Meyer T, Kang YK, Yeo W, Chopra A, Anderson J, De la Cruz C, Lang L, Neely J, Tang H, Dastani

- HB, Melero I. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet*. 2017 Jun 24;389(10088):2492-2502.
139. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, de Oliveira AC, Santoro A, Raoul JL, Forner A, Schwartz M, Porta C, Zeuzem S, Bolondi L, Greten TF, Galle PR, Seitz JF, Borbath I, Häussinger D, Giannaris T, Shan M, Moscovici M, Voliotis D, Bruix J; SHARP Investigators Study Group. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*. 2008 Jul 24;359(4):378-90.
140. VVAA. Memoria General de Resultados 2021. Registro Español de Trasplante Hepático. Disponible en: <http://www.ont.es>
141. Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *Lancet Oncol*. 2009 Jan [;10(1):35–43.
142. Chok KS, Chan SC, Cheung TT, Chan AC, Fan ST, Lo CM. Late recurrence of hepatocellular carcinoma after liver transplantation. *World J Surg*. 2011 Sep;35(9):2058-62.
143. Agopian VG, Morshedi MM, McWilliams J, Harlander-Locke MP, Markovic D, Zarrinpar A, et al. Complete pathologic response to pretransplant locoregional therapy for hepatocellular carcinoma defines cancer cure after liver transplantation: Analysis of 501 consecutively treated patients. *Ann Surg*. 2015 Sep 1;262(3):536–43.
144. Lee DD, Samoylova M, Mehta N, Musto KR, Roberts JP, Yao FY, Harnois DM. The mRECIST Classification Provides Insight into Tumor

- Biology for Patients With Hepatocellular Carcinoma Awaiting Liver Transplantation. *Liver Transpl.* 2019 Feb;25(2):228-241.
145. Halazun KJ, Rosenblatt RE, Mehta N, Lai Q, Hajifathalian K, Gorgen A, et al. Dynamic α -Fetoprotein Response and Outcomes after Liver Transplant for Hepatocellular Carcinoma. *JAMA Surg.* 2021 Jun 1;156(6):559–67.
 146. Cruz-Ramírez M, Hervás-Martínez C, Fernández JC, Briceño J, de la Mata M. Predicting patient survival after liver transplantation using evolutionary multi-objective artificial neural networks. *Artif Intell Med.* 2013 May;58(1):37–49.
 147. Briceño J, Ayllón MD, Ciria R. Machine-learning algorithms for predicting results in liver transplantation: the problem of donor-recipient matching. *Curr Opin Organ Transplant* [Internet]. 2020 Aug 1;25(4):406–11.
 148. Guijo-Rubio D, Briceño J, Gutiérrez PA, Ayllón MD, Ciria R, Hervás-Martínez C. Statistical methods versus machine learning techniques for donor-recipient matching in liver transplantation. *PLoS One.* 2021 May 1;16(5).
 149. Ivanics T, Patel MS, Erdman L, Sapisochin G. Artificial intelligence in transplantation (machine-learning classifiers and transplant oncology). *Curr Opin Organ Transplant.* 2020 Aug 1;25(4):426–34.
 150. Goldberg D, Mantero A, Newcomb C, Delgado C, Forde KA, Kaplan DE, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma using the LiTES-HCC score. *J Hepatol.* 2021 Jun 1;74(6):1398–406.
 151. Kornberg A, Witt U, Matevossian E, Küpper B, Assfalg V, Drzezga A,

- Hüser N, Wildgruber M, Friess H. Extended postinterventional tumor necrosis-implication for outcome in liver transplant patients with advanced HCC. *PLoS One*. 2013;8(1):e53960.
152. Vibert E, Azoulay D, Hoti E, Iacopinelli S, Samuel D, Salloum C, et al. Progression of alphafetoprotein before liver transplantation for hepatocellular carcinoma in cirrhotic patients: a critical factor. *Am J Transplant*. 2010 Jan;10(1):129–37.
153. Lai Q, Avolio AW, Graziadei I, Otto G, Rossi M, Tisone G, et al. Alpha-fetoprotein and modified response evaluation criteria in solid tumors progression after locoregional therapy as predictors of hepatocellular cancer recurrence and death after transplantation. *Liver Transpl*. 2013 Oct;19(10):1108–18.
154. Choi HJ, Kim DG, Na GH, Hong TH, You YK. Extended criteria for living donor liver transplantation in patients with advanced hepatocellular carcinoma. *Transplant Proc*. 2012 Mar;44(2):399–402.
155. Joong Choi H, Goo Kim D, Hyung Na G, Hyun Han J, Ho Hong T, Kyoung You Ho Joong Choi Y, et al. Clinical outcome in patients with hepatocellular carcinoma after living-donor liver transplantation. *World J Gastroenterol*. 2013;19(29):4737–44.
156. Wan P, Xia Q, Zhang J-J, Li Q-G, Xu N, Zhang M, et al. Liver transplantation for hepatocellular carcinoma exceeding the Milan criteria: a single-center experience. *J Cancer Res Clin Oncol*. 2014;140:341–8.
157. Taketomi A, Sanefuji K, Soejima Y, Yoshizumi T, Uhciyama H, Ikegami T, et al. Impact of des-gamma-carboxy prothrombin and tumor size on the recurrence of hepatocellular carcinoma after living donor liver

- transplantation. Transplantation. 2009 Feb 27;87(4):531–7.
158. Cillo U, Vitale A, Bassanello M, Boccagni P, Brolese A, Zanusi G, Burra P, Fagiuoli S, Farinati F, Rugge M, D'Amico DF. Liver transplantation for the treatment of moderately or well-differentiated hepatocellular carcinoma. Ann Surg. 2004 Feb;239(2):150-9.
159. Halazun KJ, Najjar M, Abdelmessih RM, Samstein B, Griesemer AD, Guarrera J V., et al. Recurrence after liver transplantation for hepatocellular carcinoma. Ann Surg. 2017 Mar 1;265(3):557–64.
160. Briceño J, Calleja R, Hervás C. Artificial intelligence and liver transplantation: Looking for the best donor-recipient pairing. Hepatobiliary Pancreat Dis Int. 2022 Aug;21(4):347-353.

