





VNIVERSITAT  
E VALÈNCIA (ò 人)  
Facultat de Medicina i Odontologia

Programa de Doctorado de Medicina 3042

**ANÁLISIS DE LA PROGRESIÓN DE LA  
ENFERMEDAD RENAL CRÓNICA AVANZADA  
MEDIANTE TÉCNICAS DE APRENDIZAJE MÁQUINA**

TESIS DOCTORAL

Presentada por:  
Amparo Soldevila Orient

Directores:  
Dr. Julio Hernández Jaras  
Dr. Marcelino Martínez Sober  
Dr. Alfonso Miguel Carrasco



DEPARTAMENTO DE MEDICINA

**D. JULIO HERNANDEZ JARAS.** Doctor en Medicina. Jefe del Area Clínica de Riñón y Vías Urinarias Hospital Politècnic i Universitari la Fe de València.

**D. MARCELINO MARTINEZ SOBER.** Doctor en Física. Profesor Titular de Universidad del Departamento de Ingeniería Electrónica. Escuela Técnica Superior de Ingeniería. Universitat de València.

**D. ALFONSO MIGUEL CARRASCO.** Doctor en Medicina. Profesor Titular de Universidad del Departamento de Medicina. Facultad de Medicina. Universitat de València.

Informan:

Que el trabajo titulado "ANALISIS DE LA PROGRESIÓN DE LA ENFERMEDAD RENAL CRÓNICA AVANZADA MEDIANTE TÉCNICAS DE APRENDIZAJE MÁQUINA", presentado por Dña. Amparo Soldevila Orient para optar al grado de Doctor por la Universitat de Valencia, ha sido realizado bajo nuestra dirección y reúne los requisitos necesarios para su presentación y lectura.

Y para que conste firmamos la presente en Valencia a 20 de Julio de 2017



Fdo: D. J. Hernández Jaras



Fdo: D. M. Martínez Sober



Fdo: D. A. Miguel Carrasco





## **AGRADECIMIENTOS**

Quiero expresar mi agradecimiento a todas las personas que han participado en la realización de esta tesis.

En primer lugar a mis directores Julio, Alfonso y Marcelino por su inestimable colaboración y la ilusión transmitida por este trabajo.

A Emilio y Marcelino por ser los promotores de este proyecto y por haber conseguido introducirme en el mundo de la estadística y de las redes neuronales. Gracias además por su infinita paciencia.

A mi familia y amigos por su apoyo incondicional y por el tiempo robado en muchas ocasiones.

A mis compañeros del Hospital La Fe, en especial a los que junto conmigo forman parte de la Unidad de ERCA, en la que intentamos día a día mejorar la vida de nuestros enfermos.

Y, por último, a los pacientes que sufren esta compleja enfermedad, ellos son el verdadero objeto de nuestro esfuerzo y a ellos debemos agradecerles nuestro continuo aprendizaje.



*A Pedro y a nuestros hijos  
Nuria, Pedro y María*



## ÍNDICE

---



# ÍNDICE

1. INTRODUCCIÓN .....	3
1.1. JUSTIFICACIÓN DEL TRABAJO .....	3
1.2. OBJETIVO .....	5
1.3. HIPÓTESIS .....	5
1.4. FILTRADO GLOMERULAR .....	6
1.5. ENFERMEDAD RENAL CRÓNICA .....	7
1.5.1. DEFINICIÓN .....	7
1.5.2. CLASIFICACIÓN .....	7
1.5.3. MODELO CONCEPTUAL PARA LA ENFERMEDAD RENAL CRÓNICA .....	10
1.6. PROGRESIÓN DE ENFERMEDAD RENAL .....	12
1.6.1. DEFINICIÓN .....	12
1.6.2. FACTORES DE PROGRESIÓN .....	13
1.6.2.1. <i>Concepto</i> .....	13
1.6.2.2. <i>Importancia</i> .....	14
1.6.2.3. <i>Descripción de los diferentes factores de progresión</i> .....	14
a) <i>Hipertensión arterial</i> .....	14
b) <i>Proteinuria</i> .....	15
c) <i>Raza</i> .....	16
d) <i>Diabetes</i> .....	17
e) <i>Dislipemia</i> .....	19
f) <i>Obesidad</i> .....	20



g) Hiperfosforemia .....	22
h) Tabaquismo .....	23
i) Hiperuricemia .....	24
j) Anemia .....	26
k) Acidosis metabólica .....	27
l) Inhibición del sistema renina-angiotensina-aldosterona ..	28
1.6.2.4. Predicción de la progresión en la enfermedad renal crónica .....	29
1.7. APRENDIZAJE MÁQUINA .....	32
2. METODOLOGÍA .....	37
2.1. DISEÑO DEL ESTUDIO .....	37
2.2. ANÁLISIS DE LA BASE DE DATOS .....	37
2.2.1. SELECCIÓN DE LA MUESTRA .....	37
2.2.2. VARIABLES ANALIZADAS .....	38
2.2.2.1. Características epidemiológicas .....	38
a) Raza .....	38
b) Sexo .....	38
c) Edad .....	38
2.2.2.2. Variables antropométricas .....	39
a) Peso.....	39
b) Talla .....	39
c) Índice de masa corporal .....	39
2.2.2.3. Enfermedad de base .....	39
2.2.2.4. Comorbilidad .....	40

2.2.2.5. Hipertensión arterial .....	40
2.2.2.6. Anemia .....	41
2.2.2.7. Hipertrofia ventricular izquierda .....	42
2.2.2.8. Insuficiencia cardíaca .....	42
2.2.2.9. Hábito tabáquico .....	42
2.2.2.10. Dislipemia .....	43
2.2.2.11. Diabetes mellitus .....	43
2.2.2.12. Alteraciones del metabolismo óseo-mineral .....	44
2.2.2.13. Acidosis metabólica .....	45
2.2.2.14. Tratamiento diurético .....	45
2.2.2.15. Tratamiento con IECAs/ARAII .....	45
2.2.2.16. Función renal .....	46
2.2.2.17. Hiperuricemia .....	47
2.2.2.18. Proteinuria .....	47
2.2.2.19. Albúmina .....	47
2.2.2.20. Sodio .....	48
2.2.2.21. Potasio .....	48
2.2.2.22. Cloro .....	48
2.2.2.23. Proteína C Reactiva (PCR) .....	48
2.3. ANÁLISIS ESTADÍSTICO .....	49
2.3.1. ETAPAS DEL ANÁLISIS .....	49
2.3.2. ANÁLISIS EXPLORATORIO DE DATOS .....	51
2.3.3. CONTRASTE DE HIPÓTESIS .....	52

2.3.4. MAPAS AUTOORGANIZADOS .....	54
2.3.4.1. Introducción .....	54
2.3.4.2. <i>Ejemplo</i> .....	55
2.3.5. SELECCIÓN DE VARIABLES. BORUTA .....	57
2.3.6. MODELOS MULTIVARIANTES .....	60
2.3.6.1. <i>Regresión logística</i> .....	60
2.3.6.2. <i>Naïve Bayes</i> .....	61
2.3.6.3. <i>TAN (Tree Augmented Network)</i> .....	62
2.3.7. MEDIDAS DE RENDIMIENTO DE LOS MODELOS .....	63
2.3.7.1. <i>Matriz de confusión</i> .....	63
2.3.7.2. <i>Área bajo la curva ROC (AUC)</i> .....	65
2.3.8. ÁRBOLES DE DECISIÓN .....	66
3. RESULTADOS Y DESARROLLO ARGUMENTAL .....	71
3.1. RESULTADOS .....	71
3.1.1. DESCRIPCIÓN DE LA MUESTRA .....	71
3.1.2. VARIABLES ANALIZADAS .....	74
3.1.2.1. <i>Enfermedad de base</i> .....	74
3.1.2.2. <i>Función renal</i> .....	75
3.1.2.3. <i>Comorbilidad</i> .....	78
3.1.2.4. <i>Hipertensión arterial</i> .....	79
3.1.2.5. <i>Tabaco</i> .....	80
3.1.2.6. <i>Obesidad</i> .....	80
3.1.2.7. <i>Anemia</i> .....	81
3.1.2.8. <i>Dislipemia</i> .....	85

3.1.2.9. <i>Diabetes</i> .....	86
3.1.2.10. <i>Acidosis</i> .....	88
3.1.2.11. <i>Metabolismo óseo-mineral</i> .....	90
3.1.2.12. <i>Hiperuricemia</i> .....	95
3.1.2.13. <i>Proteinuria</i> .....	96
3.1.2.14. <i>Albúmina en sangre</i> .....	97
3.1.2.15. <i>PCR</i> .....	97
3.1.2.16. <i>Diuréticos</i> .....	98
3.1.2.17. <i>Inhibidores del sistema renina-angiotensina-aldosterona</i> .....	99
3.1.2.18. <i>Hipertrofia ventricular izquierda</i> .....	100
3.1.2.19. <i>Insuficiencia cardíaca congestiva</i> .....	101
3.1.2.20. <i>Tratamiento antiagregante/anticoagulante</i> .....	102
3.1.3. MEDICIÓN DE LA PROGRESIÓN DE LA ENFERMEDAD RENAL .....	103
3.1.4. CORRELACIÓN DE LAS VARIABLES CON LA PROGRESIÓN DE LA ENFERMEDAD RENAL .....	104
3.1.5. CLASES DE PROGRESIÓN DE LA ENFERMEDAD RENAL .....	107
3.1.6. ANÁLISIS COMPARATIVO DE LOS DOS GRUPOS DE PROGRESIÓN .....	108
3.1.7. PROCEDIMIENTO EXPERIMENTAL EN EL AJUSTE DE MODELOS .....	111
3.1.8. MODELOS LINEALES UNIVARIANTES .....	112

3.1.9. MODELOS MULTIVARIANTES .....	120
3.1.9.1. <i>Regresión logística</i> .....	120
3.1.9.2. <i>Naïve Bayes</i> .....	122
3.1.9.3. <i>TAN (Tree augmented network)</i> .....	123
3.1.9.4. <i>Comparación de modelos</i> .....	125
3.1.10. SELECCIÓN DE VARIABLES .....	127
3.1.10.1. <i>Regresión logística</i> .....	127
3.1.10.2. <i>Naïve Bayes</i> .....	130
3.1.10.3. <i>TAN (Tree Augmented Network)</i> .....	132
3.1.10.4. <i>Comparación de modelos</i> .....	134
3.1.11. SELECCIÓN DE VARIABLES SIGNIFICATIVAS CON BORUTA	136
3.1.11.1. <i>Regresión logística</i> .....	136
3.1.11.2. <i>Naïve Bayes</i> .....	138
3.1.11.3. <i>TAN (Tree Augmented Network)</i> .....	139
3.1.12. ÁRBOL DE DECISIÓN .....	141
3.1.13. MAPAS AUTOORGANIZADOS .....	144
3.1.13.1. <i>Análisis global</i> .....	144
3.1.13.2. <i>Lípidos y progresión de enfermedad renal</i> .....	152
3.1.13.3. <i>Anemia y progresión de enfermedad renal</i> .....	154
3.1.13.4. <i>Hipertensión arterial y progresión de enfermedad renal</i>	156
3.1.13.5. <i>Equilibrio acido-base y progresión de enfermedad renal</i>	158

3.1.13.6. <i>Metabolismo óseo-mineral</i>	
<i>y progresión de enfermedad renal</i> .....	160
3.2. DESARROLLO ARGUMENTAL .....	162
4. CONCLUSIONES .....	181
5. BIBLIOGRAFÍA .....	185
6. ANEXOS .....	203

## 1. INTRODUCCIÓN

---





## 1. INTRODUCCIÓN

### 1.1. JUSTIFICACION DEL TRABAJO

La enfermedad renal crónica (ERC) constituye actualmente un importante problema de salud. Hoy en día tenemos una prevalencia de insuficiencia renal de un 10.6 % a nivel mundial y de un 11.86 % en Europa (Hill 2016). Según el estudio *EPIRCE* (Otero 2010) realizado en la población española, la prevalencia de esta enfermedad en las categorías 3-5 fue del 6.8 % y, más recientemente, se reporta un incremento de la misma de hasta un 8.7 % (Vargas 2015). Esto conlleva implicaciones socio-sanitarias y económicas, puesto que, por una parte, la enfermedad renal crónica se asocia a una importante comorbilidad, y, por otra, el coste económico que genera es muy elevado, suponiendo en la actualidad un 2.5 - 3% del presupuesto del Sistema Nacional de Salud y más del 4 % del de la atención especializada (Vargas 2015).

Teniendo en cuenta estos aspectos, es crucial el manejo adecuado de este grupo de pacientes para intentar frenar la progresión de la enfermedad renal. De esta forma retrasaremos el inicio del tratamiento sustitutivo de la función renal (diálisis o trasplante renal) y reduciremos complicaciones asociadas a esta enfermedad en estadios avanzados.

La velocidad de progresión de la enfermedad renal crónica no es similar para todos los pacientes, de forma que unos evolucionan a estadios avanzados mucho más rápido que otros, en función de diferentes variables (Yuste 2013, Grams 2015).

Se han implicado múltiples factores clínicos y analíticos en la progresión de la enfermedad renal, como hipertensión arterial, proteinuria mayor de 1.5 g/día, nefropatía de base, raza negra, sexo masculino, obesidad, diabetes mellitus (DM), dislipemia, tabaco, dietas hiperproteicas, retención de fosfato, acidosis metabólica, hiperuricemia (Klahr 2003, Taal 2006, Collister 2016). Queda por determinar de una forma más definitiva, de qué manera pueden influir los diferentes factores o la combinación de varios de ellos en la progresión de la enfermedad renal crónica.

En este sentido se han realizado múltiples estudios para intentar predecir la evolución de la enfermedad renal usando variables clínicas y analíticas. En todos ellos se han utilizado técnicas que podríamos llamar clásicas: contraste de hipótesis, modelos multivariantes lineales, ANOVA, etc. Sin embargo, la evolución del *hardware/software* permite la aplicación de técnicas más modernas y potentes a grandes bases de datos. Actualmente se aplican en todas las áreas del conocimiento, siendo la clínica una de ellas.

En el presente trabajo, además de las pruebas estadísticas “clásicas”, se emplean técnicas denominadas de aprendizaje máquina en la caracterización de los pacientes con enfermedad renal crónica para intentar determinar la influencia de diferentes variables clínicas y analíticas en la progresión de la misma. Además, se desarrollan modelos predictivos de progresión de enfermedad renal.

### 1.2. OBJETIVO

El objetivo de esta tesis doctoral es analizar mediante técnicas de aprendizaje máquina diferentes variables en un grupo de pacientes con enfermedad renal crónica en categorías 3b, 4 y 5. Las conclusiones derivadas de este análisis nos permitirán determinar de una manera más precisa la progresión de esta enfermedad de dos maneras:

1. Determinando la influencia que tienen las diferentes variables estudiadas sobre la evolución de la enfermedad renal.

2. Desarrollando un modelo predictivo de progresión de enfermedad renal basado en estas variables que ayudará en la práctica clínica a determinar cuando se va a alcanzar un determinado nivel de filtrado glomerular para poder programar el inicio de terapia sustitutiva en estos pacientes.

### 1.3. HIPÓTESIS

En la presente tesis se plantea:

**Hipótesis nula:** Las técnicas de aprendizaje máquina no van a aportar ningún aspecto novedoso en la caracterización de los pacientes con enfermedad renal crónica en categorías 3b, 4 y 5 ni en la predicción de su evolución.

**Hipótesis alternativa:** La aplicación de las citadas técnicas a nuestra base de datos va a posibilitar una caracterización minuciosa de los pacientes mediante la utilización de redes neuronales (mapas autoorganizados), que nos va a permitir incidir en determinados aspectos de su manejo para mejorar su pronóstico. Asimismo, se podrán desarrollar modelos predictivos mediante modelos probabilísticos avanzados que nos ayudarán a programar su necesidad de inicio de tratamiento sustitutivo de la función renal.

#### 1.4. FILTRADO GLOMERULAR

El riñón tiene funciones excretoras, metabólicas y endocrinas. El filtrado glomerular (FG) es un componente de la función excretora del riñón, pero se considera el mejor índice global de la función renal, puesto que generalmente se reduce después de un daño estructural renal y la mayoría de las funciones restantes del riñón descienden en paralelo con el filtrado glomerular.

La tasa de filtrado glomerular es la herramienta que utilizamos en la práctica clínica para definir la función renal de un paciente. Su valor normal depende de la edad, el sexo y el tamaño corporal, y es aproximadamente de 130 y 120 ml/min/1.73 m<sup>2</sup> para hombres y mujeres respectivamente. A partir de los 40 años de edad se produce una reducción fisiológica del FG del orden de 0,75 ml/min/1.73 m<sup>2</sup>/año aproximadamente.

Hoy en día disponemos de diferentes métodos de medición del filtrado glomerular, siendo en la actualidad los más empleados las fórmulas estimativas desarrolladas a partir de la creatinina sérica junto a variables demográficas y/o antropométricas, como las derivadas del estudio *MDRD* o la *CKD-EPI*, desarrollada en 2009 y más precisa en pacientes con FG mayor de 60 ml/min/1.73 m<sup>2</sup> (Levey 2009).

## 1.5. ENFERMEDAD RENAL CRÓNICA

### 1.5.1. DEFINICIÓN

La enfermedad renal crónica en el adulto se define como la presencia de una alteración estructural o funcional renal que persiste más de 3 meses, o por un filtrado glomerular  $<60 \text{ ml/min/1.73 m}^2$  sin otros signos de lesión renal. Este concepto fue propuesto por la *NFK-KDOQI* en 2002 y se ha ratificado en posteriores revisiones (*KDOQI* 2005, *KDIGO* 2012). En las guías publicadas en 2012 se añade además que la alteración renal debe tener implicaciones para la salud y se habla de marcadores de daño renal para referirse a las alteraciones estructurales (*KDIGO* 2012, Gorostidi 2014).

Como marcadores de daño renal se han incluido: Presencia de albuminuria mayor de 30 mg/día, alteraciones en el sedimento urinario, alteraciones electrolíticas u otros trastornos de origen tubular, anomalías histológicas, alteraciones estructurales detectadas por pruebas de imagen o trasplante renal (*KDIGO* 2012).

La enfermedad renal crónica se considera el final común de una constelación de alteraciones que afectan al riñón de forma crónica e irreversible (Lorenzo 2010).

### 1.5.2. CLASIFICACIÓN

La enfermedad renal crónica se clasificó en 2002 en 5 estadios en función del filtrado glomerular (*KDOQI* 2002):

- Estadio 1: Filtrado glomerular mayor o igual a  $90 \text{ ml/min/1.73 m}^2$  (normal o elevado), acompañado de lesión estructural renal.
- Estadio 2: Filtrado glomerular entre  $60$  y  $89 \text{ ml/min/1.73 m}^2$  (descenso leve).
- Estadio 3: Filtrado glomerular entre  $30$  y  $59 \text{ ml/min/1.73 m}^2$ .
- Estadio 4: Filtrado glomerular entre  $15$  y  $29 \text{ ml/min/1.73 m}^2$  (descenso severo).

- Estadio 5: Filtrado glomerular por debajo de  $15 \text{ ml/min/1.73 m}^2$  (fallo renal).

En revisiones posteriores se han realizado pequeñas modificaciones a esta clasificación. Así, en 2005, el grupo *KDIGO* añadió dos categorías complementarias: la T en cualquier estadio, si el paciente era portador de un trasplante renal funcionan y la D, en estadio 5, si el paciente recibía tratamiento dialítico. (Levey 2005).

Más tarde, las guías *NICE* en 2008 propusieron dividir el estadio 3 en dos subestadios:

- 3a: Filtrado glomerular entre  $45$  y  $59 \text{ ml/min/1.73 m}^2$ .

- 3b: Entre  $30$  y  $44 \text{ ml/min/1.73 m}^2$ .

En la última revisión del grupo *KDIGO* se modifica la nomenclatura y se habla de grados o categorías en lugar de estadios. Todos los grados descritos se van a clasificar a su vez, dependiendo de la presencia de albuminuria, en (*KDIGO* 2012):

- A1: Albuminuria menor de  $30 \text{ mg/día}$ .

- A2: Albuminuria entre  $30$  y  $300 \text{ mg/día}$ .

- A3: Albuminuria  $> 300 \text{ mg/día}$ .

Además, se incluye la enfermedad de base en la clasificación. Los autores, a partir del grado de FG y de albuminuria, establecen una clasificación de riesgo (Figura 1.1) pero advierten que es importante considerar además la enfermedad de base y otros factores de riesgo y comorbilidad para establecer una predicción de progresión de la enfermedad renal. Estos autores recomiendan el uso de scores de riesgo para mejorar el manejo de la enfermedad.

Prognosis of CKD by GFR and Albuminuria Categories: KDIGO 2012				Persistent albuminuria categories Description and range		
				A1	A2	A3
				Normal to mildly increased	Moderately increased	Severely increased
				<30 mg/g <3 mg/mmol	30-300 mg/g 3-30 mg/mmol	>300 mg/g >30 mg/mmol
GFR categories (ml/min/1.73m <sup>2</sup> ) Description and range	G1	Normal or high	≥90			
	G2	Mildly decreased	60-89			
	G3a	Mildly to moderately decreased	45-59			
	G3b	Moderately to severely decreased	30-44			
	G4	Severely decreased	15-29			
	G5	Kidney failure	<15			

Figura 1.1. Pronóstico de la enfermedad renal crónica en función de las categorías del FG y de la albuminuria: Verde: Riesgo muy bajo de progresión si no existen otros marcadores de enfermedad renal. Amarillo: Riesgo moderado. Naranja: Riesgo elevado. Rojo: Riesgo muy elevado de desarrollar fallo renal (KDIGO-2012).

### 1.5.3. MODELO CONCEPTUAL PARA LA ENFERMEDAD RENAL CRÓNICA

La *National Kidney Foundation (NKF)* ya en 2002 desarrolló el modelo conceptual de enfermedad renal crónica que representa esta entidad como un proceso continuo en su desarrollo, progresión y complicaciones. El final de este proceso sería el fallo renal, definido como  $FG < 15$  ml/min o por la necesidad de tratamiento sustitutivo renal (*Figura 1.2*).

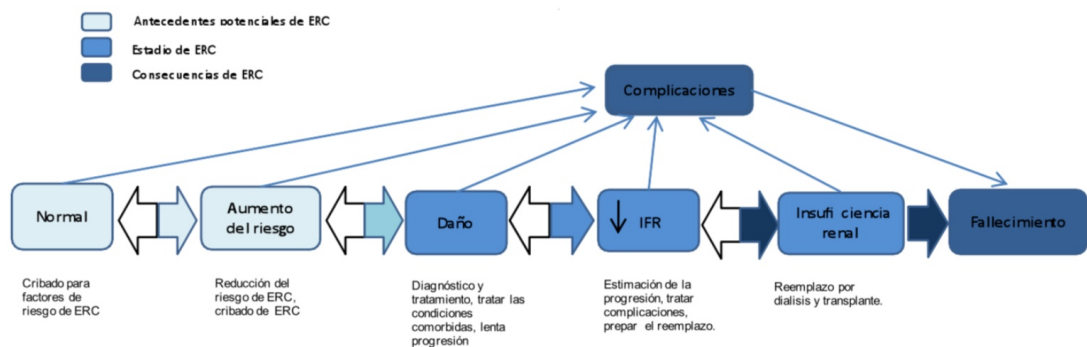


Figura 1.2. Modelo conceptual para la enfermedad renal crónica (NKF)(Levey 2005).

Como vemos en la *Figura 1.2*, en las primeras fases únicamente tendríamos existencia de factores de riesgo para enfermedad renal o antecedentes familiares de la misma, y es en este momento donde son importantes las medidas de *screening* o de prevención primaria, para frenar la evolución del proceso. Posteriormente se establece el daño renal, en una primera fase con función renal normal y, posteriormente, con disminución progresiva de la misma hasta llegar a la situación de fallo renal. En estas fases intermedias de daño renal o descenso de la función renal, se deben implementar medidas de prevención secundaria para frenar la evolución de la enfermedad y reducir complicaciones. Si no se consigue frenar el proceso, llegamos a la situación de fallo renal, donde se instauran las medidas sustitutivas de la función renal. Sin embargo, aunque se trata de un proceso continuo,



la velocidad de progresión de la enfermedad es variable de unos pacientes a otros, y no todos ellos progresan, por lo que el diagnóstico de enfermedad renal crónica no implica necesariamente el desarrollo de fallo renal.

La evolución entre las diferentes fases puede ser bidireccional, aunque la posibilidad de regresión a una fase previa siempre es menor. Esta situación de regresión se puede producir en los estadios más precoces y también en pacientes que han sido sometidos a un trasplante renal.

En cualquier momento de la enfermedad se pueden desarrollar complicaciones que van a influir en la morbimortalidad del paciente. El modelo subraya la posibilidad de que se produzcan complicaciones en las fases más precoces de la enfermedad, que pueden conducir a la muerte de los pacientes sin llegar a la situación de fallo renal.

El modelo conceptual es importante porque identifica una población con factores de riesgo para el desarrollo de enfermedad renal crónica, alguno de los cuáles son modificables. La intervención sobre estos factores de riesgo podría frenar la evolución de la enfermedad renal. Además, enfatiza el hecho de que desarrollar estrategias para la prevención, detección precoz y tratamiento de las complicaciones de la enfermedad renal crónica puede prolongar la supervivencia y aumentar la calidad de vida de los pacientes, aunque no sea posible frenar la progresión de la enfermedad renal (Levey 2009).

## 1.6. PROGRESIÓN DE ENFERMEDAD RENAL

### 1.6.1. DEFINICIÓN

La progresión de la enfermedad renal es la pérdida de función renal que se va produciendo a lo largo del tiempo, y se caracteriza por el descenso progresivo del filtrado glomerular. La tasa media anual de disminución del filtrado glomerular en población sana es de 0,7-1 ml/min/1,73 m<sup>2</sup> a partir de los 40 años de edad (Stevens 2006). Esta progresión es más rápida en pacientes con enfermedad renal crónica grados 3-5, siendo especialmente importante a partir del grado 4 (FG < 30 ml/min), donde habitualmente se produce un deterioro más rápido hasta llegar a la situación de fallo renal con requerimiento de tratamiento sustitutivo de la función renal.

A lo largo de los últimos años se han establecido diferentes criterios para intentar cuantificar la progresión de la enfermedad renal. Ya en 2002, las guías *KDOQI* definieron la progresión como el descenso del nivel de función renal en un paciente a lo largo del tiempo, o como la necesidad de inicio de tratamiento de la función renal. Posteriormente, en diferentes ensayos clínicos realizados en pacientes con función renal normal, se ha seguido el criterio establecido por la *FDA* según el cuál se considera que existe progresión de enfermedad renal si se duplica la creatinina sérica sobre su nivel basal o si se precisa tratamiento sustitutivo de la función renal.

Las guías *NICE 2008* establecen la progresión de la ERC como el descenso del filtrado glomerular mayor de 5 ml/min/1,73 m<sup>2</sup>/año o mayor de 10 ml/min/1,73 m<sup>2</sup> en cinco años (*NICE 2008*).

En el documento de consenso publicado en 2012 por varias sociedades científicas españolas se ratifica este concepto de progresión (Documento de consenso sobre ERC).

Finalmente las guías *KDIGO* de 2012 definen la progresión de la enfermedad renal como el descenso de categoría de FG acompañado de una pérdida del 25 % del basal, o bien por un descenso anual del FG  $\geq 5$  ml/min/1.73 m<sup>2</sup> (*KDIGO 2012*).

---

La velocidad de la pérdida de función renal varía considerablemente entre individuos con la misma enfermedad renal de base o con el mismo grado de función renal. Esta variabilidad individual es típica de enfermedades complejas y refleja la naturaleza multifactorial de los mecanismos biológicos implicados en la evolución de la enfermedad subyacente (Mccellan 2003, Collister 2016).

Así, el estudio "*Modification of Diet in Renal Disease (MDRD)*" reportó que la velocidad media de descenso de la función renal durante el seguimiento fue mayor en los pacientes que partían de un FG más bajo, pero también se encontraron casos en los que no se evidenció ninguna progresión durante el seguimiento, lo que traduce la amplia variabilidad interindividual en la misma (Hunsicker 1997).

Posteriormente se han publicado otros trabajos que ratifican este hecho, en los que se han descrito diferentes patrones de descenso del filtrado glomerular, con diferentes velocidades de progresión, y con comportamientos lineales o mixtos, pero en los que no se determina que factores condicionan los diferentes comportamientos (O'Hare 2012, Li 2012).

### **1.6.2. FACTORES DE PROGRESIÓN**

#### **1.6.2.1. Concepto**

Los factores de progresión de la enfermedad renal crónica son aquellos que empeoran el daño renal y aceleran el deterioro funcional. Entre los factores de riesgo de progresión los mejor documentados son la proteinuria y la hipertensión arterial (Tsai 2016). Además se han implicado otros factores en el inicio y la progresión de la enfermedad renal crónica, la mayoría de los cuáles han demostrado asociación más que causalidad, y muchas veces de forma inconsistente. Entre estos factores los más importantes son: edad, raza negra, diabetes, sexo masculino, hipoalbuminemia, dislipidemia, sobrepeso, hemoglobina glicosilada, hiperfosforemia, anemia, nacimiento con bajo peso, tabaquismo e hiperuricemia.

Se han realizado múltiples estudios para intentar determinar la influencia de estos factores de riesgo, de forma aislada o en combinación, sobre la progresión de la enfermedad renal, con el objeto de buscar una explicación a la variabilidad de

progresión entre diferentes pacientes y modificar estos factores para ralentizar esta evolución. Por ejemplo, en el estudio anteriormente citado “Modification of Diet in Renal Disease study” (Hunsicker 1997), realizado en una población de 840 pacientes con diversas enfermedades renales, se encontró en el análisis multivariante que las siguientes características eran predictores independientes de un descenso más rápido en la tasa de filtración glomerular: proteinuria severa, presión arterial elevada, raza negra, colesterol HDL descendido y niveles bajos de transferrina. En estudios posteriores se han confirmado como predictores independientes de progresión de insuficiencia renal la proteinuria elevada y la hipertensión arterial.

Con respecto a otros factores de riesgo los resultados no han sido concluyentes.

#### **1.6.2.2. Importancia**

El conocimiento de los diferentes factores de progresión y su papel es importante para identificar individuos con mayor riesgo de progresión hacia enfermedad renal terminal y así establecer estrategias de intervención que puedan frenar esta evolución.

#### **1.6.2.3. Descripción de los diferentes factores de progresión**

##### *a) Hipertensión arterial*

La hipertensión arterial se define como la presencia de cifras de tensión arterial por encima de 140/90 mmHg (Mancia 2013). Tiene una elevada prevalencia entre los pacientes con enfermedad renal crónica y, además es un factor de riesgo para el desarrollo de nefropatía (Klag 1996, Drawz 2013). Por este motivo se ha asumido, basándose en estudios observacionales, que el descenso importante de las cifras de TA puede reducir la progresión de la enfermedad renal. Sin embargo, diferentes ensayos clínicos realizados en pacientes con enfermedad renal crónica (AASK, REIN-2, MDRD) no han podido demostrar que cifras muy bajas de presión arterial puedan ralentizar la progresión de la insuficiencia renal, salvo en aquellos casos que además presentan proteinuria y, sin embargo, se han encontrado efectos adversos como consecuencia del tratamiento antihipertensivo o de mantener unas

cifras tensionales más bajas (Upadhyay 2011).

Por tanto, se desconoce la cifra óptima de tensión arterial para conseguir ralentizar la progresión de la enfermedad renal, aunque se recomienda para estos pacientes mantener una tensión arterial menor de 140/90 mmHg y, en aquellos pacientes que además presenten proteinuria, por debajo de 130/80 mmHg (KDIGO 2012).

### *b) Proteinuria*

La proteinuria, definida como la presencia en orina de más de 500 mg diarios de proteínas, es una manifestación común de muchas nefropatías crónicas independientemente de su causa. Además, es uno de los factores de progresión de enfermedad renal mejor establecidos. Por este motivo, las guías clínicas para el manejo de la enfermedad renal crónica, la han añadido junto al filtrado glomerular para su clasificación y estratificación del riesgo (KDIGO-2012).

En una larga serie de estudios observacionales y ensayos clínicos se ha encontrado una correlación entre la cantidad de proteínas excretadas en la orina y la velocidad de descenso del filtrado glomerular, tanto en pacientes diabéticos como no diabéticos. Así, un estudio observacional realizado en un período de 20 años en pacientes de raza caucásica encontró que la presencia de proteinuria en una tira reactiva fue un factor de riesgo independiente para progresión de enfermedad renal y para mortalidad global (Tarver-Carr 2000). Otro estudio (realizado en Japón en una población de casi 100.000 pacientes) encontró que el riesgo de progresión a enfermedad renal avanzada tras siete años de seguimiento prácticamente se limitaba a aquellos pacientes con proteinuria en tira reactiva al inicio del seguimiento, independientemente de la función renal que presentaran en ese momento (Iseki 2003).

El estudio *RENAAL* (Zeeuw 2004), realizado con 1532 pacientes con DM tipo II y nefropatía, reveló que aquellos con proteinuria > 3 g/g creatinina (tras ajustar por datos demográficos, analíticos y clínicos) tenían un riesgo 5.2 veces mayor de precisar terapia sustitutiva de la función renal y 8.1 veces mayor de progresión de la enfermedad renal. Asimismo, se demostró que si se conseguía reducir la albuminuria

---

un 50% en los seis primeros meses se producía un descenso de un 36 % en el riesgo de desarrollar una ERC terminal y un descenso de un 45 % en el riesgo de progresión de la misma.

El estudio *REIN*, realizado en pacientes no diabéticos, también demostró una mayor progresión de enfermedad renal en el grupo de pacientes con proteinuria mayor de 3 gramos/día. Se objetivó que este efecto era independiente del control de la tensión arterial. En estos pacientes se demostró además un menor descenso en el FG en el grupo de mayor proteinuria cuando se randomizaban a ramipril (Remuzzi 2004).

Más recientemente se han publicado varios metaanálisis que ratifican el papel de la proteinuria como factor de progresión de enfermedad renal (Astor 2011, Inker 2014, Tsai 2016). De estos estudios se concluye que la presencia de proteinuria es un factor de progresión de enfermedad renal, tanto mayor cuanto más elevado es el grado de proteinuria y que su control es prioritario para frenar la progresión de la enfermedad.

### c) Raza

En algunos estudios se ha encontrado relación entre la raza negra y la progresión de la enfermedad renal crónica. En un subanálisis del estudio *MRFIT* se describió un mayor riesgo de insuficiencia renal avanzada en pacientes de raza negra independientemente de otros factores. También un análisis de las características basales de los pacientes del estudio *MDRD* identificó la raza negra como un predictor independiente para un descenso más rápido del filtrado glomerular.

Se pensó que estas diferencias podían atribuirse a un peor acceso de los pacientes de raza negra a los sistemas de salud, puesto que eran estudios realizados en E.E.U.U. Sin embargo, en un trabajo realizado en Holanda, donde todos los pacientes tienen un acceso similar a la sanidad, se comparaba la evolución de pacientes en prediálisis de raza caucásica y raza negra, y se encontró una progresión de la enfermedad renal de 0.18 ml/min/1.73 m<sup>2</sup> mensuales mayor en los pacientes de raza negra (Van de Beukel 2013). Los resultados de este trabajo apoyan el papel de la raza negra como factor de progresión de la enfermedad renal.

---

Se han propuesto además de aspectos socioambientales, factores genéticos para explicar este hecho. Parte de esta mayor susceptibilidad podría justificarse por una mayor activación de citoquinas, puesto que se han encontrado niveles de *ET-1* y *TGF- $\beta$*  más elevados en pacientes de raza negra e hipertensos (Yu 2003).

### d) *Diabetes*

La diabetes y su control influyen en la rapidez de progresión de la enfermedad renal. En 1997, Branchati *et al.* concluyeron que la diabetes es un factor de riesgo independiente para el desarrollo de enfermedad renal terminal, aunque ésta esté producida por otras causas (Branchati 1997).

Lorenzo *et al.* encontraron resultados similares en un estudio retrospectivo realizado sobre 333 pacientes en el que se determinó que la diabetes era un factor independiente para el desarrollo de enfermedad renal, aunque este hecho se asociaba a la presencia de albuminuria (Lorenzo 2010).

Tanto en diabetes tipo I como en la tipo II se puede desarrollar nefropatía diabética, y, en este caso, el grado de albuminuria va a condicionar la velocidad de descenso del filtrado glomerular.

En la *DM* tipo I, a los 15 años del diagnóstico de la enfermedad, de un 20 a 30% de pacientes tienen albuminuria levemente incrementada. De estos, un 50% desarrollarán albuminuria severamente elevada y progresarán a enfermedad renal (Stanton 2014). En los últimos años, con el mayor control de la glucemia y el empleo de *IECAs* o *ARA II* ha disminuido la velocidad de progresión de la nefropatía diabética en *DM* tipo I. Así, en 2011, en un estudio prospectivo randomizado realizado en pacientes diabéticos de tipo I en el que los pacientes se asignaban a dos grupos, uno con un control intensivo de la diabetes y otro con un control convencional, se demostró que el grupo con un control más agresivo presentaba una evolución de la enfermedad renal un 50 % más lenta que el otro grupo (*DCCT/EDIC group* 2011). Este estudio proporciona una de las evidencias más fuertes acerca del beneficio del manejo agresivo de la diabetes.

En DM tipo II se ha visto que el tiempo de aparición de proteinuria desde el comienzo de la enfermedad, y el tiempo de desarrollo de enfermedad renal terminal desde la aparición de la proteinuria son similares a la tipo I. Este hecho viene apoyado por los resultados del estudio *UKPDS* realizado en Reino Unido de forma prospectiva en 5000 pacientes diabéticos en los que se recogieron datos de función renal, albuminuria, etc. durante 25 años. En este estudio se describió que los pacientes sin albuminuria se mantenían libres de enfermedad durante 19 años, los que presentaban albuminuria leve desarrollaban albuminuria severa en unos 10 años, y la mitad de estos precisaban tratamiento sustitutivo de la función renal en 10 años (Adler 2003).

El papel de la diabetes en el inicio de la enfermedad renal está ampliamente demostrado. Su papel como factor de progresión de la nefropatía ya establecida también se ha analizado, aunque son menos los estudios que lo demuestran y con resultados no tan claros. En 2008 Johnson *et al.* realizaron un estudio retrospectivo en 9782 pacientes con enfermedad renal crónica con un FG entre 15 y 60 ml/min para diseñar un modelo predictivo de progresión hacia enfermedad renal terminal, evidenciando que la diabetes era uno de los factores de progresión (Johnson 2008). Sin embargo, otros trabajos no han podido confirmar este dato, como el realizado por Yuste en 2013, en el que se estudian de forma retrospectiva 300 pacientes con ERC en categorías 4 y 5 para determinar factores de progresión, en el que no se encontraron diferencias con respecto a la diabetes (Yuste 2013). En un meta-análisis publicado en 2016 en el que se revisan 40 estudios sobre progresión de enfermedad renal crónica, se concluye que la diabetes puede ejercer un papel menor en la evolución de estos pacientes (Tsai 2016). Por todo ello, son necesarios más estudios para confirmar el papel de la diabetes mellitus en la progresión de la enfermedad renal cuando ésta ya está establecida.



### e) *Dislipemia*

En los pacientes con enfermedad renal crónica se producen alteraciones en el metabolismo de las lipoproteínas que resultan en niveles elevados de triglicéridos y de colesterol-LDL (Kalaitzidis 2011). Se han realizado estudios experimentales en animales que sugieren que niveles elevados de lípidos pueden promover progresión de enfermedad renal, y que su descenso mediante una estatina reduce la progresión de daño renal. En estudios epidemiológicos realizados en población general, la presencia de dislipemia se ha asociado con mayor riesgo de desarrollar enfermedad renal crónica (Schaeffner 2003).

Sin embargo, en pacientes con enfermedad renal crónica el papel de la dislipemia no acaba de estar claro. Existen dudas sobre si la hiperlipidemia es causa de progresión de la enfermedad renal o bien, si el daño renal y la proteinuria son los causantes de la progresión de la enfermedad renal y de la alteración del perfil lipídico. En alguna pequeña serie se ha podido demostrar el papel de la dislipemia como factor de progresión de la enfermedad renal (Samuelsson 1997). Sin embargo, el estudio *CRIC*, prospectivo, observacional, multicéntrico, en el que se incluyeron 3983 pacientes con enfermedad renal leve-moderada, para examinar factores de riesgo de progresión de enfermedad renal y cardiovascular, no encontró ninguna asociación como factor de riesgo independiente entre los niveles de colesterol total, *HDL-C*, *LDL-C*, *Apo A1*, *Apo B*, *Lp(A)* y *VLDL-C* y la progresión de la enfermedad renal. Únicamente se encontró una relación inversa entre los niveles de colesterol total y la evolución de la función renal en pacientes con proteinuria menor de 0.2 g/día (Rahman 2014).

En cuanto a la influencia del tratamiento con estatinas en la progresión no hay resultados claros. Se ha demostrado un descenso en la morbimortalidad cardiovascular en estas fases de la enfermedad renal, pero su influencia sobre la progresión de la misma no está bien definida (Kalaitzidis 2011). Los resultados no son concluyentes porque los estudios realizados para evaluar el efecto de las estatinas sobre la progresión de la enfermedad renal son subanálisis de ensayos diseñados para analizar la eficacia de estos fármacos sobre la enfermedad

cardiovascular. Por ejemplo, en un análisis *post-hoc* del estudio *CARE* (diseñado para evaluar mortalidad cardiovascular en más de 4000 pacientes hiperlipidémicos que habían sufrido un infarto de miocardio) se analizaron 690 pacientes con FG < 60 ml/min y se encontró una menor progresión de la enfermedad renal en pacientes en tratamiento con pravastatina y un FG < 40 ml/min (Tonelli 2003). Sin embargo, el estudio *SHARP* (Baigent 2011), realizado de forma prospectiva sobre 9720 pacientes con insuficiencia renal, 6247 de ellos no en diálisis, en el que se randomizaban los pacientes a un grupo de tratamiento con simvastatina y ezetimiba, y un segundo grupo con placebo, demostró una reducción de un 17% en la aparición de eventos aterotrombóticos, pero no se pudo demostrar una reducción en la velocidad de progresión de la enfermedad.

En cuanto a los triglicéridos, su papel como factor de progresión de la enfermedad renal también es dudoso. En el estudio *ARIC* la elevación de triglicéridos se asoció a un aumento en la creatinina sérica (Muntner 2000). También en otro estudio realizado en un grupo de pacientes con nefropatía IgA, la hipertrigliceridemia se asoció con una elevación del 20 % en los niveles de creatinina sérica. Sin embargo en otros estudios realizados no se ha podido demostrar esta relación.

#### *f) Obesidad*

La obesidad favorece el desarrollo y complica el curso de varias enfermedades como la diabetes, la hipertensión arterial, enfermedades cardiovasculares, síndrome metabólico, algunas neoplasias y la enfermedad renal.

Un número considerable de estudios han examinado la asociación entre obesidad y enfermedad renal, sugiriendo que la obesidad aumenta el riesgo de desarrollarla y también acelera su progresión. En 2008, Wang *et al.* publicaron un metaanálisis (Wang 2008) basado en estudios prospectivos realizados en diferentes poblaciones, concluyendo que las personas con sobrepeso tenían un riesgo mayor de desarrollar enfermedad renal, y este riesgo aumentaba cuanto mayor era el grado de obesidad, con un riesgo relativo (RR) de 1.40 (1.30–1.50) para pacientes con índice de masa corporal (IMC) de 25 a 30 y de 1.83 (1.57–2.13) para aquellos con IMC mayor de 30 kg/m<sup>2</sup>.

En este metaanálisis también se evidenció que la asociación entre obesidad y enfermedad renal era más fuerte en mujeres que en hombres.

En uno de los mayores estudios realizados en una población de E.E.U.U. e incluidos en el metaanálisis citado, Hsu *et al.* documentaron que el IMC se relaciona con el riesgo de ERC. En este estudio, personas con un IMC >40 mostraron un riesgo para desarrollar ERC siete veces mayor que aquellos que tenían un IMC normal (Hsu 2006).

La asociación entre obesidad y progresión de la enfermedad renal crónica se demostró también en una serie de pacientes con nefropatía IgA, en los que un IMC > 25 kg/m<sup>2</sup> predijo una progresión más rápida de la insuficiencia renal (Bonnet 2001). Sin embargo, estudios realizados posteriormente en pacientes con ERC estadios 3 y 4 demuestran resultados contradictorios con respecto a la relación entre la obesidad y la progresión de la enfermedad renal. Así, Lu *et al.* reportaron una relación entre el normopeso y un menor deterioro de función renal, pero en los pacientes obesos no encontraron diferencias. Otros autores, como Munter *et al.* y De Nicola *et al.*, en sus respectivos trabajos, no encontraron asociación entre obesidad y mayor progresión de la enfermedad renal, sin embargo, Dalrymple *et al.* sí pudieron demostrar esta asociación (Ahmadi 2016).

La obesidad puede promover daño renal directamente a través de efectos hemodinámicos y hormonales (disminución del flujo plasmático renal que conduciría a vasoconstricción de la arteriola eferente e hiperfiltración, elevación de dimetil arginina asimétrica, hiperaldosteronismo con producción de citoquinas proinflamatorias por los adipocitos, etc), y también de forma indirecta favoreciendo el desarrollo de hipertensión arterial y de diabetes tipo II, entidades con influencia negativa sobre la progresión de la enfermedad renal. No obstante, en los estudios en los que se han corregido factores de confusión como la presencia de diabetes e hipertensión arterial se ha visto que la obesidad es un factor de riesgo independiente para la progresión de la enfermedad renal (Eknoyan 2011).

La obesidad se considera una enfermedad multifactorial en la cuál el tejido adiposo, más que constituir un almacén de energía, funciona como un órgano

exocrino y endocrino con efectos vasoactivos y neurohumorales que están implicados en la génesis del daño orgánico asociado a la obesidad, incluyendo el daño renal. La lesión común en los pacientes obesos con enfermedad renal crónica es una glomerulosclerosis segmentaria y focal y se considera que es consecuencia de la hiperfiltración que se produce en estos pacientes (Mallamaci 2013).

La primera manifestación clínica de esta hiperfiltración es la aparición de albuminuria, que se ha visto que es mayor cuanto mayor es el índice de masa corporal y que, además, aumenta si coexiste la presencia de hipertensión arterial y/o diabetes (Eknoyam 2011).

*g) Hiperfosforemia*

En la enfermedad renal crónica se produce una retención de fosfato desde estadios precoces. Esta retención de fósforo, además de intervenir en la enfermedad ósea, parece que también puede contribuir a la progresión de la enfermedad renal.

En un estudio observacional realizado en 985 pacientes seguidos durante una media de 2 años, Schwarz *et al.* observaron que la *hazard ratio* ajustada para duplicar la creatinina sérica fue 1.3 por cada 1 mg/dl de incremento en el fósforo sérico, y se encontró una relación similar para el producto calcio-fósforo (Schwarz 2006). También Yuste *et al.* realizaron un estudio retrospectivo sobre 300 pacientes con ERC avanzada para determinar el índice de progresión de la enfermedad renal e identificar predictores, y concluyeron que los niveles elevados de fósforo en sangre se correlacionaban de forma independiente con un descenso del FG, al igual que la proteinuria y los niveles de PTH (Yuste 2013).

En un análisis *post-hoc* del estudio *REIN*, se sugiere que el nivel basal de fósforo es un factor de riesgo independiente para la progresión de la enfermedad renal (Zocalli 2011). Posteriormente Li *et al.* realizan un trabajo en 74 pacientes con enfermedad renal e hipertensión en el que comparan periodos de estabilidad de la función renal con periodos de deterioro más rápido y encontraron también niveles más elevados de fósforo en los periodos de mayor descenso del FG (Li 2014).

Todos estos estudios no permiten demostrar una relación causa-efecto entre la hiperfosfatemia y el deterioro de la función renal, y no existen datos que demuestren una mejoría de la función renal tras mejorar el control del fósforo, por lo que en la actualidad no se puede considerar un factor independiente de progresión de enfermedad renal (Collister 2016).

Un posible mecanismo causal podría ser la precipitación de fosfato cálcico en el intersticio renal, que podría iniciar una reacción inflamatoria que resultaría en fibrosis intersticial y atrofia tubular.

### *h) Tabaquismo*

Además de ser un factor de riesgo para enfermedades neoplásicas, cardiovasculares y pulmonares, en los últimos años se ha demostrado que el tabaquismo es un factor de riesgo independiente para progresión de enfermedad renal (Hallan 2011).

El efecto deletéreo del tabaco sobre el riñón podría estar mediado, al menos de forma parcial, por la acción de la nicotina. En modelos animales se ha demostrado que a nivel vascular promueve arteriosclerosis y angiogénesis, y a nivel mesangial induce proliferación. Además, tras la administración de nicotina se producía un empeoramiento del daño renal. Los mecanismos implicados son complejos y estarían mediados por activación a diferentes niveles de receptores nicotínicos de la acetilcolina (Jain 2013).

En estudios realizados en humanos con función renal normal se ha encontrado una fuerte asociación entre el consumo de tabaco y la excreción urinaria de albúmina, tanto en sujetos sanos como en pacientes afectados de diabetes mellitus y/o hipertensión arterial, sugiriendo la implicación del tabaco en la generación de daño renal “de novo” (Pinto-Sietsma 2000, Halimi 2000).

En 1998, Orth *et al.* publicaron el primer trabajo en el que se relacionaba al tabaco directamente con la enfermedad renal. Se trataba de un estudio multicéntrico en el que demostraron que varones fumadores con nefropatía IgA, o con poliquistosis

renal autosómica dominante, tenían un mayor riesgo de progresión a enfermedad renal crónica estadio 5 cuando se comparaban con no fumadores (Orth 1998).

Posteriormente, en otros trabajos realizados se han confirmado estos hallazgos. Uno de los más importantes ha sido el publicado por Hallan y Orth en 2011, en el que encontraron que, en menores de 70 años, los pacientes fumadores y exfumadores tenían un riesgo 4 y 3.3 veces mayor de alcanzar enfermedad renal crónica estadio 5 que los no fumadores. No se encontraron diferencias en mayores de 70 años (Hallan 2011).

En una revisión publicada en 2014 sobre 12 trabajos publicados tanto de estudios casos-controles como de cohortes en pacientes con enfermedad renal, los autores concluyen que el tabaco es un factor de riesgo de progresión de la enfermedad renal, y que su efecto es dosis-dependiente, de forma que se ha encontrado una mayor progresión en aquellos pacientes fumadores de > 15 paquetes/año. Además se ha demostrado una correlación entre el abandono del tabaco y la mejoría de la función renal, excepto en aquellos pacientes que fumaban más de la cantidad citada anteriormente (Elihimas 2014).

*i) Hiperuricemia*

La hiperuricemia se define como la acumulación de ácido úrico en suero por encima de su punto de solubilidad en agua (6.8 mg/dl) y puede producirse por sobreproducción, infraeliminación, o ambos. En los pacientes con enfermedad renal crónica se produce como consecuencia del descenso en la eliminación urinaria de ácido úrico.

La hiperuricemia puede inducir enfermedad renal por diferentes mecanismos: a través de la formación de cristales de ácido úrico que se depositan en el túbulo colector, y también por mecanismos independientes de la formación de cristales, que pueden inducir disfunción endotelial (por reducción de producción de óxido nítrico), arteriopatía a nivel de la arteria aferente y fibrosis tubulointersticial (por activación del sistema renina-angiotensina-aldosterona) e inflamación (por producción de citoquinas) (Jalal 2013).

Se han realizado numerosos estudios epidemiológicos en pacientes con función renal basal normal (Weiner 2008, Obermayr 2008), que han mostrado una asociación entre los niveles elevados de ácido úrico y el desarrollo de nefropatía, independientemente de otros factores.

También varios estudios epidemiológicos han analizado si los niveles elevados de ácido úrico pueden favorecer la progresión de la enfermedad renal crónica, así Hsu *et al.* estudiaron una cohorte de 177.570 voluntarios durante un período de 25 años y encontraron que la hiperuricemia se asoció con un riesgo elevado de enfermedad renal terminal independientemente de otros factores como edad, raza, sexo, índice de masa corporal, nivel educativo, presión arterial, diabetes, creatinina sérica, hemoglobina y proteinuria (Hsu 2009).

Esta asociación no se ha confirmado en estudios realizados con pacientes con FG < 60 ml/min, por ejemplo, un análisis de 840 pacientes con ERC en categorías 3-4, participantes en el estudio *MDRD* no encontró el ácido úrico como factor de riesgo independiente de progresión de enfermedad renal (Madero-2009). Tampoco Nacak encuentra esta asociación en un estudio realizado en Suecia con 2466 pacientes con ERC en categorías 3 a 5, encontrando incluso menor descenso de la función renal en pacientes con niveles de ácido úrico más elevados al inicio del estudio. Una posible explicación al hecho de que la hiperuricemia acelere la progresión en estadios precoces de enfermedad renal pero no en fases más avanzadas es su acción a través de hipertensión glomerular y sistémica, que en estadios más avanzados estaría instaurada igualmente por otros mecanismos como retención de agua y sodio (Nacak 2015).

Asimismo, estudios recientes sugieren que el descenso de los niveles de ácido úrico mediante tratamiento puede ser beneficioso en la ralentización de la progresión de la ERC. Siu *et al.* realizaron un estudio randomizado en 54 pacientes hiperuricémicos con enfermedad renal leve o moderada, en el que un grupo recibía tratamiento con alopurinol y otro grupo no se trataba, observando a los 12 meses que un mayor número de participantes en el grupo control (16% vs 46% ) sufrieron elevación de creatinina, diálisis o muerte (Siu-2006). Goicoechea *et al.* realizaron

posteriormente un estudio similar en 113 pacientes hiperuricémicos con enfermedad renal crónica, observando a los dos años un mayor descenso del filtrado glomerular en el grupo control (Goicoechea 2010). En un análisis post-hoc de este estudio se confirma que el tratamiento prolongado con alopurinol puede frenar la progresión de la enfermedad renal (Goicoechea 2015). Kenji *et al.* publican en 2015 un metaanálisis en el que parece confirmarse esta ralentización de la progresión de la enfermedad renal en relación con terapias para reducir el ácido úrico (Kanji 2015), aunque por tratarse de estudios con ciertas limitaciones metodológicas, se requieren nuevos estudios para confirmar este hecho.

### j) *Anemia*

También se han relacionado la anemia y su corrección con la progresión de la ERC, sugiriéndose que la normalización de los niveles de hemoglobina en estos pacientes podría atenuar la progresión de la enfermedad renal. Así, una análisis *post hoc* del estudio *RENAAL* (Keane 2006) demostró que los niveles mayores de hemoglobina se asociaron a menor velocidad de progresión de la enfermedad renal en pacientes con diabetes tipo 2. Sin embargo, se han realizado estudios prospectivos randomizados, con resultados conflictivos en cuanto a la relación entre la cifra diana de hemoglobina y la progresión de la enfermedad renal crónica.

Un estudio prospectivo en 88 pacientes no diabéticos en prediálisis reportó efectos renoprotectores asociados con el tratamiento con epoetina, de manera que el grupo de pacientes que iniciaba tratamiento de forma precoz, tenía menor progresión de la insuficiencia renal, así como menor mortalidad (Gouva 2004). Sin embargo, un estudio realizado con 390 pacientes randomizados a una hemoglobina diana de 13-15, *versus* 11-12 g/dl, no mostró diferencias en la tasa de descenso del filtrado glomerular entre ambos grupos (Rossert 2006).

Los principales ensayos clínicos realizados sobre anemia y repercusión orgánica, como el *CHOIR*, el *CREATE* y el *TREAT*, han demostrado un efecto neutral o negativo de la corrección de la anemia con eritropoyetina sobre la progresión de la enfermedad renal (Inrig 2012). Más recientemente, en el estudio *NADIR-3*, prospectivo, realizado en pacientes con enfermedad renal crónica en grado 3, se ha



encontrado que los pacientes que desarrollaban anemia tenían un deterioro más rápido de la función renal (Portolés 2013).

### *k) Acidosis metabólica*

La acidosis metabólica es una complicación frecuente en los pacientes con enfermedad renal crónica, sobre todo cuando el filtrado glomerular cae por debajo de 30 ml/min. Aunque tradicionalmente se ha definido como el mantenimiento en estos pacientes de una concentración de bicarbonato en sangre menor de 22 mEq/l, en diferentes estudios se ha encontrado retención de ácido a pesar de mantener niveles normales de bicarbonato.

Con el deterioro de la función renal disminuye la capacidad del riñón para excretar amonio o reabsorber bicarbonato. En general, la excreción de amonio disminuye cuando el FG cae por debajo de 45 ml/min y es la principal causa de acidosis en la enfermedad renal crónica (Kraut 2016).

A medida que disminuye el número de nefronas funcionantes, aumenta la excreción fraccional de amonio por nefrona. El acúmulo local de amonio puede activar el complemento, lo cuál puede producir daño tubulointersticial secundario, y, por tanto, favorecer la progresión de la enfermedad renal (Drawz 2013). Este hecho se ha demostrado en estudios animales (Nath 1985), y también en humanos. Son muchos los estudios observaciones que han confirmado el papel de la acidosis en la progresión de la enfermedad renal, demostrando un mayor deterioro de la función renal en aquellos pacientes que mantenían unos niveles de bicarbonato por debajo de 22 mEq/l (Shah 2009, Dobre 2015).

Independientemente de su papel en la progresión de la enfermedad renal, el mantenimiento de la acidosis metabólica tiene consecuencias negativas para los pacientes que justifican su corrección, como alteraciones en el crecimiento, osteopenia, pérdida de masa muscular, aumento de catabolismo proteico, etc (Kraut 2016). Además, varios estudios han demostrado aumento de mortalidad en presencia de acidosis en pacientes con enfermedad renal crónica (Menon 2010, Raphael 2013).

Por ello, las guías clínicas vigentes en la actualidad recomiendan mantener los niveles de bicarbonato sérico por encima de 22 mEq/l (KDIGO 2012).

El tratamiento con bicarbonato ha demostrado una reducción del daño renal en estudios animales y parece que también en pacientes con enfermedad renal avanzada, como han demostrado algunos trabajos. Brito-Ashurst *et al.* realizaron un estudio prospectivo con 134 pacientes con FG 15 a 30 ml/min y bicarbonato de 16-20 mmol/l, los dividieron en dos grupos, uno con corrección de la acidosis y otro no, y realizaron un seguimiento de 2 años, concluyendo que el grupo de pacientes que recibía tratamiento experimentó un menor descenso del filtrado glomerular que el grupo no tratado (5.93 versus 1.88 ml/min/1.73 m<sup>2</sup>; p<0,0001) (Brito Arhurst 2009).

Jeong *et al.* también encontraron menor descenso del FG en pacientes tratados con bicarbonato con enfermedad renal crónica estadio 4, pero no en los pacientes con un FG menor de 15 ml/min, aunque los resultados de este estudio son menos concluyentes por tener menos pacientes (Jeong 2014).

Además de los niveles de bicarbonato, también se ha relacionado el anión GAP con la progresión de la enfermedad renal, considerándose la elevación del mismo un factor de riesgo independiente para la misma (Togawa 2013).

En los pacientes con enfermedad renal crónica, la elevación del anion GAP indica un aumento en la concentración de aniones no cuantificables, incluyendo algunas toxinas urémicas, como indoxil-sulfato o p-cresil sulfato.

#### *l) Inhibición del sistema renina-angiotensina-aldosterona*

La angiotensina II tiene efectos hemodinámicos que inducen vasoconstricción de la arteriola eferente con hipertensión glomerular, y además tiene efectos no hemodinámicos que parecen contribuir al desarrollo de fibrosis tubulo-intersticial, a través de receptores tipo IB presentes en los glomérulos, cuya activación genera factores protombóticos como *TGF-beta* y otras citoquinas (Ruiz-Ortega 2006).

En los últimos años se ha acumulado evidencia tanto en estudios en humanos como en experimentación en ratas, de que la progresión de la enfermedad

renal puede detenerse e incluso revertirse con la inhibición del sistema renina-angiotensina-aldosterona. En un estudio realizado en ratas con enfermedad renal crónica y publicado recientemente se ha demostrado que tanto los IECAs como los ARA II mejoraron claramente la histología renal, con regresión del daño tubular y glomerular y con regeneración de la vasculatura renal (Remuzzi 2016).

En un metaanálisis también publicado recientemente, en el que se incluyen 119 trabajos de pacientes en tratamiento con *IECAs/ARA II* y con enfermedad renal crónica, se concluye que el empleo de este tipo de fármacos reduce el riesgo de fallo renal y de eventos cardiovasculares, y que los IECAs tienen incluso un mayor efecto beneficioso que los ARA II, también reduciendo la mortalidad de todas las causas, por lo que deberían ser un fármaco de primera elección en este tipo de pacientes (Xie 2016). En cuanto a la asociación de ambos grupos farmacológicos, que en algunos estudios pareció beneficiosa en cuanto a reducir de proteinuria, el estudio *ONTARGET*, demostró que los pacientes que recibían tratamiento con ambos fármacos, presentaban mayor necesidad de diálisis y mayor mortalidad, así como mayor número de reacciones adversas como hiperkalemia e hipotensión. Por todo ello, en la actualidad se desaconseja el doble bloqueo del SRAA en estos pacientes (Drawz 2013).

#### **1.6.2.4. Predicción de la progresión en la enfermedad renal crónica**

Se han realizado múltiples trabajos para intentar establecer modelos predictivos de evolución de la enfermedad renal en función de determinados factores de riesgo. Estos modelos pueden ser de gran utilidad, puesto que pueden servir de guía para implementar medidas y tratamientos para frenar la progresión de la enfermedad renal, y también para planificar de forma adecuada la remisión del paciente a diálisis o la realización de un trasplante renal. Sin embargo, hasta la fecha no se emplean en la práctica diaria de forma sistemática, puesto que la mayoría de los creados presentan algún tipo de limitación, en primer lugar porque se han generado de forma retrospectiva a partir de bases de datos generadas con otros fines, o bien porque están basados en poblaciones muy concretas, que los hace

difíciles de generalizar, y también por posibles fallos metodológicos en su creación (dicotomización de variables continuas, asunción de linealidad sin tests de confirmación, exclusión de participantes por variables ausentes, etc) (Echouffo-Tcheugui 2012).

Keane *et al.*, basándose en el estudio *RENAAL*, y mediante un modelo de regresión logística multivariante, crearon un *score* para predecir el riesgo a los 4 años de desarrollar enfermedad renal terminal (*ERT*) en pacientes con diabetes tipo II y nefropatía establecida, a partir de cuatro variables analíticas (albuminuria, hipoalbuminemia, elevación de creatinina y descenso de hemoglobina). Este estudio tuvo limitaciones puesto que todos los pacientes eran diabéticos e hipertensos con tratamiento randomizado a losartan o placebo, lo que hace que el *score* creado no sea ampliamente reproducible en otras poblaciones (Keane 2006). Basándose en este estudio, Johnson *et al.* realizaron un estudio retrospectivo para identificar características clínicas basales predictoras de *ERT*. Estudiaron un total de 9782 pacientes diagnosticados de ERC estadios 3 y 4, y valoraron en ellos características clínicas, demográficas y analíticas que pudieran predecir evolución a *ERT* a los 5 años, concluyendo que tenían un valor predictivo las variables edad, filtrado glomerular, sexo, diabetes, anemia e hipertensión arterial. A partir de estos datos, y mediante un análisis de regresión de Cox, se estableció un modelo de riesgo asignando una serie de puntos en función de las características del paciente, de forma que a mayor puntuación, mayor probabilidad de iniciar tratamiento sustitutivo de la función renal a corto plazo. Este estudio tuvo limitaciones porque no se consideraron factores como la raza o los niveles de albúmina (Johnson-2008).

Tangri *et al.* también desarrollaron un modelo predictivo de evolución a *ERT*. Consideraron pacientes con ERC en estadios 3 a 5 estudiando dos grupos, con 3449 y 4942 pacientes respectivamente, que fueron seguidos por Nefrología en Canadá entre 2001 y 2008 hasta que iniciaron tratamiento sustitutivo de la función renal o fallecieron. En ellos se consideraron las siguientes variables: edad, sexo, filtrado glomerular, albuminuria, calcio, fósforo, bicarbonato sódico y albúmina séricos. Los modelos fueron desarrollados utilizando métodos de análisis de supervivencia clásicos (regresión de Cox). Aquí se determinaron como factores

---

predictivos de enfermedad renal un filtrado glomerular bajo, albuminuria elevada, menor edad, sexo varón, niveles bajos de albúmina, calcio y bicarbonato e hiperfosforemia. Las variables más significativas se integraron en varias ecuaciones con diferente número de variables de las descritas para calcular el riesgo de ERT. Estas ecuaciones se validaron en un grupo de población y están disponibles en una página web para su utilización: <http://www.qxmd.com/Kidney-Failure-Risk-Equation> (Tangri 2011). Recientemente se ha publicado un meta-análisis en el que se ha analizado la validez de estas fórmulas para predecir evolución a ERT a los 2 y 5 años en más de 720.000 pacientes de 31 cohortes y 30 países diferentes, confirmándose su validez (Tangri 2016).

Rucci *et al.*, basándose en una población de más de 2000 pacientes con ERC grados 3 a 5, con al menos cuatro determinaciones de creatinina en un año, desarrollaron un árbol clasificatorio e identificaron factores predictores asociados con el descenso del FG, en función de los cuáles establecieron siete subgrupos de pacientes con tasas específicas de descenso del FG (Rucci 2014).

En 2014, Coresh *et al.* publican un trabajo en el que, basándose en un metaanálisis en el que recogen datos de 1.7 millones de pacientes, establecen un modelo predictivo de desarrollo de ERT en función del FG basal del paciente y del porcentaje de descenso del mismo a lo largo de un periodo de dos años.

También se ha desarrollado un modelo para predicción de progresión de enfermedad renal en población pediátrica, en el que, basándose en medidas antropométricas, FG estimado, nefropatía de base, proteinuria y presencia de hipertensión arterial, se clasifica a los pacientes en tres niveles de riesgo para desarrollo de enfermedad renal terminal (Cerqueira 2014).

El modelo ideal para predecir progresión de la enfermedad renal debe ser seguro, fácil de implementar y generalizable a un gran espectro de pacientes con enfermedad renal crónica en poblaciones independientes.

## 1.7. APRENDIZAJE MÁQUINA

Esta tesis doctoral se enmarca dentro del área del aprendizaje máquina aplicado a ciencias de la salud. El aprendizaje máquina es una rama de la inteligencia artificial que simula el conocimiento humano y sus procesos de clasificación, lo cuál, a diferencia de los métodos estadísticos convencionales (análisis de correlaciones o regresión multivariante) proporciona la oportunidad de aprender de datos pasados y utilizarlo para clasificar nuevos datos. (Alpaydin 2010, Flach 2012). Incluyen todos aquellos algoritmos/técnicas que son una extensión no lineal de métodos estadísticos clásicos. (Alpaydin 2010, Flach 2012).

En los últimos años se ha producido un aumento exponencial de este tipo de trabajos debido a sus buenos resultados. Especialmente indicativo es el número de trabajos aplicando estas técnicas y que están centrados en el área de hemodiálisis (Martin 2009, Martin 2012, Titapiccolo 2012, Ferrario 2010, Bhan 2011), lo que demuestra la potencia de estos métodos.

Las ventajas del uso de estas técnicas frente a otras más clásicas (contraste de hipótesis y modelo lineal multivariante) consiste en que permiten encontrar relaciones de cualquier tipo (lineal o no lineal) entre conjuntos de variables. Además estas técnicas no parten de ninguna suposición (por ejemplo normalidad) en los datos para el establecimiento del modelo. Hay que destacar su potencia y su versatilidad frente a otras aproximaciones clásicas y esta versatilidad es la que hace necesario un control sobre el ajuste de los diferentes modelos. Para comprobar la validez de los modelos predictivos utilizados se usarán técnicas de validación cruzada, *v-fold* y *bootstrap* que son métodos muy conocidos en esta área (Alpaydin 2010). Este tipo de técnicas evitan el sobreajuste del modelo a unos determinados datos siendo incapaz de generalizar .

Hay que destacar que, en la aplicación de este tipo de técnicas, no hay referencias sobre el número de patrones a utilizar para un determinado objetivo como sí ocurre en los contrastes de hipótesis clásicos (normalidad, t-test, etc). La ausencia de expresiones para obtener el tamaño de muestra se debe a la no linealidad y versatilidad de estos modelos frente a las aproximaciones estadísticas.

---

Dentro de las técnicas de aprendizaje máquina se incluyen los mapas autoorganizados, métodos de selección automática de variables (Boruta), Naive-Bayes y TAN y árboles de decisión. Todas estas técnicas se emplean en esta tesis para el análisis de los datos.





## **METODOLOGÍA**

---



## 2. METODOLOGÍA

### 2.1. DISEÑO DEL ESTUDIO

Se ha realizado un estudio retrospectivo sobre la población de pacientes seguidos en la Unidad de Enfermedad Renal Crónica Avanzada (ERCA) del Hospital Universitari i Politécnic La Fe de Valencia. El control de los pacientes ha sido realizado por el mismo equipo de atención médica y enfermería, y la valoración de los mismos se ha llevado a cabo de forma homogénea en cuanto a la aplicación de las guías de práctica clínica publicadas en la literatura sobre enfermedad renal crónica.

### 2.2. ANÁLISIS DE LA BASE DE DATOS

#### 2.2.1. SELECCIÓN DE LA MUESTRA

Se han analizado 115 pacientes, cuyos criterios de inclusión han sido:

- Tener un filtrado glomerular menor de  $45 \text{ ml/min/1.73 m}^2$  al inicio del seguimiento, es decir, pacientes con enfermedad renal crónica en categorías 3b, 4 y 5 de la clasificación de las guías *KDIGO* de 2012.

- Haber sido controlados en la citada Unidad durante un mínimo de 6 meses, en el período comprendido entre enero de 2002 y diciembre de 2012, y tener un mínimo de tres revisiones clínicas y analíticas realizadas.

El seguimiento de los pacientes se ha iniciado la primera vez que han sido valorados en la consulta y ha finalizado cuando han precisado tratamiento sustitutivo de la función renal en cualquiera de sus modalidades (hemodiálisis, diálisis peritoneal o trasplante renal anticipado), si han fallecido o, en el resto de casos, en la fecha de finalización del estudio (Diciembre/12).

#### 2.2.2. VARIABLES ANALIZADAS

En los 115 pacientes seleccionados se han analizado las variables al inicio del seguimiento y, posteriormente, se han reevaluado como mínimo anualmente y, además, en diferentes visitas realizadas siempre que se ha producido alguna

circunstancia clínica o analítica que haya derivado en alguna modificación terapéutica, o que haya podido influir en la progresión de la enfermedad renal del paciente. Cada visita genera un registro, y en cada registro se analizan todas las variables. De este modo, el número de registros (visitas) es diferente de unos pacientes a otros ya que en los más estables se han recogido los datos de forma anual y en aquellos que han sufrido más cambios, tanto clínicos como analíticos, se han valorado en más ocasiones. Con este criterio de recogida de datos se ha pretendido no perder información relativa a las posibles oscilaciones de la función renal que pueden sufrir los pacientes en el contexto de diferentes efectos adversos clínicos o analíticos, como podría ser una intervención quirúrgica o un episodio de anemización por poner algún ejemplo.

Las diferentes variables que se han considerado en estos pacientes han sido:

**2.2.2.1. Características epidemiológicas**

a) *Raza*

De los 118 pacientes seleccionados en un principio, uno era asiático y dos de raza negra. Dada la escasa representación de otras razas diferentes a la caucásica, se desestimó esta variable y no se incluyeron estos tres pacientes en el estudio.

b) *Sexo*

c) *Edad*

Se ha considerado la edad de los pacientes en la fecha de inclusión.

**2.2.2.2. Variables antropométricas**

a) *Peso*

En kilogramos, en el momento de la inclusión y posteriormente en cada visita registrada.

b) *Talla*

En centímetros en la primera visita.

### c) Índice de masa corporal

En cada visita, calculado a partir de la fórmula:

$$IMC = \frac{\text{peso (kg)}}{\text{talla (m)}^2}$$

### 2.2.2.3. Enfermedad de base

Nefropatía que ha conducido al desarrollo de la enfermedad renal crónica. Las diferentes enfermedades se han agrupado de la siguiente manera:

1. Nefropatía vascular
2. Diabetes mellitus
3. Poliquistosis renal del adulto
4. Otras (uropatía obstructiva, pérdida de masa renal, hiperfiltración por obesidad)
5. Nefropatías glomerulares
6. Nefropatía intersticial
7. Etiología no filiada

### 2.2.2.4. Comorbilidad

Enfermedades asociadas que, sin ser la causa de la insuficiencia renal, pueden influir en su evolución. Para cualificar la comorbilidad se ha utilizado el índice de comorbilidad de Charlson ajustado a la edad. Este índice es un sistema de evaluación de la esperanza de vida a los 10 años, en función de la edad y de las patologías concomitantes. Consta de la edad y de 19 ítems que, si están presentes, influyen en la esperanza de vida del paciente en concreto. Aunque su objetivo es predecir la supervivencia, se ha empleado con otros fines, como valorar comorbilidad, cálculo de costos asociados a enfermedad entre otros (Charlson 2008).

### **2.2.2.5. Hipertensión arterial**

Se ha considerado que los pacientes presentaban hipertensión arterial si tomaban medicación antihipertensiva, o bien si presentaban cifras de TA sistólica mayor de 140 mmHg y/o de tensión arterial diastólica por encima de 90 mmHg medidas en la consulta (Mancia 2013).

Asimismo, se han dividido los pacientes en tres grupos en función del tratamiento antihipertensivo administrado:

- a) No precisan tratamiento.
- b) Hipertensión arterial controlada con menos de 3 fármacos (es decir, se consiguen cifras menores de 140/90 mmHg con el tratamiento administrado).
- c) Pacientes con hipertensión refractaria (Mancia 2013), o sea, aquellos en los que no se han conseguido unas cifras de tensión arterial inferiores a 140/90 mmHg a pesar de llevar tratamiento con tres o más fármacos, siendo uno de ellos un diurético.

Los pacientes recibían tratamiento antihipertensivo con uno ó más de los siguientes grupos farmacológicos:

- a) Antagonistas del calcio
- b) IECAs y/o ARA II
- c) Alfa-bloqueantes
- d) Beta-bloqueantes
- e) Diuréticos

Tanto el uso de IECAs/ARA como el de diuréticos se ha considerado además de forma individual por su potencial repercusión sobre la función renal.

### 2.2.2.6. *Anemia*

Para valorar la anemia se han analizado:

- Cifras de hemoglobina en sangre.
- Estudio de metabolismo férrico mediante determinación de índice de saturación de transferrina y niveles de ferritina en sangre.

- Tratamiento de la anemia con agentes eritropoyéticos. En función de esta variable se han dividido los pacientes en cuatro grupos:

0: no llevan este tratamiento.

1: epoetina alfa o beta.

2: metoxi-polietilenglicol epoetina beta (CERA).

3: darbepoetina.

- Dosis de agentes eritropoyéticos: Se ha considerado la dosis total mensual en UI/kg/mes. Para unificar las dosis se ha establecido la siguiente equivalencia (Vega 2014):

1mcg darbepoetina: 200 unidades de epoetina

1 mcg CERA: 240 unidades de epoetina

- Tratamiento con hierro: En función de este se han dividido los pacientes en tres grupos:

0: sin tratamiento.

1: tratados con hierro oral.

2: tratados con hierro endovenoso.

**2.2.2.7. Hipertrofia ventricular izquierda**

Se ha considerado como presente o ausente en función de criterios ecocardiográficos y/o electrocardiográficos. Las ecografías han sido realizadas por el servicio de Cardiología del hospital La Fe y, en función de los informes emitidos, se ha clasificado como presente o ausente. Para los criterios electrocardiográficos se ha considerado un índice de Sokolov mayor de 35 mm.

**2.2.2.8. Insuficiencia cardíaca**

Se ha considerado la presencia de insuficiencia cardíaca teniendo en cuenta criterios ecocardiográficos y criterios clínicos, empleando para estos últimos la clasificación de la NYHA (*The Criteria Committee of the New York Heart Association*). Hemos considerado cuatro situaciones diferentes:

- 0: disfunción ventricular ecocardiografía.
- 1: asintomático o grado 1.
- 2: insuficiencia cardíaca grados 2-3.
- 3: insuficiencia cardíaca grado 4.

**2.2.2.9. Hábito tabáquico**

En función de este se han clasificado los pacientes en tres grupos:

- 0: sin historia de tabaquismo.
- 1: fumadores.
- 2: exfumadores.



### **2.2.2.10. Dislipemia**

Se han clasificado los pacientes en dos grupos en función del tratamiento para la dislipemia:

0: sin tratamiento.

1: tratados con estatinas.

El criterio para indicar tratamiento con estatinas era mantener cifras de colesterol LDL por debajo de 100 mg/dl, o bien como prevención secundaria en pacientes con antecedentes de cardiopatía isquémica. La recogida de datos fue previa a las actuales recomendaciones

Además se han evaluado los valores en sangre de colesterol total, colesterol HDL, colesterol LDL y triglicéridos.

### **2.2.2.11. Diabetes mellitus**

Se ha considerado la existencia de diabetes mellitus (DM) en cualquiera de estas tres situaciones: glucemia basal > 126 mg/dl, postprandial > 200 mg/dl, hemoglobina glicosilada > 6,5 % o empleo de fármacos antidiabéticos. Se ha clasificado a los pacientes en función del tratamiento para la diabetes mellitus.

0: sin tratamiento.

1: asociación de insulina y antidiabéticos orales.

2: insulina

3. antidiabéticos orales.

4. Dieta.

Además se han considerado:

- Niveles basales de glucemia en mg/dl.

- Cifra de hemoglobina glicosilada en %.

**2.2.2.12. Alteraciones del metabolismo óseo-mineral**

Para ello se han recogido, a nivel basal y de forma evolutiva:

- Valores en sangre de hormona paratiroidea (PTH) en pg/ml y de 25 (OH) vitamina D3 en pg/ml.

- Tratamiento con derivados activos de la vitamina D, en función de este tratamiento se han clasificado los pacientes en tres grupos:

0: sin tratamiento.

1: en tratamiento con calcitriol.

2: en tratamiento con paricalcitol.

- Tratamiento con vitamina D nativa, se han dividido los pacientes en dos grupos:

0: sin tratamiento.

1: con tratamiento.

- Fósforo en sangre en mg/dl.

- Calcio en sangre en mg/dl. Se ha ajustado a niveles de albúmina sérica según la fórmula:

$\text{Calcio ajustado (mg/dl)} = \text{Calcio medido (mg/dl)} - \text{albúmina (g/dl)} + 4$

- Tratamiento con cinacalcet:

0: no llevan.

1: sí llevan.

- Tratamiento con fármacos quelantes del fósforo:

0: no llevan.

1: quelantes cálcicos.

2: quelantes no cálcicos.

**2.2.2.13. Acidosis metabólica**

Valorada por:

- Niveles en sangre de bicarbonato en mmol/l.
- Tratamiento con bicarbonato sódico:

0: no tratados.

1: tratados.

- Anion GAP: se ha calculado anión GAP tradicional, ajustado a albúmina y a calcio mediante las fórmulas (Abramowitz 2012, Togawa 2013):

*ANION GAP TRADICIONAL* = Sodio en sangre - (Cloro + bicarbonato) (mEq/l).

*ANION GAP AJUSTADO A ALBÚMINA* = Anion GAP tradicional – (2,5 x albúmina en sangre (g/dl)).

*ANION GAP COMPLETO* = Sodio (mEq/l)- Cloro (mEq/l) - Bicarbonato (mEq/l)- (2,5 x albúmina (g/dl)) + potasio (mEq/l)+ (0,5 x calcio(mg/dl)) - (0,323 x fósforo(mg/dl)) x 1,8.

**2.2.2.14. Tratamiento diurético**

En función de si estaban recibiendo tratamiento diurético los pacientes eran clasificados en dos grupos:

0: no tratados.

1: tratados con furosemida

**2.2.2.15. Tratamiento con IECAs/ARAI**

0: no.

1: IECAs o ARA II.

2: Asociación de ambos.

### 2.2.2.16. Función renal

Para el estudio de la función renal se han evaluado:

- Cifra de urea en sangre en mg/dl.
- Cifra de creatinina en sangre en mg/dl. En el hospital La Fe se utilizan técnicas estandarizadas para determinación de creatinina sérica.

- Filtrado glomerular estimado mediante la fórmula *CKD-EPI* y expresado en ml/min/1.73 m<sup>2</sup>. Aunque esta ha sido la fórmula empleada para los test estadísticos, también se ha calculado mediante fórmula MDRD-4 y MDRD-IDMS y se han comprobado equivalencias.

- Fórmula CKD-EPI para pacientes de raza blanca:

Mujeres

$$\text{Si la creatinina es } < 0.7 \text{ mg/dl} \rightarrow FGe = 144 \times \left( \frac{\text{creatinina}}{0.7} \right)^{-0.329} \times (0.993)^{\text{edad}}$$

$$\text{Si la creatinina es } > 0.7 \text{ mg/dl} \rightarrow FGe = 144 \times \left( \frac{\text{creatinina}}{0.7} \right)^{-1.209} \times (0.993)^{\text{edad}}$$

Hombres

$$\text{Si la creatinina es } < 0.9 \text{ mg/dl} \rightarrow FGe = 141 \times \left( \frac{\text{creatinina}}{0.9} \right)^{-0.411} \times (0.993)^{\text{edad}}$$

$$\text{Si la creatinina es } > 0.9 \text{ mg/dl} \rightarrow FGe = 144 \times \left( \frac{\text{creatinina}}{0.9} \right)^{-1.209} \times (0.993)^{\text{edad}}$$

- Ecuación MDRD-4

$$FGe = 186 \times (\text{creatinina})^{-1.154} \times (\text{edad})^{-0.203} \times 0,742 (\text{si es mujer}) \times 1.21 (\text{si raza negra})$$

- Ecuación MDRD-IDMS

$$FGe = 175 \times (\text{creatinina})^{-1.154} \times (\text{edad})^{-0.203} \times 0,742 (\text{si es mujer}) \times 1.21 (\text{si raza negra})$$

- Pendiente de descenso del *FG* a lo largo de la evolución para cada paciente. Se ajusta un modelo lineal por mínimos cuadrados (Peña 08) y se determina la pendiente de dicho modelo.

- Pendiente de descenso del *FG* entre los dos primeros controles realizados en consulta.

#### **2.2.2.17. Hiperuricemia**

Evaluada mediante:

- Niveles de ácido úrico en sangre en mg/dl.

- Tratamiento con alopurinol:

0: no tratados.

1: tratados.

#### **2.2.2.18. Proteinuria**

Proteínas en muestra de orina de 24 horas medida en g/día.

#### **2.2.2.19. Albúmina**

Valores en sangre en g/dl.

**2.2.2.20. Sodio**

Valores en sangre en mEq/l.

**2.2.2.21. Potasio**

Valores en sangre en mEq/l.

**2.2.2.22. Cloro**

Valores en sangre en mEq/l.

**2.2.2.23. Proteína C Reactiva (PCR)**

Valores en sangre en mg/dl.

### 2.3. ANÁLISIS ESTADÍSTICO

#### 2.3.1. ETAPAS DEL ANÁLISIS

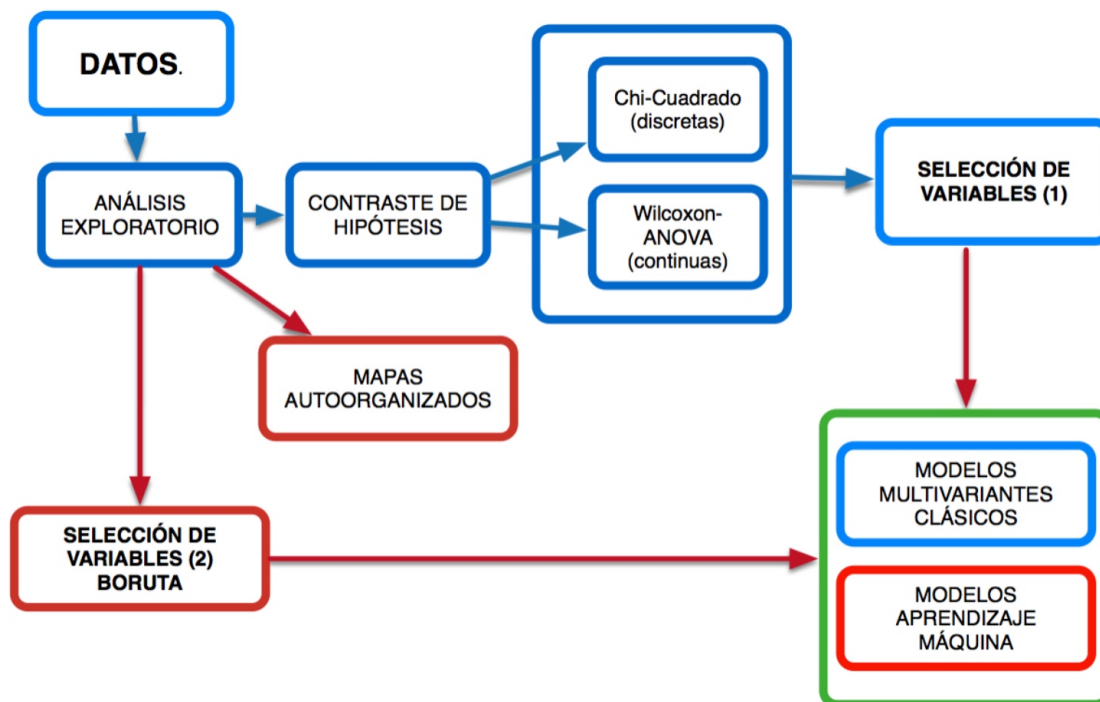


Figura 2.1. Esquema de los métodos planteados en el análisis de los datos de la presente tesis doctoral.

Las etapas en el análisis de datos quedan reflejadas en la figura 2.1, donde aparecen los diferentes métodos que se han aplicado para analizarlos. Se han utilizado colores para distinguir los procedimientos utilizados: en azul aparecen procedimientos estadísticos clásicos y en rojo aparecen métodos de aprendizaje máquina más avanzados. A continuación se describen de forma somera las diferentes etapas para pasar a continuación a describir los métodos de forma más detallada (Riffenburgh 06).

*Análisis exploratorio de datos.* En este caso el objetivo de los métodos es caracterizar los datos recogidos mediante el uso de estadísticos y gráficas. En el caso de los estadísticos se usan frecuencias para variables discretas, y media, desviación estándar, rango intercuartil, etc, para variables continuas. Las figuras escogidas para su representación son diagramas de barras, cajas, sectores, etc.

*Contraste de hipótesis.* El objetivo que se persigue es encontrar diferencias significativas en los dos grupos analizados en el valor de las variables elegidas. Si la variable a analizar es discreta (por ejemplo sexo) se utiliza un Chi-Cuadrado y su variante más exacta (test de Fisher), y si la variable es continua se puede usar un test de Student si las variables siguen una distribución normal o un test de Wilcoxon si la distribución no es normal. De la aplicación de estos contrastes aparecen una serie de variables significativas que serán usadas por los modelos multivariantes tanto clásicos (regresión logística) como avanzados (modelos de aprendizaje máquina).

*Mapas autoorganizados.* Son un instrumento de minería de datos visual de datos multidimensionales. Son capaces de establecer relaciones entre más de tres variables (su gran diferencia con otros métodos gráficos) de forma conjunta.

*Selección de variables mediante técnicas de aprendizaje máquina.* Existen métodos numéricos para establecer qué variables son las importantes a la hora de resolver un problema. La diferencia con respecto a la aplicación de los métodos de contraste de hipótesis clásicos es que estos métodos de selección de variables consideran la interacción conjunta de todas las variables (en un contraste de hipótesis como los definidos en el punto 2 se analizan relaciones univariantes). La selección obtenida en este punto se utilizará también como entradas para los modelos multivariantes (tanto los clásicos como los avanzados).

*Modelos multivariantes.* Tras el análisis univariante de las variables hay que estudiar su comportamiento de forma conjunta. Por una parte se tiene el modelo estadístico clásico de clasificación (modelos lineal, regresión logística) junto con modelos más avanzados, modelos de aprendizaje máquina, que consideran interacciones más complejas.



A continuación se exponen los diferentes métodos utilizados; se describen primero los métodos básicos para, posteriormente, describir los métodos más avanzados en apartados concretos.

### 2.3.2. ANÁLISIS EXPLORATORIO DE DATOS

El paso inicial en todo análisis de datos es caracterizar la muestra. Para ello se utilizan herramientas numéricas y herramientas gráficas. Cada una de estas herramientas es diferente si se quieren caracterizar variables discretas (variables con un bajo número de posibles valores, por ejemplo sexo) o variables continuas (variables con un mayor espectro de posibles valores, por ejemplo peso o talla de un paciente).

Los parámetros numéricos (estadísticos) para caracterizar las variables discretas son las frecuencias de cada posible valor de la variable; de aquí se deriva la moda, que define el valor de la variable que aparece con mayor frecuencia en la muestra.

Para las variables continuas el número de estadísticos posible es mayor; podríamos destacar entre ellos los siguientes (Peña 08):

*Estadísticos de posición.* Estos estadísticos sirven para situar la distribución de los datos en un determinado nivel. Normalmente se usan los que definen la tendencia central, esto es, los que definen los valores centrales de la distribución de datos; los estadísticos más usados para definir la tendencia central son la media aritmética (conocida como valor medio) y la mediana (valor central de una serie de datos ordenados) siendo éste último más robusto cuando se tienen valores extremos muy grandes/pequeños (*outliers*, o datos atípicos).

*Estadísticos de dispersión.* Miden el rango de variación de la variable a analizar; los más usados son la varianza, la desviación estándar y el rango intercuartílico.

En cuanto a los gráficos, para las variables discretas se usan de forma clásica los diagramas de barras y sectores. Para las variables continuas los gráficos

que se utilizan son los histogramas y los diagramas de cajas que dan idea de la distribución de los datos en cada una de las variables analizadas (Peña 08).

En esta tesis se proponen visualizaciones directamente relacionadas con las anteriores pero se han introducido variaciones en los formatos de las gráficas para facilitar la visualización de los datos. Entendemos que ésta se puede considerar como una aportación de la tesis.

Dentro de este apartado de análisis exploratorio de datos se incluye el cálculo de correlaciones lineales entre variables continuas; más concretamente se calcula el coeficiente de correlación lineal que establece la posible relación lineal entre variables. Si aparecen dependencias lineales se pueden desechar variables de cara a establecer un modelo de predicción ya que, en definitiva, lo que indica esta dependencia es que la información contenida en las dos variables altamente correlacionadas es la misma.

### **2.3.3. CONTRASTE DE HIPÓTESIS**

Una vez caracterizados los datos hay que establecer otras relaciones entre las variables además de las posibles dependencias lineales comentadas anteriormente. Para ello se usará lo que se conoce como contraste de hipótesis, que consiste, como su nombre indica, en establecer una hipótesis de partida y, analizando las distribuciones/parámetros estadísticos de las variables, establecer si se cumple, o no, dicha hipótesis. Para ello se determina un parámetro estadístico que se obtiene de los datos y se compara con su valor tabulado si la hipótesis se cumpliera, de tal forma que se tiene una medida de incertidumbre probabilística sobre la hipótesis de partida. En este trabajo se han aplicado diferentes contrastes de hipótesis que se enumeran a continuación (Riffenburgh 06):

*Test de normalidad.* En este caso el objetivo es comprobar si una variable continua sigue una distribución gaussiana o normal. La razón de este test es que muchos procedimientos estadísticos dependen del cumplimiento de esta condición. El test aplicado es el de Kolmogorov-Smirnov.

*Chi-cuadrado.* En este test se comprueba si dos variables discretas son independientes o no. La idea es construir una tabla donde aparecen todos los posibles casos cruzados de las dos variables y compararla con otra que sería aleatoria, esto es, donde todas las casillas contienen el número de datos que se tienen dividido por el número de casillas. Esta comparación es la que permitirá establecer un factor numérico de independencia. El objetivo de este test es, en primer lugar, determinar posibles dependencias entre variables de entrada al modelo para reducir el número de entradas (de la misma manera que se hacía con las correlaciones). En segundo lugar se busca establecer relaciones entre la variable a predecir que es discreta (progresar o no) y el resto de variables discretas de entrada al modelo predictor. De esta manera se tiene una primera “preselección” de entradas adecuadas al modelo.

*Test de comparación de medianas/medias entre grupos.* Este tipo de test se aplica en variables continuas y el objetivo es determinar si existen diferencias estadísticamente significativas en los estadísticos de tendencia central (media/mediana) entre los grupos considerados. Esto es, buscamos encontrar esas diferencias entre los pacientes que progresan y los que no progresan. Dependiendo de si la variable tiene que cumplir una serie de requisitos o no existen métodos paramétricos (*t-test* o *t de Student*, se debe cumplir normalidad) o no paramétricos (*Wilcoxon*). En el primer caso se buscan diferencias significativas en las medias y, en el segundo caso, se buscan diferencias significativas en las medianas de las variables continuas. El objetivo con este test es el mismo que en el caso anterior, encontrar aquellas variables continuas que tienen una mayor relación con la variable discreta que se quiere predecir.

Tras esta etapa se obtiene una selección de variables que serán usadas por los modelos multivariantes tanto clásicos como los avanzados de aprendizaje máquina. Dentro de estos modelos se sitúa el modelo que se explica a continuación, los mapas autoorganizados.

## 2.3.4. MAPAS AUTOORGANIZADOS

### 2.3.4.1. Introducción

Los mapas autoorganizados son un tipo de red neuronal, integrado en el llamado aprendizaje-máquina. Las redes neuronales son una herramienta de tipo computacional, diseñadas para imitar el pensamiento humano, que se caracteriza por ser impreciso pero adaptativo.

Está demostrado que uno de los mejores elementos para extraer conocimiento de los datos es el ojo humano; de aquí se deriva todo un *corpus* de conocimiento conocido como minería de datos visual. El objetivo es representar de forma avanzada los datos para poder extraer el máximo de conocimiento de los datos. Si pensamos en las visualizaciones clásicas (algunas de ella ya comentadas, como son diagramas de sectores y de barras) se tiene el problema del bajo número de variables que se pueden representar de forma conjunta; como mucho podremos representar tres variables de forma conjunta pero no podemos representar más usando ese tipo de técnicas.

Lo que se plantea con el uso de los *Mapas Autoorganizados (SOM)*, o *Mapas de Kohonen* en honor a su inventor, es la posibilidad de *representar de forma conjunta datos multidimensionales* (Kohonen 01). Este tipo de representación consiste en una especie de panel de abejas donde se van a ir metiendo pacientes en cada celda. Si dos pacientes se parecen mucho (esto significa que tienen los valores de las variables analizadas muy parecidas) se meten en la misma celda. Conforme disminuye este parecido los pacientes se alojan en diferentes celdas aumentando la distancia entre pacientes dentro del panel conforme la desigualdad entre ellos aumenta. El objetivo es agruparlos en ese panel (se realiza un *clustering*).

Este procedimiento de acercar-alejar pacientes dentro de un plano bidimensional según parecido se aplica en otras técnicas estadísticas clásicas ya utilizadas en clínica como es el Escalado Multidimensional (MDS, Multidimensional Scaling); la novedad introducida con los SOM consiste en la visualización de las variables usando este modelo y que queda explicada en la siguiente sección mediante un ejemplo.

**2.3.4.2. Ejemplo**

A modo de ejemplo se presenta un caso ficticio; se han generado datos de forma artificial de 1500 pacientes definidos por 5 variables: peso, altura, índice de masa corporal, colesterol y triglicéridos. Con herramientas clásicas no podemos representar las variables de forma conjunta y para todos los patrones; sin embargo, con el SOM sí que vamos a poderlo hacer. En primer lugar se colocan los diferentes pacientes en el panel comentado anteriormente; si dos pacientes se parecen mucho caen en la misma celda; si el parecido es menor aparecen en diferentes celdas pero cercanas alejándose en el panel conforme ese parecido disminuye. Se obtiene entonces una primera gráfica, *mapa de ganadoras*, donde se representa el número de pacientes que caen en cada celda del panel.

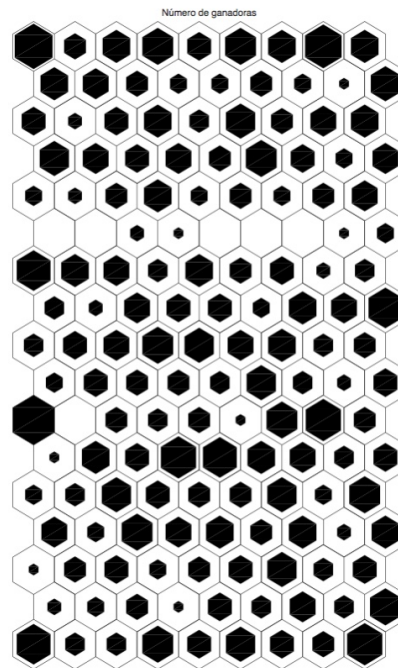
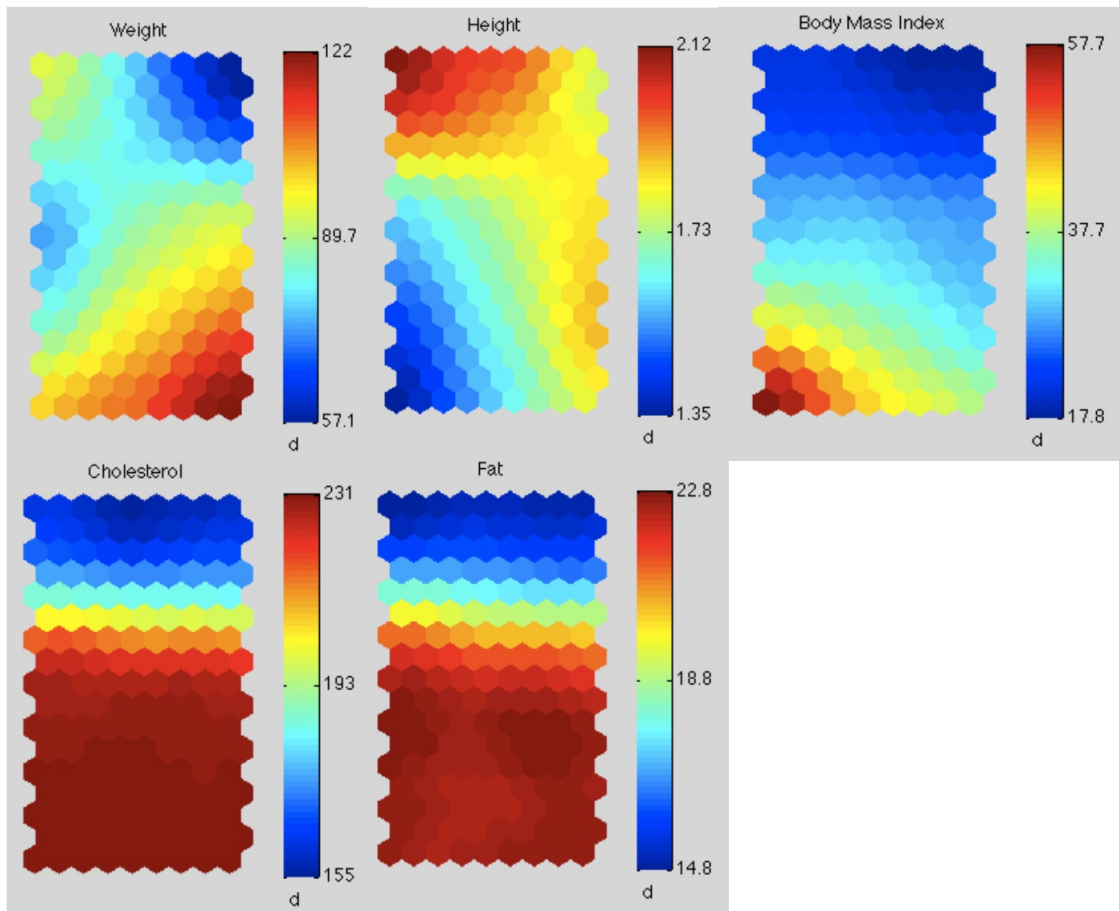


Figura 2.2. Mapa de ganadoras del ejemplo propuesto.

En la figura 2.2 se rellena cada celda del panel de forma proporcional al número de pacientes que caen en dicha celda. De esta forma se tiene una representación visual de la distribución de los pacientes en el mapa.

El paso final sería representar ese panel tantas veces como variables se tienen. En nuestro caso se llega a la figura 2.3. El análisis de esta figura consiste en fijarse en la misma zona espacial del SOM y ver qué valores toman las diferentes variables en esa zona.



*Figura 2.3. Mapa de componentes del ejemplo propuesto.*

En la figura 2.3 aparecen zonas que proporcionan información sobre diferentes comportamientos en los datos:

*Zona inferior izquierda.* Pacientes con un IMC muy alto; se corresponden con pacientes con un alto peso (en torno a los 90-95 kgs, misma zona, esquina superior izquierda pero para la variable peso) y con una menor altura (esos mismos pacientes tienen una altura entre 1.35 y 1.5 metros. Además tienen altos índices de triglicéridos y colesterol.

*Zona inferior derecha.* En este caso se tienen IMC bajos y se corresponden a pacientes con muy poco peso (entre 55 y 60 kg) y más altos (de 1.75 a 1.85 metros de altura). Estos pacientes presentan valores bajos de colesterol y de triglicéridos.

Se podría continuar analizando todas las zonas del mapa que nos interesaran. Además, de la interrelación entre variables del SOM se pueden extraer correlaciones entre variables y, lo más importante, no tienen por qué ser lineales. Así, si se observa el mapa correspondiente a triglicéridos y el correspondiente a colesterol son iguales, lo que indica que las dos variables dan la misma información y que, por tanto, se puede eliminar una de ellas en análisis posteriores.

### **2.3.5. SELECCIÓN DE VARIABLES. BORUTA**

A la hora de desarrollar un modelo basado en datos del tipo que sea, predictivo/clasificadorio, el número de variables de entrada al modelo es un factor muy importante ya que impacta directamente en (Lee 07):

*Cantidad de patrones necesarios para ajustar el modelo.* Es bien conocido en estadística que el número de patrones necesarios para ajustar un modelo crece conforme se aumenta el número de variables. Existen varias reglas más o menos heurísticas en las que teniendo el número de variables y el modelo a ajustar se definen el número de patrones necesarios para ajustar el modelo. En todas esas reglas el número de patrones necesario sigue una relación creciente con el número de variables que, en algunos casos, es una relación de crecimiento mayor que una lineal.

Maldición de la dimensionalidad. Relacionado con la anterior; conforme el número de variables aumenta, el número de posibilidades crece de manera exponencial. Por ejemplo, pensemos dos variables discretas que tienen 4 posibles valores, las posibilidades son 16; si son 3 nos vamos a 64 posibilidades, si tuviera 10 variables discretas tendríamos  $4^{10}$  ¡más de un millón de posibilidades!. No puedo tener todos los casos posibles y tengo que inferirlos con los modelos pero esto se hace más difícil conforme crece el número de posibilidades.

*Calidad del modelo.* Se ha demostrado que conforme el número de variables crece no sólo es más difícil ajustar el procedimiento (crece la carga computacional y la complejidad de los algoritmos) sino que, además, los modelos suelen ser peores.

Por todo lo comentado se hace necesario reducir el número de variables antes de desarrollar el modelo predictivo. Según la figura 2.4 se ha seguido un camino que podríamos definir como clásico al tiempo que se plantea una aproximación más moderna usando un procedimiento automático de selección de características conocido como *Boruta* cuyo esquema se muestra en la figura 2.4 (Kursa 10).

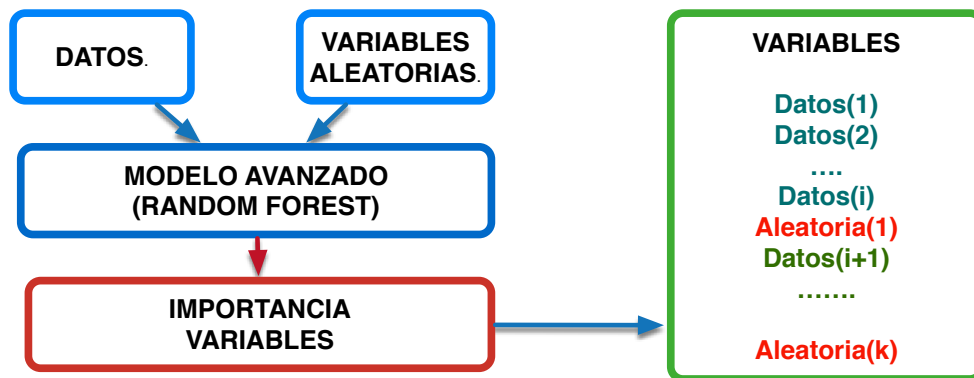


Figura 2.4. Esquema del procedimiento Boruta.



El procedimiento Boruta tiene las siguientes etapas:

1. Se construye un modelo predictivo avanzado (Random Forest) que se alimenta con dos tipos diferentes de entradas: a) las variables que se tienen sobre el problema a resolver y b) variables aleatorias que no tienen nada que ver con el problema a resolver.

2. Se analiza la importancia de cada una de las variables del modelo. Existen muchas formas diferentes de hacer esto; una de las formas de hacerlo es ir anulando de forma secuencial las entradas y ver cómo afecta su eliminación a la capacidad predictiva del modelo. Cuanto más baje esta capacidad más importante es la variable.

3. Como última etapa se ordenan las entradas del modelo y se seleccionan las entradas de acuerdo al siguiente criterio; *toda aquella entrada del conjunto de datos a analizar que sea menos importante que una variable aleatoria será eliminada*. Este hecho es lógico, si una entrada es menos importante que una variable aleatoria (ruido) significa que no hay que considerarla en el modelo.

### 2.3.6. MODELOS MULTIVARIANTES

#### 2.3.6.1. Regresión logística

Es un modelo estadístico clásico, incluido en la categoría de los llamados modelos lineales generalizados. Estos modelos vienen representados por los siguientes pares de ecuaciones (Peña-08):

$$u = w_0 + w_1 \cdot x_1 + \dots w_L \cdot x_L$$

$$y = g(u)$$

Donde  $g(u)$  es una función no lineal. En el caso de una regresión logística se tiene lo siguiente:

$$y = \frac{1}{1 + e^{-u}}$$

Este valor de salida,  $y$ , da la probabilidad de tener un determinado suceso dado el conjunto de variables de entrada  $x$  siendo  $w$  los parámetros del modelo a ajustar según los datos que se tienen. En nuestro caso la variable a predecir es si el paciente será progresor rápido o no, dadas unas determinadas variables de entrada.

### 2.3.6.2. Naïve Bayes

Es un modelo probabilístico (red bayesiana) basado en la hipótesis de que las variables de entrada son condicionalmente independientes de la variable a predecir (Koller 09). A pesar de que esta hipótesis es muy restrictiva, este clasificador es uno de los más usados mostrando resultados similares a modelos más avanzados (como redes neuronales) en determinados problemas. El gran problema de este modelo aparece cuando no se cumplen las condiciones de independencia condicional y se tienen variables de entrada con una fuerte dependencia entre ellas. El modelo de Naïve Bayes queda representado por el esquema de la figura 2.5.

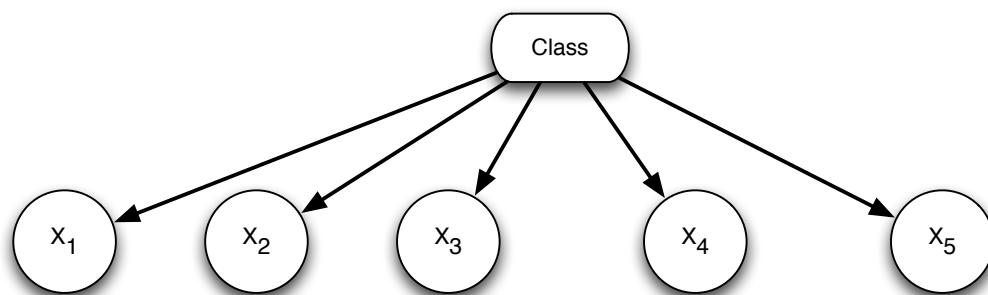


Figura 2.5. Esquema del modelo Naïve Bayes.

Aquí la probabilidad de obtener  $y$  (progresor o no progresor) condicionada a las variables de entrada queda factorizada en un producto de variables condicionadas más simples:

$$P(y|x_1, x_2, \dots, x_L) = P(y|x_1) \cdot P(y|x_2) \cdot \dots \cdot P(y|x_L)$$

### 2.3.6.3. TAN (Tree Augmented Network)

Esta estructura permite que cada variable de entrada se relacione con otras; de esta manera se tiene un árbol probabilístico entre variables de entrada para acabar estableciendo una relación con la variable de salida de forma similar a un Naïve Bayes (Koller 09). La figura 2.6 muestra el esquema de este modelo:

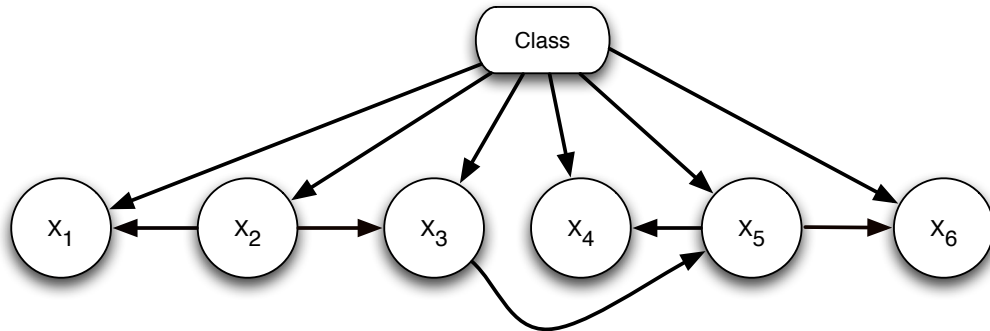


Figura 2.6. Esquema del TAN (Tree Augmented Network).

**2.3.7. MEDIDAS DE RENDIMIENTO DE LOS MODELOS****2.3.7.1. Matriz de confusión**

Los modelos que se van a desarrollar son modelos de clasificación binaria, esto es, su objetivo es discriminar entre dos clases de pacientes, los progresores y los no progresores. En este tipo de problemas utilizar como medida de rendimiento el acierto en la clasificación no suele ser lo óptimo por varias razones:

Si las clases están desbalanceadas, esto es, si hay una clase con un gran número de patrones el acierto no indica un rendimiento óptimo. Por ejemplo, si tenemos una proporción del 99% de una clase decir que tenemos un clasificador que acierta el 99% nos podría inducir a error (es un sistema que siempre tiene como salida la clase predominante).

En determinados problemas el error en una clase es más peligroso que el error en otra. Por ejemplo, en un problema de detección de cáncer es más peligroso decirle a un paciente que no lo padece que al contrario. En el primer caso el paciente se marcha sin saber que lo tiene y en el segundo se realizarán otras pruebas que determinarán la presencia de dicho cáncer.

Para resolver estos dos problemas se plantea reconsiderar el problema de clasificación binaria de la forma expresada por la siguiente tabla:

		MODELO	
		<i>PROBLEMAS</i>	<i>SANOS</i>
REALIDAD	<i>PROBLEMAS</i>	<b>A</b>	<b>B</b>
	<i>SANOS</i>	<b>C</b>	<b>D</b>

En esta tabla se representa en las filas lo que se tiene en la realidad y, en las columnas lo que predice el modelo. Es decir, los pacientes con problemas serían la suma de A + B. A serían los pacientes que tienen el problema a estudio y el modelo lo acierta y B serían aquellos que estando enfermos el modelo dice que no. La población sana vendría representada por C+D, donde C son pacientes sanos que el modelo identifica como enfermos y D son los realmente sanos.

Así se definen los siguientes parámetros (Koller 09):

*Sensibilidad.* Proporción de pacientes con problemas que el sistema acierta sobre el total de pacientes con problemas que se tienen; su expresión viene dada por:

$$\text{Sensibilidad} = \frac{A}{A + B}$$

*Especificidad.* Proporción de pacientes sanos que el sistema acierta sobre el total de pacientes sanos que se tienen; su expresión viene dada por:

$$\text{Especificidad} = \frac{D}{C + D}$$

*Valor Predictivo Positivo.* Proporción de pacientes enfermos que el sistema acierta sobre el total de pacientes enfermos que el sistema predice; su expresión viene dado por:

$$\text{Valor Predictivo Positivo} = \frac{A}{A + C}$$

*Valor Predictivo Negativo.* Proporción de pacientes sanos que el sistema acierta sobre el total de pacientes sanos que el sistema predice; su expresión viene dado por:

$$\text{Valor Predictivo Negativo} = \frac{D}{B + D}$$

### 2.3.7.2. Área bajo la curva ROC (AUC)

Los parámetros anteriores se determinan si el modelo tiene una salida discreta (0/1). En el caso de tener una salida continua se plantea otro tipo de medida relacionada con lo que se conoce como curva ROC (Krzanowski 09) . La forma de construirla es la siguiente:

- Se fija un umbral para discretizar las salidas del modelo; por ejemplo se considera 0.5; si la salida es mayor que 0.5 la consideramos igual a 1 y si la salida es menor o igual a 0.5 la salida la tomamos como 0. De esta forma se tiene una salida discreta.

- Una vez discretizada se puede formar la matriz de confusión y calcular la sensibilidad/especificidad. Estos dos parámetros se corresponden a un punto en la curva ROC.

- La forma de realizarla es variar el umbral de 0 a 1 de tal forma que se van obteniendo diferentes puntos de la curva ROC. Por convenio se representa sensibilidad frente a 1-especificidad.

La figura 2.7 muestra un ejemplo donde se aplican dos modelos a un determinado problema.

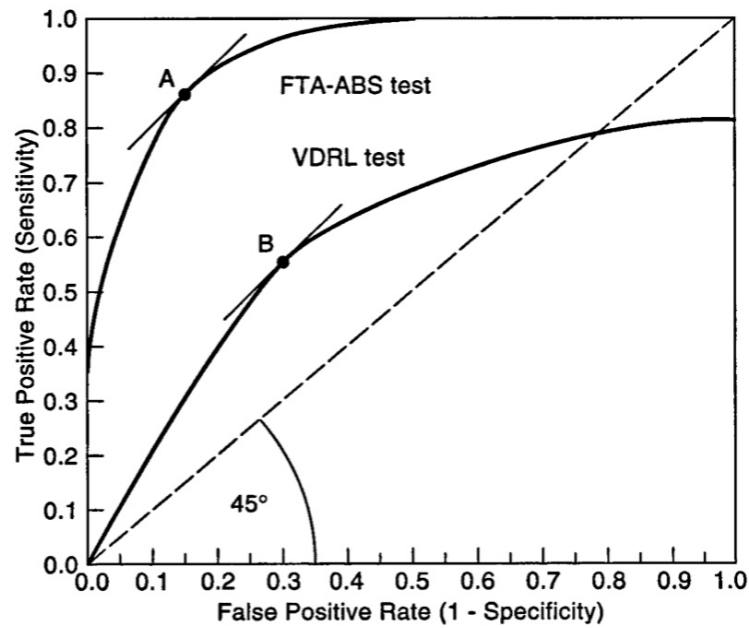


Figura 2.7. Esquema de dos curvas ROC.

El parámetro que deriva de la curva ROC y que da una medida del rendimiento del clasificador es el área bajo la curva (AUC). Este parámetro está entre 0 y 1. Normalmente el clasificador de referencia es un modelo aleatorio que tiene una AUC de 0.5. Un clasificador perfecto tendría una AUC de 1. Hay que destacar que esta medida no tiene los problemas comentados anteriormente y que sí tiene el acierto.

### 2.3.8. ÁRBOLES DE DECISIÓN

El árbol de decisión es uno de los modelos analíticos más usados en Ciencias de la Salud porque es muy visual y es un elemento de clasificación/modelización no lineal. Consiste en un conjunto encadenado de condiciones sobre las variables que se tienen de tal forma que, al final de ese conjunto se obtiene una decisión final, en nuestro caso el paciente tendrá una



progresión rápida o no. Un esquema típico de árbol de decisión es el representado en la figura 2.8 (se da un ejemplo simulado).

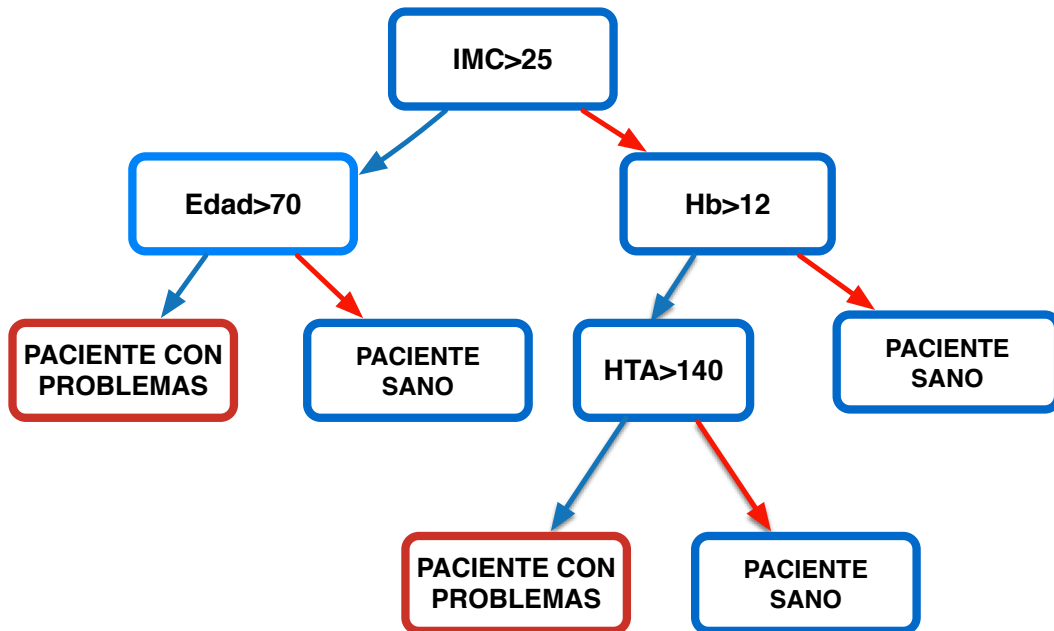


Figura 2.8. Esquema de un árbol de decisión (problema simulado).

En cada uno de los nodos de la figura 2.8 aparece una determinada condición de tal manera que, si se cumple seguimos la flecha azul y, si no se cumple, seguimos la roja. Al final de todo tenemos una decisión, paciente sano o enfermo.

Existen algoritmos para crear de forma automática un árbol de decisión; en estos algoritmos se plantea una medida de complejidad conocida como *Ganancia de Información* y, a partir de esa medida para las diferentes variables se va construyendo el árbol (Hastie 09).



### **3. RESULTADOS Y DESARROLLO ARGUMENTAL**

---



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

#### 3.1. RESULTADOS

##### 3.1.1. DESCRIPCIÓN DE LA MUESTRA

Se han analizado datos de 115 pacientes seguidos en la consulta de ERCA durante un tiempo medio de 3,44 años con un total de 786 registros recogidos. El número de registros de cada paciente ha variado en función del tiempo de seguimiento y de la concurrencia de eventos que hayan condicionado visitas extra, oscilando entre 3 y 16. En la *Figura 3.1* se agrupan todos los pacientes representándose sexo, tiempo de seguimiento y resultado al final del mismo. En cada registro se han recogido datos clínicos y analíticos de los pacientes. Aunque de entrada puede parecer menos uniforme que una recogida de datos periódica, de este modo se refleja mejor la evolución real del proceso.

El tiempo de seguimiento medio de los pacientes ha sido de  $3.44 \pm 2.64$  años, con un rango de 0.65 a 11.4 años. De los 115 pacientes, 53 son hombres y 62 mujeres, todos ellos de raza caucásica. La edad media de los pacientes estudiados ha sido de  $66.28 \pm 15.54$  años. Si los dividimos por sexo, las mujeres tienen una edad media de  $68.34 \pm 15.4$  años y los hombres de  $63.87 \pm 15.4$  (*Figura 3.2*). En cuanto a la evolución de los pacientes, al final del estudio 81 pacientes (42 hombres y 39 mujeres) continuaban en seguimiento, sin necesidad de tratamiento sustitutivo de la función renal, y 30 pacientes (26 % del total) iniciaron tratamiento sustitutivo de la función renal: 8 (6 hombres y 2 mujeres) se habían incluido en programa de diálisis peritoneal, 21 (3 hombres y 18 mujeres) habían iniciado hemodiálisis y un hombre recibió un trasplante renal anticipado de donante vivo. Cuatro pacientes (1 hombre y 3 mujeres) fallecieron durante el seguimiento (*Figura 3.3*).

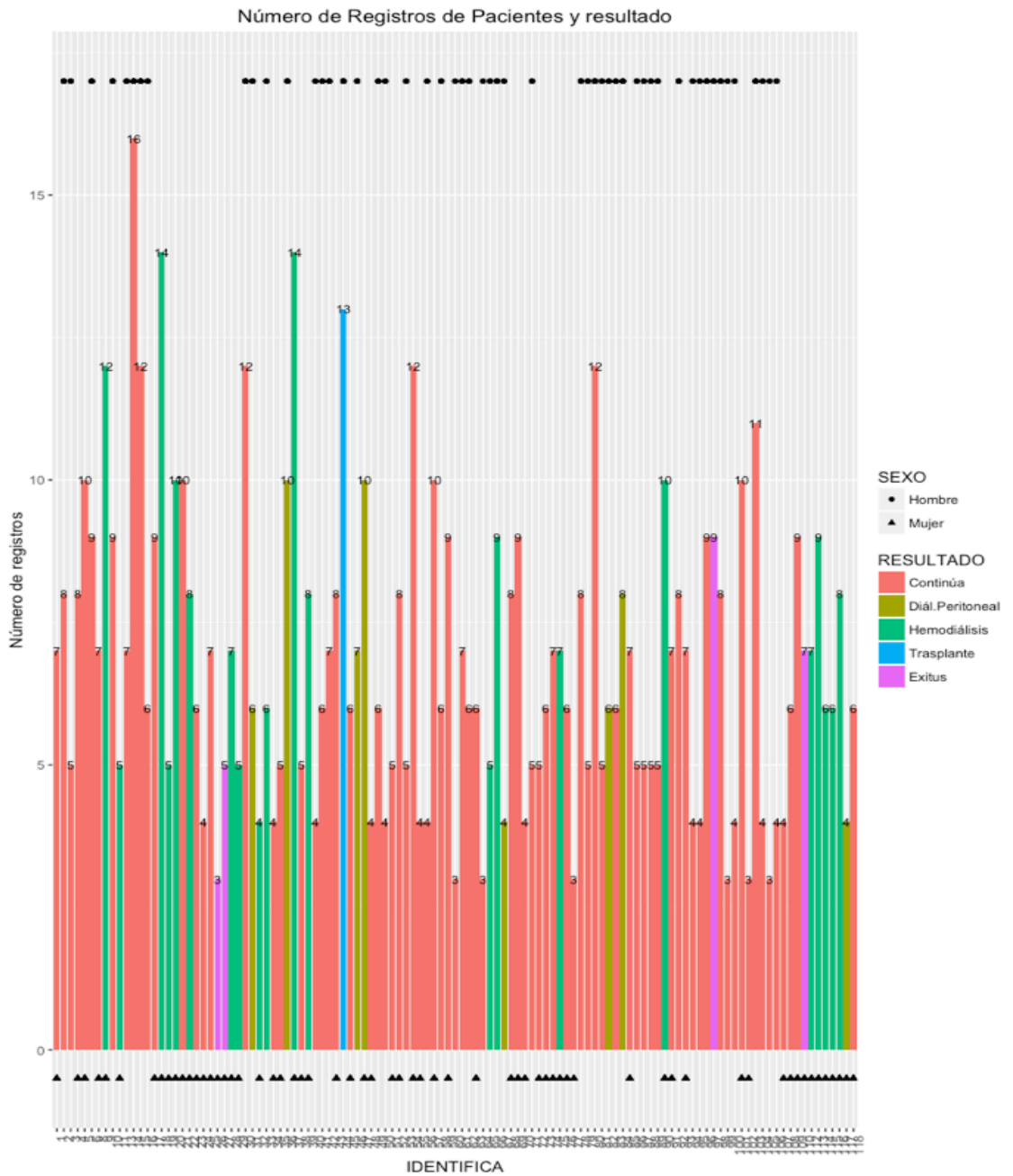


Figura 3.1. Población estudiada y número de visitas realizadas por paciente. En el gráfico se representa además el sexo del paciente y su situación al finalizar el estudio.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

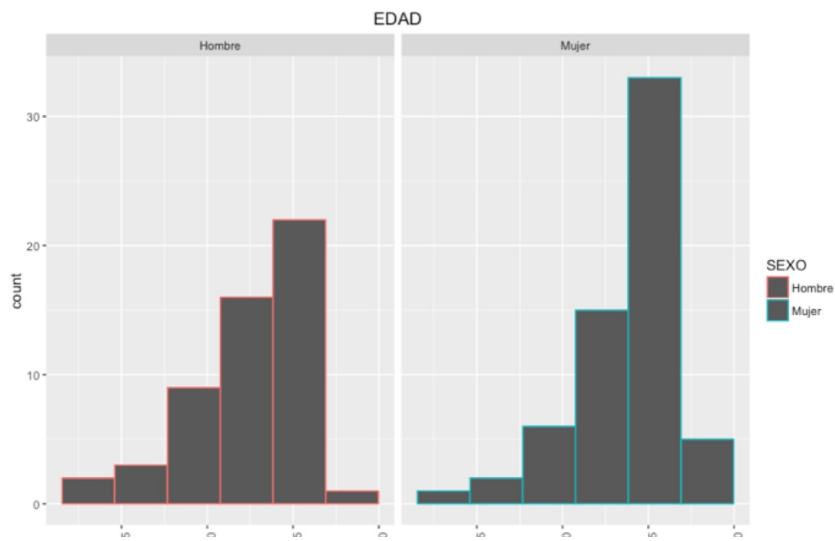


Figura 3.2. Histograma que representa la edad y sexo de los pacientes estudiado. En el eje de ordenadas se representa la edad y en el eje de abscisas el número absoluto de pacientes.

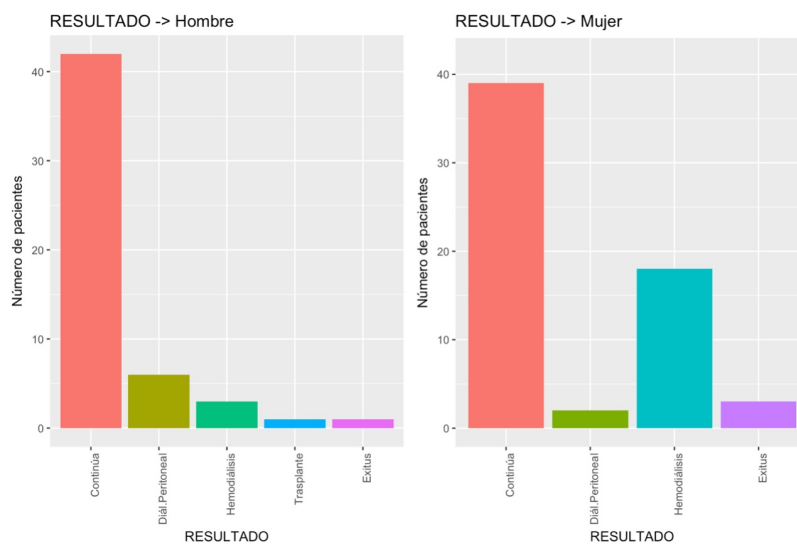


Figura 3.3. Resultado de los pacientes al finalizar el estudio.

### 3.1.2. VARIABLES ANALIZADAS

#### 3.1.2.1. Enfermedad de base

La causa más frecuente de enfermedad renal en los pacientes estudiados ha sido la nefropatía vascular secundaria a hipertensión arterial, seguida de la nefropatía diabética. En la *tabla 3.1* y *figura 3.4* se representan la nefropatía de base de los pacientes en global y también distribuida por sexos.

Nefropatía vascular	50
Diabetes Mellitus	15
Otras	14
Glomerulonefritis	13
Nefropatía intersticial	7
Poliquistosis renal	8
Etiología no filiada	8

Tabla 3.1. Nefropatía de base.

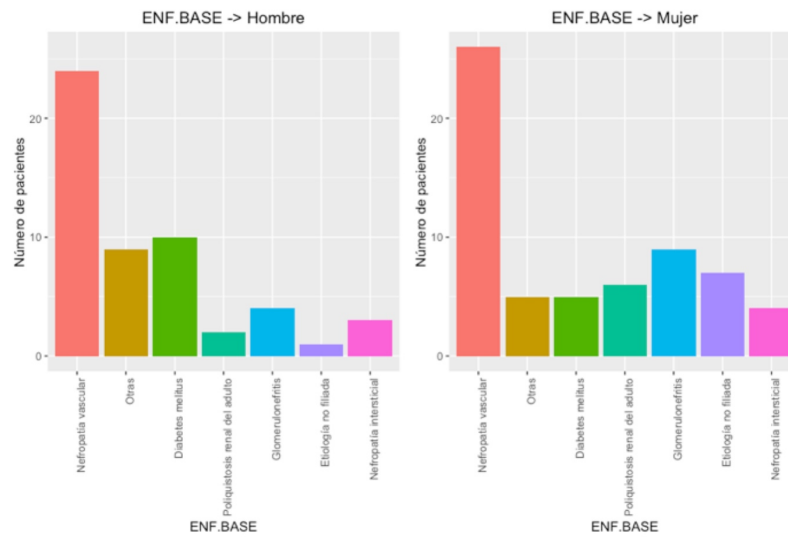


Figura 3.4. Nefropatía de base: a) global, b) distribución por sexos.



3.1.2.2. **Función renal**

En cuanto a la función renal de este grupo de pacientes, la cifra media de urea durante el seguimiento ha sido de  $118.6 \pm 35.8$  mg/dl (109.6 en hombres y 125.4 mg/dl en mujeres) y la de creatinina de  $3.43 \pm 1.24$  mg/dl (3.51 en varones y 3.35 en mujeres). La media del filtrado glomerular durante el seguimiento estimado por *CKD-EPI* ha sido de  $17.3 \pm 6.4$  ml/min/1.73 m<sup>2</sup> y por *MDRD-IDMS* de  $17.5 \pm 6.6$  ml/min/1.73 m<sup>2</sup>, siendo algo más elevado en los varones (19.3 versus 15.6 ml/min/1.73 m<sup>2</sup> por *CKD-EPI*). La media de los valores del FG al inicio del seguimiento estimado por *CKD-EPI* ha sido de  $18.7 \pm 5.5$  ml/min/1.73 m<sup>2</sup>. Al inicio del estudio 11 pacientes presentaban ERC grado 3b, 74 pacientes grado 4 y 30 pacientes grado 5.

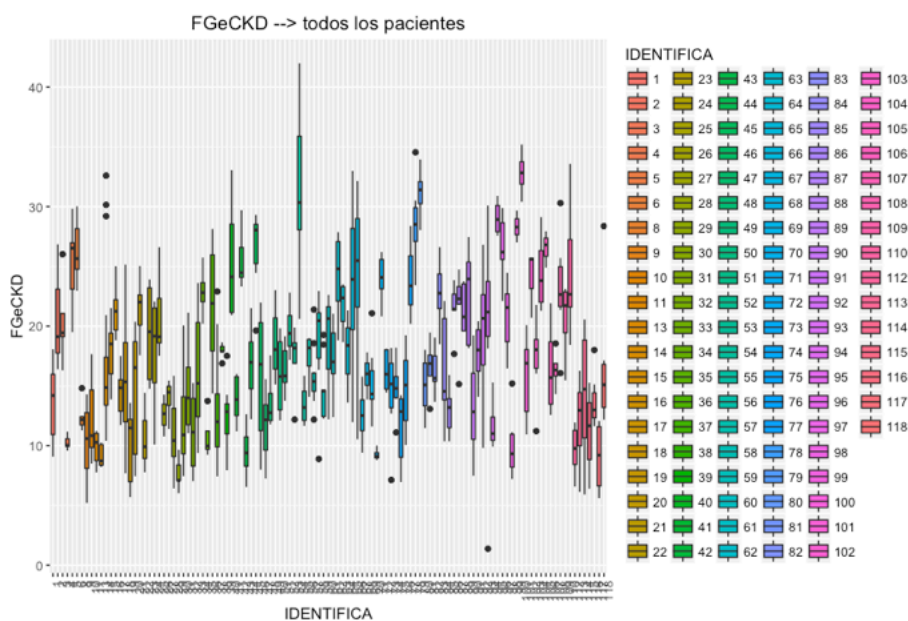


Figura 3.5. Representación de los diferentes diagramas de cajas (boxplot) de FG estimado por *CKD-EPI* para cada uno de los pacientes estudiados.

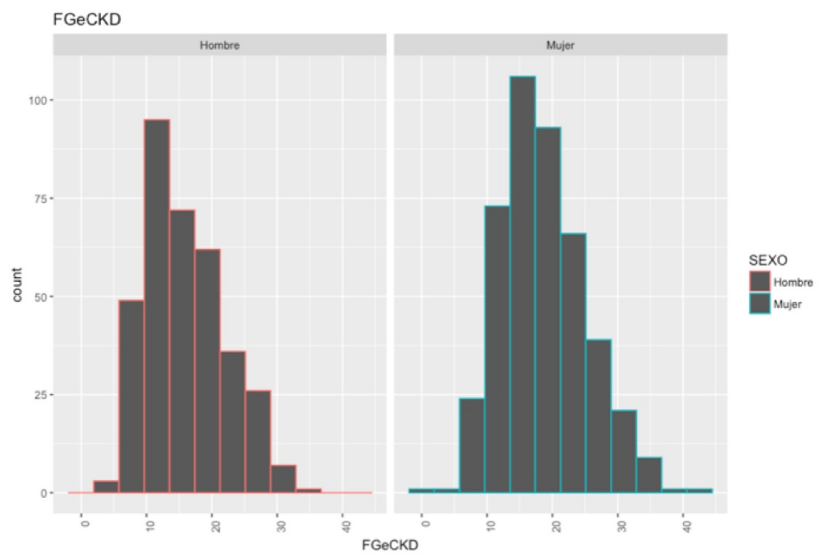


Figura 3.6. Histograma que representa la distribución de los valores del FG estimado por sexo. En el eje de ordenadas se representa el valor del FG estimado por CKD-EPI y en el eje de abscisas el número de registros realizados.

Aunque se ha utilizado el valor del FG estimado por CKD-EPI como referencia, de acuerdo con las recomendaciones actuales (Levey 2009), también se ha calculado por MDRD-4 y MDRD-IDMS, comprobando que los resultados son superponibles, como se aprecia en la Figura 3.7.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

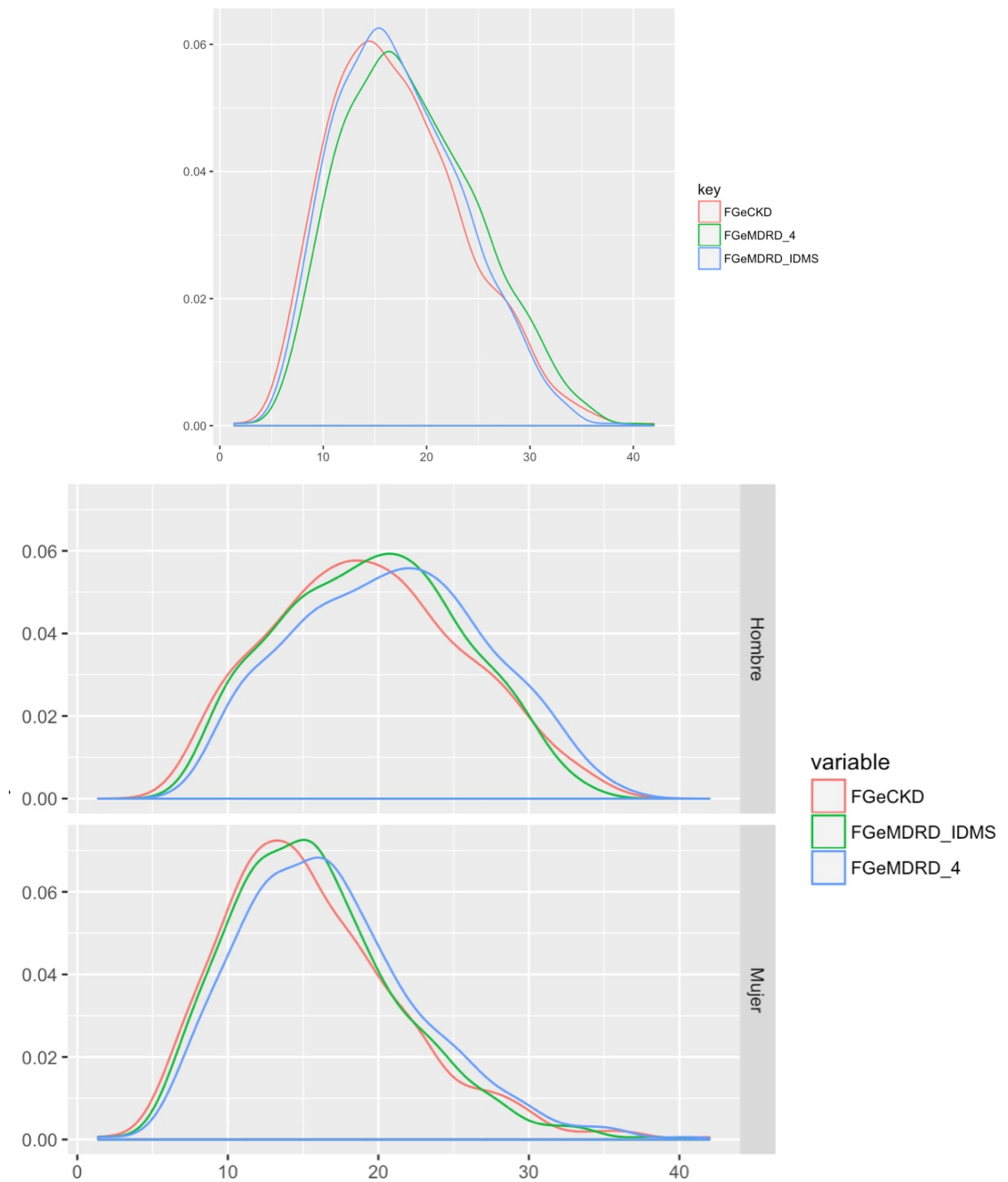


Figura 3.7. Gráficos comparativos del filtrado glomerular estimado por las diferentes fórmulas empleadas.

### 3.1.2.3. Comorbilidad

Se ha utilizado el índice de comorbilidad de Charlson ajustado a edad para cuantificar y homogeneizar otras patologías que han presentado estos pacientes. Los pacientes analizados han presentado una puntuación media en dicho índice de  $5.23 \pm 1.98$  (Figura 3.8).

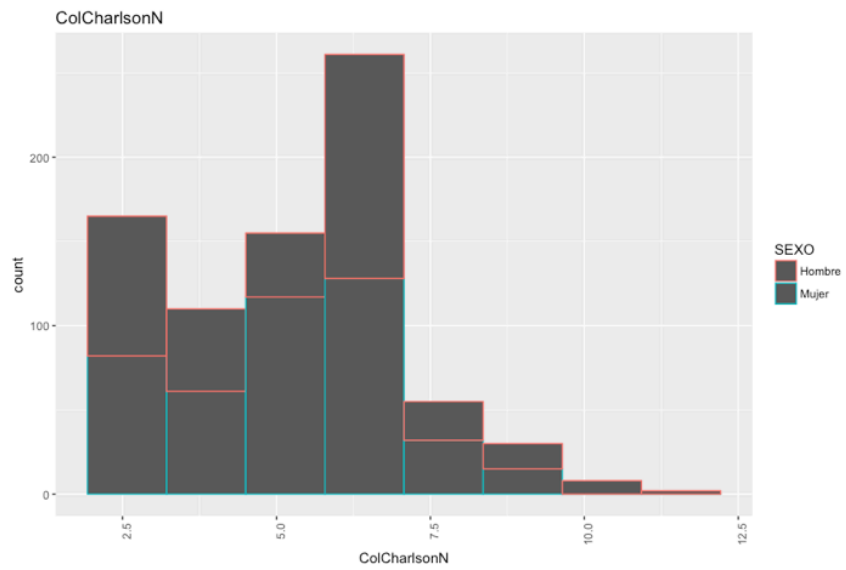
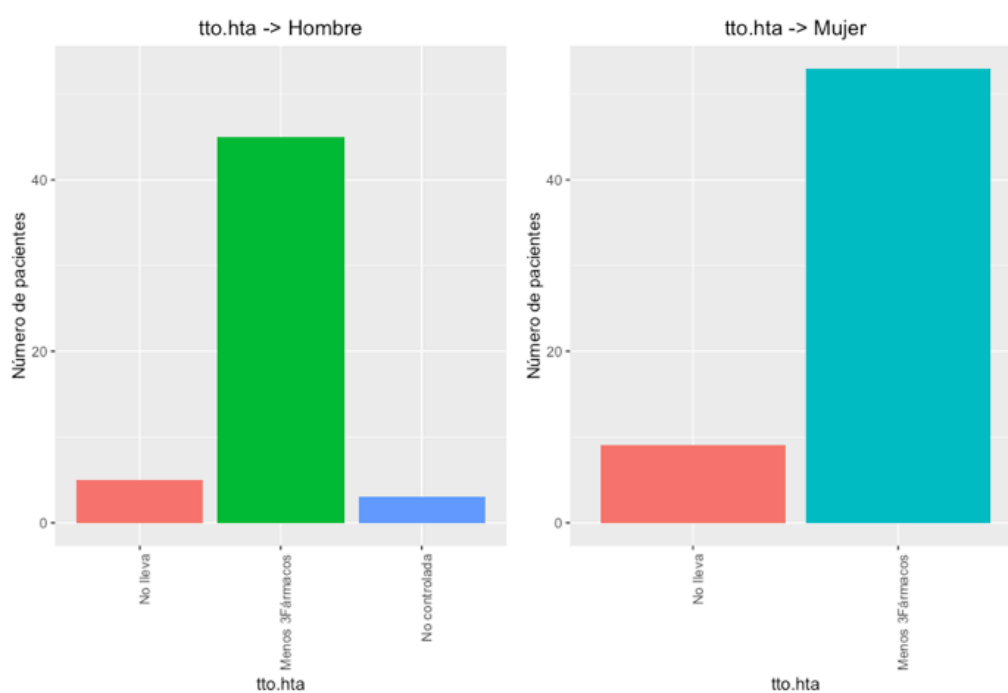


Figura 3.8. Representación del índice de comorbilidad de Charlson y distribución por sexos.

**3.1.2.4. Hipertensión arterial**

De los 115 pacientes analizados, 103 (89.5%) presentaron hipertensión arterial en algún momento del seguimiento: 49 de los 53 hombres (92.4%) y 54 de las 62 mujeres (87.1%). Entre los pacientes hipertensos, la mayoría se controlaban con tres o menos fármacos antihipertensivos, y solo 3 pacientes, todos ellos varones, lo que suponía un 2.9% del total de los pacientes hipertensos, presentaban una hipertensión arterial refractaria al tratamiento (*Figura 3.9*).



*Figura 3.9. Representación de la hipertensión arterial y su tratamiento.*

### 3.1.2.5. Tabaco

76 de los 115 pacientes no eran fumadores, lo cual supone un 66.1 % de pacientes sin hábito tabáquico. Entre los no fumadores la mayoría eran mujeres (67.1% versus 32.9%). Diez hombres y cuatro mujeres eran fumadores activos durante el seguimiento y 25 (18 hombres y 7 mujeres) habían abandonado el tabaco al inicio del estudio.

### 3.1.2.6. Obesidad

Los pacientes presentaron un IMC medio de  $28 \pm 4.87 \text{ kg/m}^2$  ( $27.8 \text{ kg/m}^2$  en varones y  $28.3 \text{ kg/m}^2$  en mujeres) (Figura 3.10).

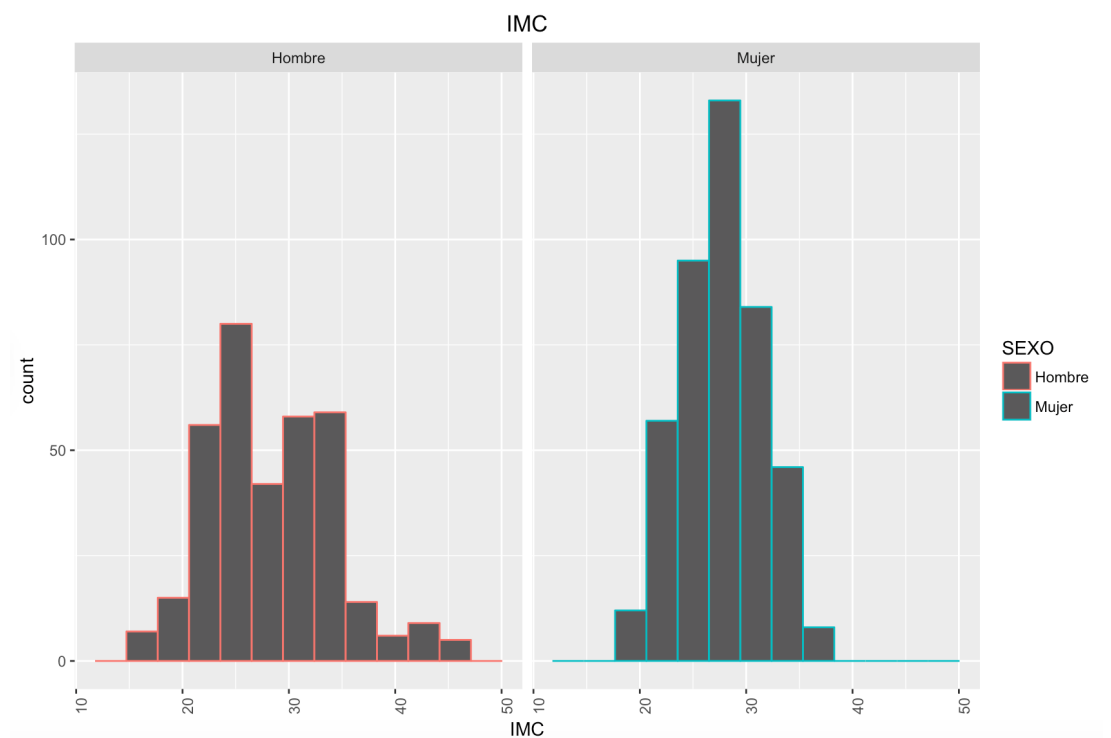


Figura 3.10. Representación del índice de masa corporal.

**3.1.2.7. Anemia**

La cifra de Hb media mantenida durante el seguimiento ha sido de  $12.4 \pm 1.35$  g/dl ( $12.7$  en varones y  $12.3$  en mujeres) (Figura 3.11). En cuanto a parámetros de metabolismo férrico, los pacientes han presentado un *IST* medio de  $26.15 \pm 9.78\%$  y una ferritina de  $192.9 \pm 190$  ng/ml.

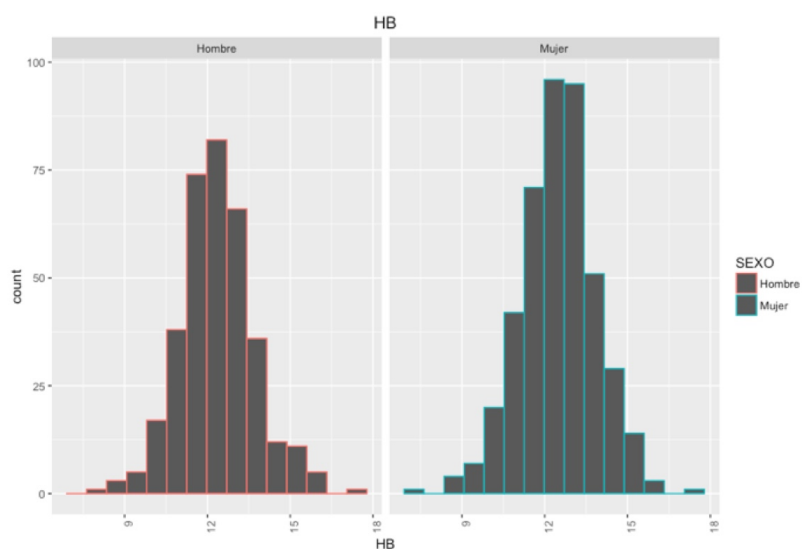
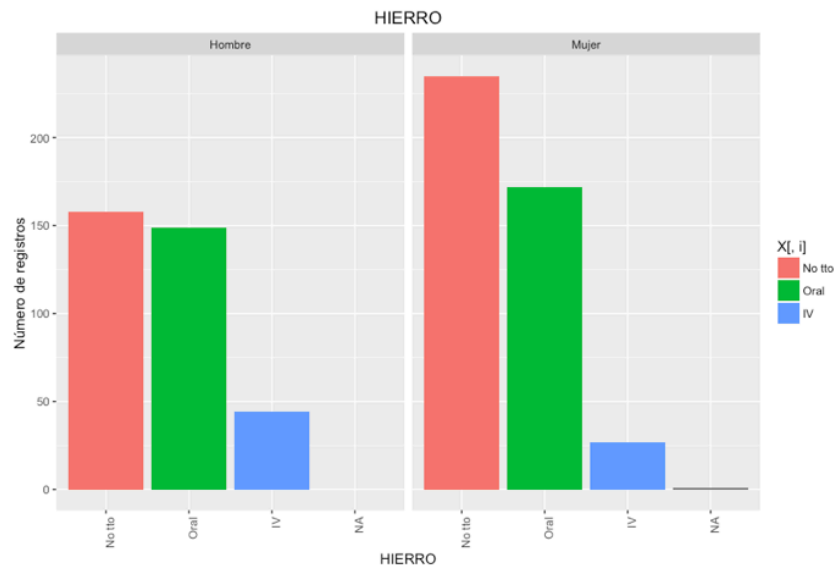


Figura 3.11. Histogramas que muestran la distribución de las cifras de hemoglobina por sexos.

En cuanto a los aportes de hierro para el manejo de la anemia, en 392 registros los pacientes no recibieron este tratamiento, y entre los que lo llevaban, recibían suplementos por vía oral mayoritariamente (321 registros). Se administró hierro endovenoso en 71 registros, y en un control no se recogió este dato (*Figura 3.12*).



*Figura 3.12. Tratamiento con suplementos de hierro.*



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

De los 115 pacientes estudiados, 105 han precisado tratamiento con derivados eritropoyéticos en algún momento de la evolución, es decir un 91,5 % de los pacientes recibieron este tratamiento.

En un 46.3% de los registros realizados los pacientes recibían tratamiento con derivados eritropoyéticos. De ellos un 75.3% habían llevado tratamiento con epoetina alfa o beta, un 5.4% con darbepoetina y un 19.3 % con CERA (Figura 3.13).

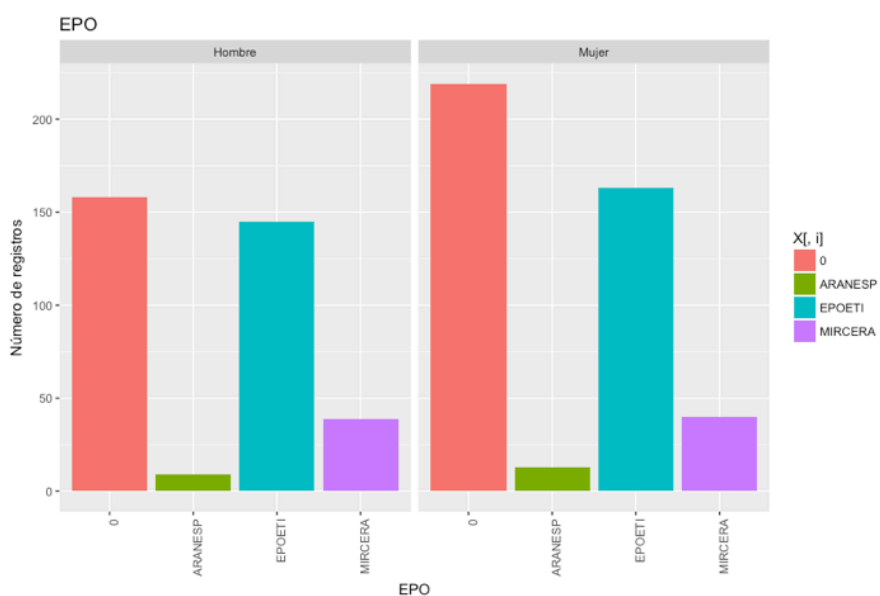
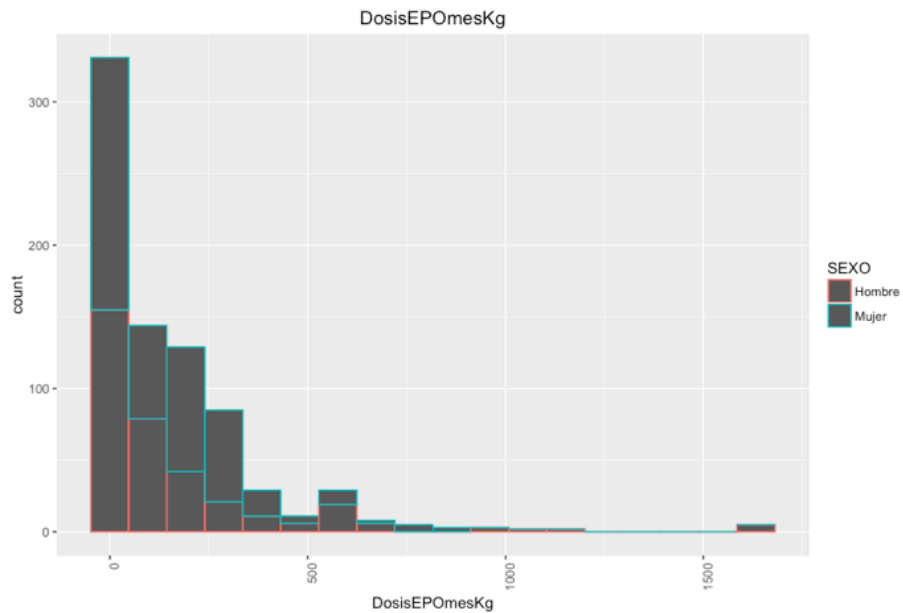


Figura 3.13. Tratamiento con derivados eritropoyéticos.

Para unificar los datos se han realizado equivalencias entre los diferentes preparados farmacológicos de forma que la dosis del tratamiento administrado se ha expresado como unidades de epoetina/kg/mes.

La dosis media de agentes eritropoyéticos que han recibido los pacientes ha sido de  $268.7 \pm 227.3$  UI/kg/mes. En la *figura 3.14* se representa la dosis total de epoetina recibida por los pacientes.



*Figura 3.14. Histograma que representa la dosis recibida de agentes eritropoyéticos. En el eje de ordenadas se representa la dosis en UI/Kg/mes y en el de abscisas los registros realizados.*

3.1.2.8. *Dislipemia*

Los pacientes presentaron una cifra media de colesterol total de  $180 \pm 43.4$  mg/dl, siendo algo más elevada en mujeres (188.6 versus 170.2 mg/dl). La cifra media de colesterol HDL ha sido de  $51.2 \pm 14.93$  mg/dl (46.5 mg/dl en hombres y 55.8 mg/dl en mujeres), y la de colesterol LDL de  $102 \pm 35.08$  mg/dl (97.7 mg/dl en hombres y 107.9 mg/dl en mujeres).

Los niveles de triglicéridos en sangre han sido de  $136.8 \pm 66.42$  mg/dl (145.1 mg/dl en hombres y 130.3 mg/dl en mujeres).

En cuanto al tratamiento, un 49.1 % de los pacientes recibieron tratamiento con estatinas en algún momento, siendo las mujeres las más tratadas (56.9 % versus 43.1%).

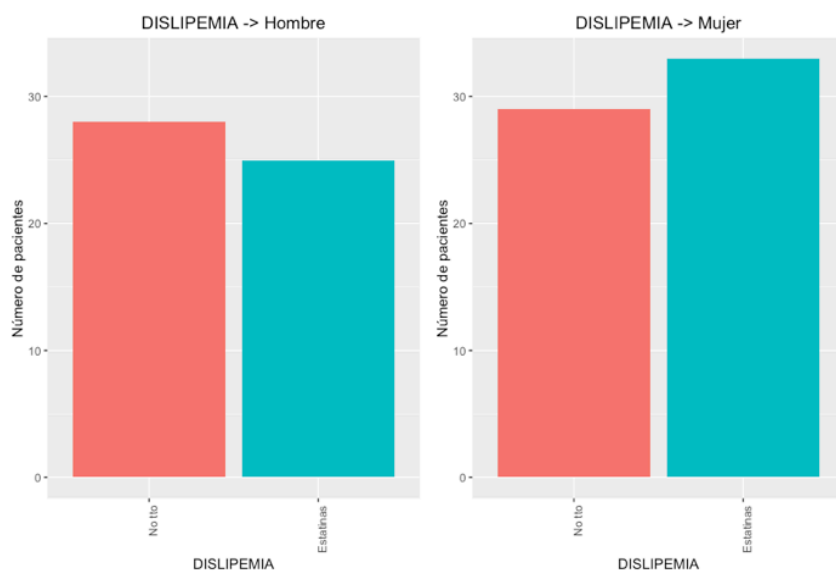


Figura 3.15. *Tratamiento con estatinas.*

**3.1.2.9. Diabetes**

En la muestra estudiada 37 pacientes (33.17 % del total) fueron diagnosticados de diabetes mellitus en algún momento de la evolución. En la figura 3.16 se ha representado la distribución del tratamiento antidiabético que presentaron los pacientes a lo largo del seguimiento. En 514 de los 786 registros realizados (73%), los pacientes no precisaban tratamiento alguno. En el resto se emplearon distintas opciones terapéuticas (un 40.56 % recibieron tratamiento con antidiabéticos orales (ADO), un 37.79 % llevaron tratamiento con insulina y un 8.49 con asociación de insulina y ADO. En un 14.5 % fueron tratados únicamente con medidas higiénico-dietéticas. La cifra media de glucosa obtenida fue de  $104.61 \pm 29.73$  mg/dl.

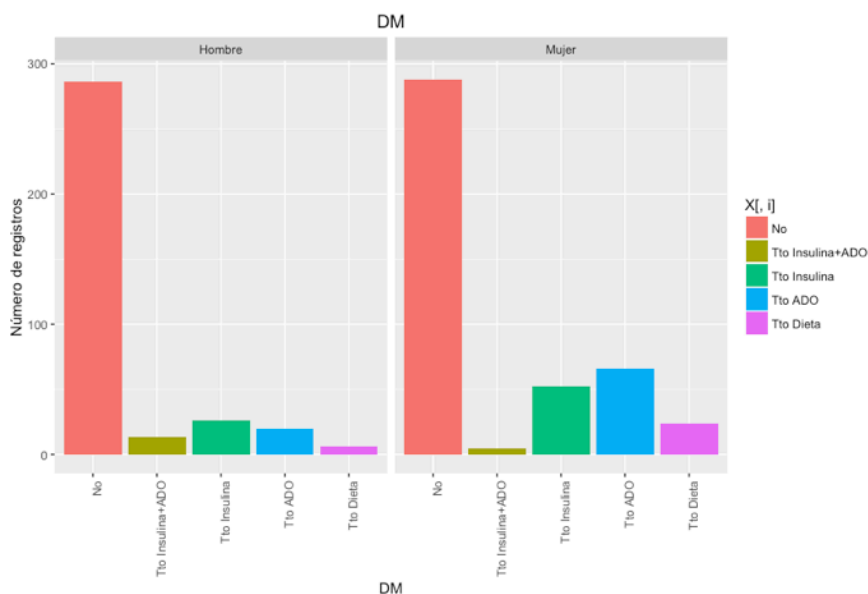


Figura 3.16. Tratamiento de diabetes mellitus.

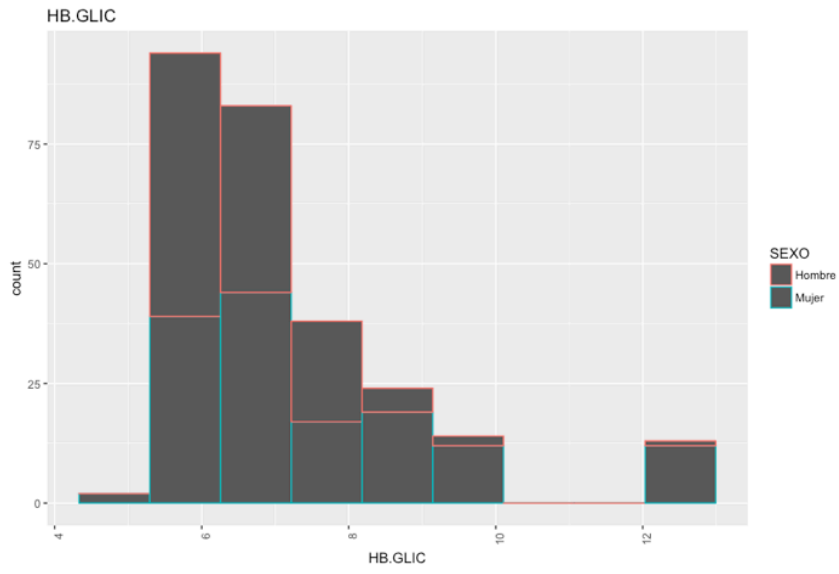


Figura 3.17. Hemoglobina glicosilada (%)

**3.1.2.10. Acidosis**

Los pacientes presentaron una cifra media de bicarbonato en sangre de  $23.6 \pm 3.42$  mmol/l. Recibían tratamiento con suplementos de bicarbonato sódico en un 41 % de los registros.

En la *Figura 3.18* se aprecia la distribución de los valores de anión GAP en los registros realizados. Los valores medios obtenidos mediante las diferentes fórmulas han sido: A GAP tradicional  $11,6 \pm 4,25$ , A GAP ajustado a albúmina  $1,38 \pm 4,25$ , A GAP completo  $8,67 \pm 4,14$ ).

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

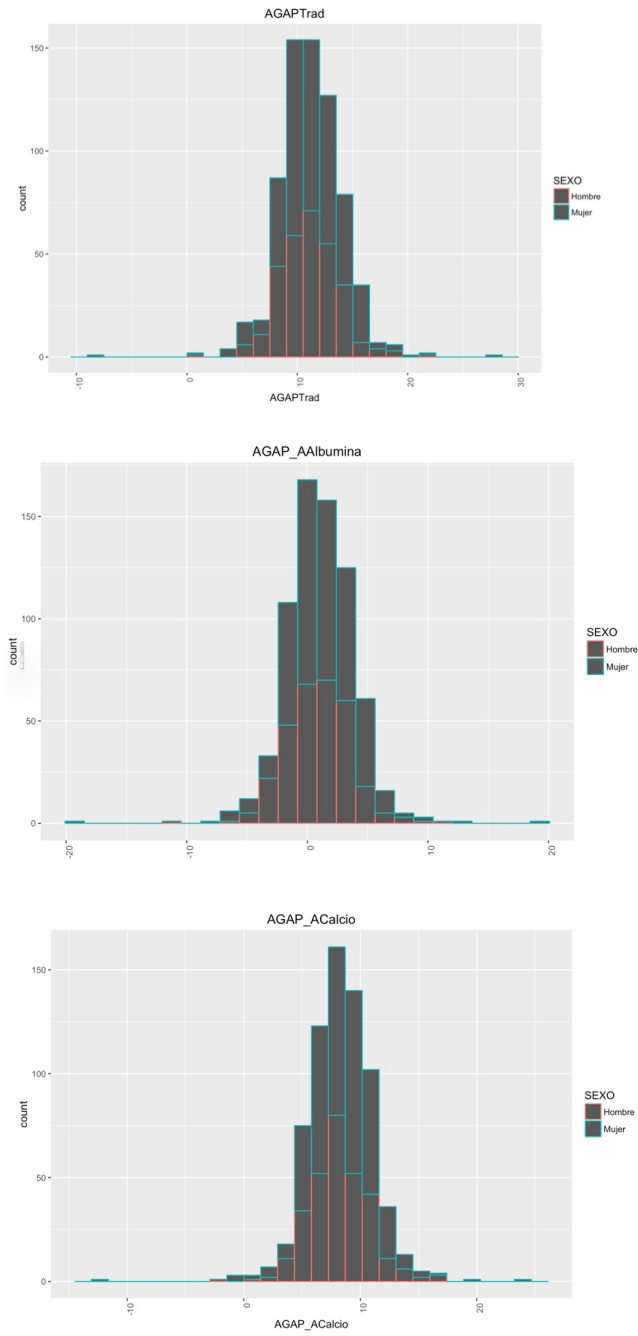
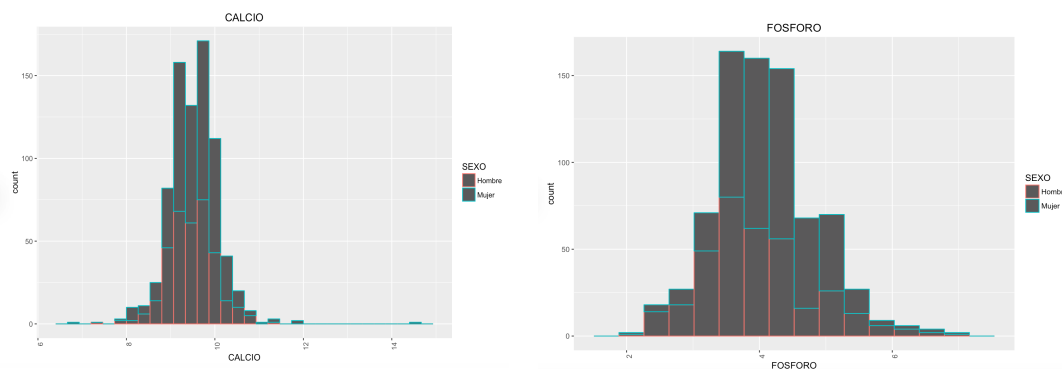


Figura 3.18. Distribución de Anion GAP (A GAP tradicional  $11,6 \pm 4,25$ , A GAP ajustado a albúmina  $1,38 \pm 4,25$ , A GAP completo  $8,67 \pm 4,14$ ).

### 3.1.2.11. Metabolismo óseo-mineral

Los niveles medios de calcio han sido de  $9.5 \pm 0.6$  mg/dl, con un rango de 6.9 a 14.5 mg/dl. En los hombres la calcemia media era de 9.5 y en mujeres de 9.5 mg/dl (*Figura 3.19*).

La media de los niveles de fósforo registrados ha sido de  $4.07 \pm 0.7$  mg/dl, con un rango de 2 a 6.9 mg/dl. En mujeres, la media del fósforo ha sido de 4.1 y en hombres de 3.9 mg/dl. A pesar de que el valor medio de la fosfemia encontrado está dentro del rango deseado, en el diagrama adjunto en el que vienen representados todos los pacientes, se puede apreciar que ha habido una importante variabilidad individual en los niveles de fósforo obtenidos (*Figura 3.20*).



*Figura 3.19. Niveles de calcio y fósforo. En el eje de ordenadas se representan los niveles en mg/dl y en el de abscisas en número de registros.*



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

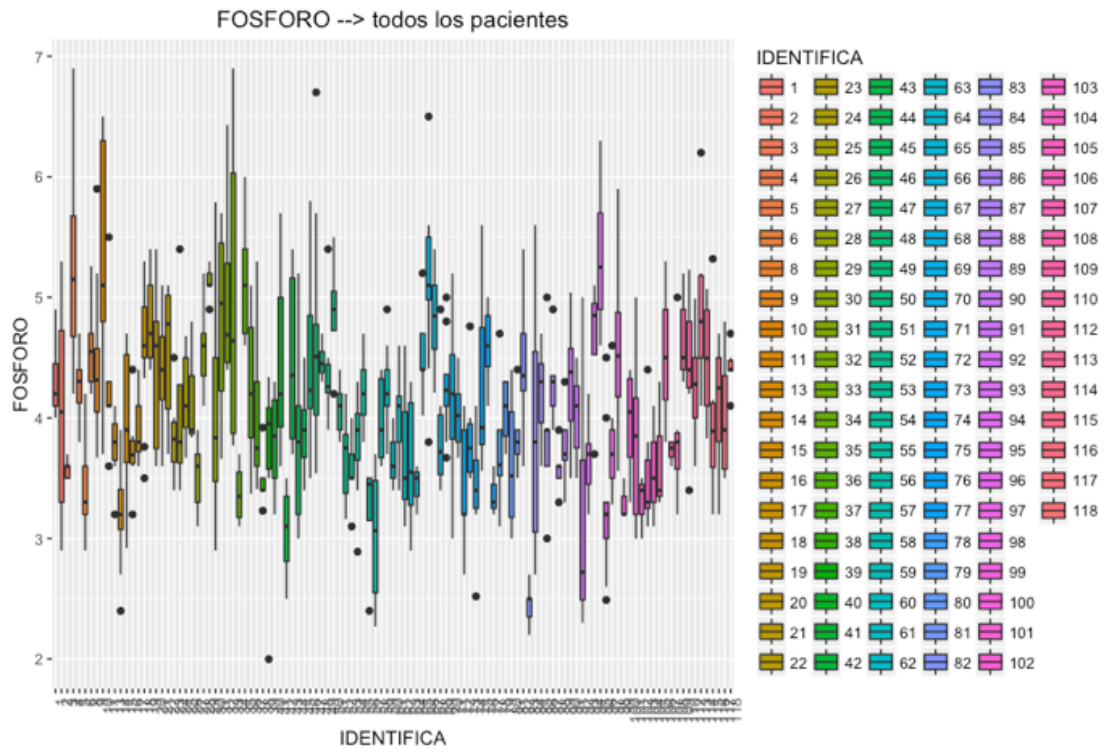


Figura 3.20. Diagramas de cajas que representan el rango de fosfatemia para cada paciente.

Se han obtenido unos niveles de vitamina D3 de  $23.8 \pm 14.7$  pg/ml con una amplia variabilidad (4.5 – 105 pg/ml). Los niveles medios de PTH han sido de  $191 \pm 125$  pg/ml (6-920 pg/ml), y también han mostrado una variabilidad importante como se observa en la figura 3.21.

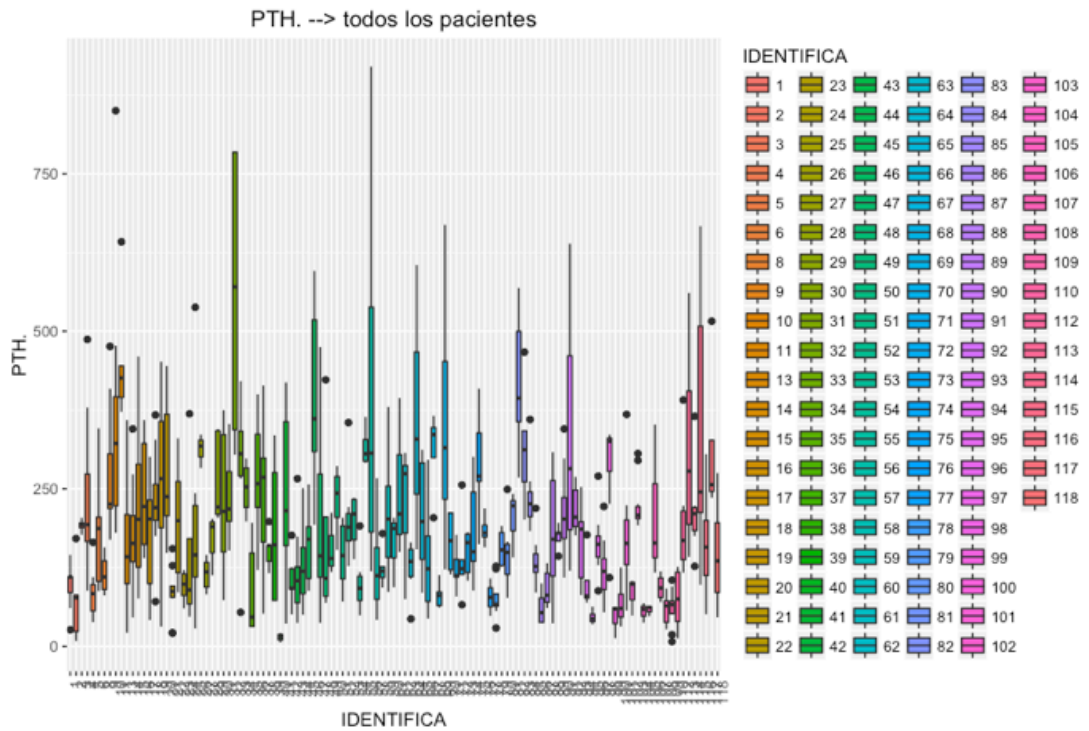


Figura 3.21. Variabilidad individual en los niveles de PTH (diagrama de cajas).

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

En cuanto al tratamiento recibido para el control de las alteraciones del metabolismo óseo-mineral, en un 80% de los registros los pacientes no recibían tratamiento con fármacos captadores del fósforo. Entre los tratados, un 90% llevaban quelantes cálcicos, y en 11 de los registros no se obtuvo este dato (Figura 3.22).

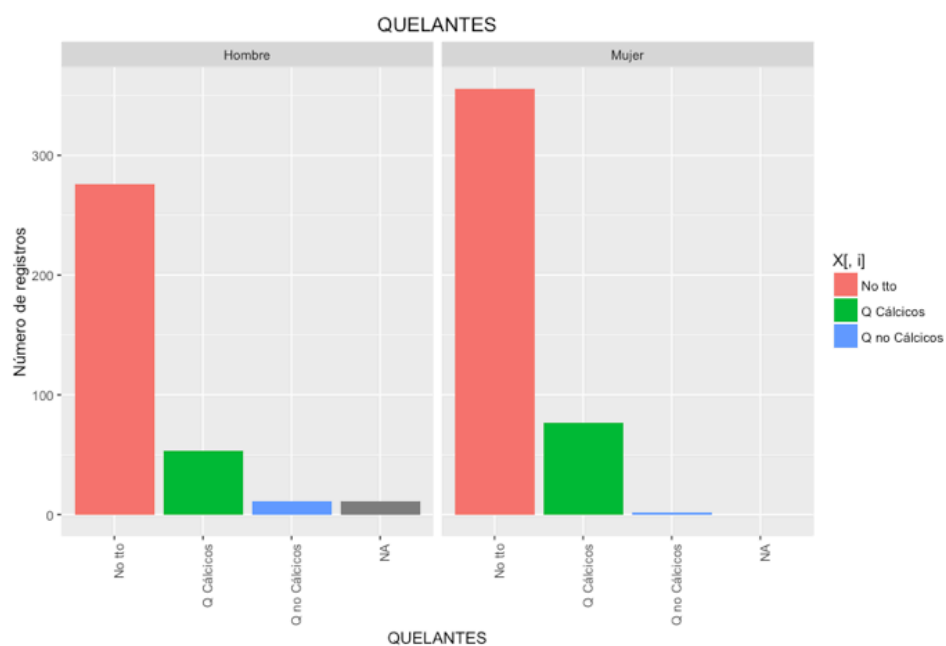


Figura 3.22. Tratamiento con fármacos quelantes del fósforo.

En un 56 % de los registros realizados los pacientes recibieron tratamiento con derivados activos de la vitamina D o con agonistas selectivos del receptor de la vitamina D, distribuido como se aprecia en la figura 3.24. En un 8,9% de los registros los pacientes recibían tratamiento con suplementos de vitamina D nativa.

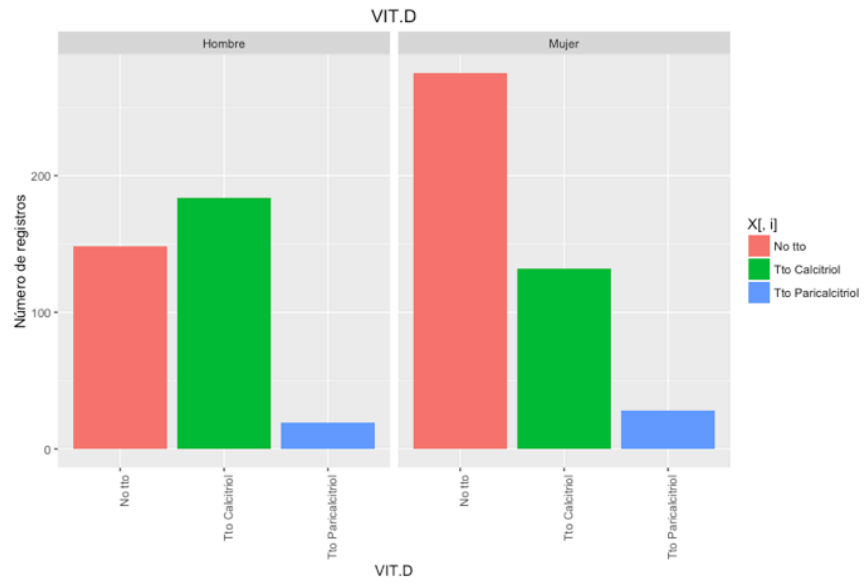


Figura 3.23. Tratamiento con derivados activos de vitamina D.

3.1.2.12. *Hiperuricemia*

La media de los niveles de ácido úrico en sangre obtenidos ha sido de 7.13 ± 1,39 mg/dl, con un rango entre 3.4 y 13.5 mg/dl.

En un 31.9 % de los registros realizados, los pacientes recibieron tratamiento con alopurinol.

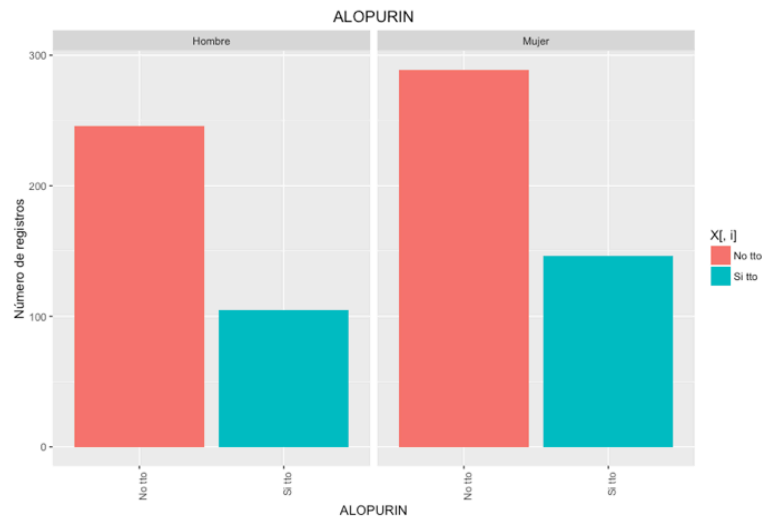


Figura 3.24. *Tratamiento de hiperuricemia con alopurinol.*

### 3.1.2.13. Proteinuria

La media de la proteinuria encontrada ha sido de  $1.5 \pm 1.8$  g/día (0-22.9), siendo en mujeres de  $1.2 \pm 1.5$  y en hombres de  $1,8 \pm 2,1$  g/día (Figura 3.25).

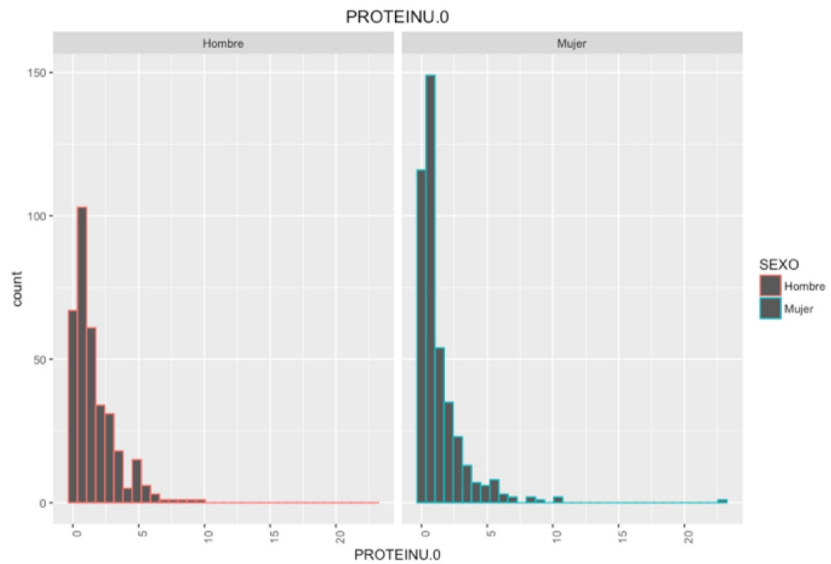


Figura 3.25. Distribución de proteinuria. En el eje de ordenadas se representa el valor de proteinuria en g/24 h y en las abscisas número de registros.

#### 3.1.2.14. Albúmina en sangre

Los niveles medios de albúmina han sido de  $4.1 \pm 10.3$  (2 – 6.9) g/dl (Figura 3.26).

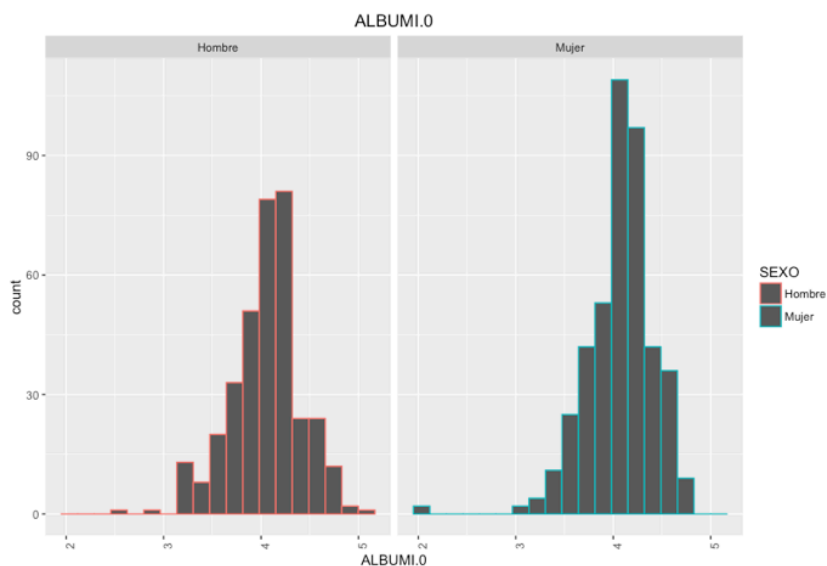


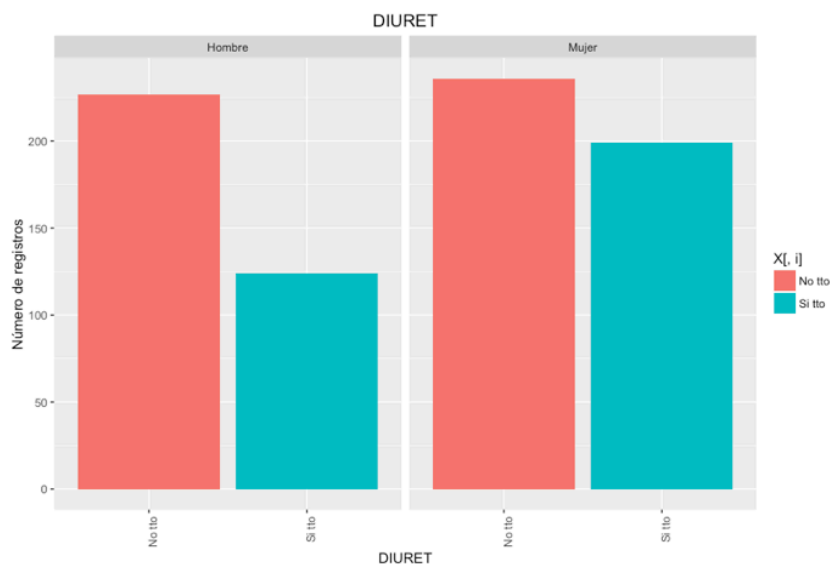
Figura 3.26. Distribución de los niveles de albúmina por sexos. En el eje de ordenadas niveles de albúmina en g/dl y en el eje de abscisas número de registros.

#### 3.1.2.15. PCR

La cifra media de PCR que han presentado los pacientes ha sido de 21.9 +/- 53.7 mg/dl (0-284), destacando que no se pudo recopilar este dato en 480 de los 786 registros realizados. A destacar que ha sido más elevada en hombres (24,45 versus 19,84 mg/dl).

### 3.1.2.16. Diuréticos

Un 41 % de los pacientes precisaron tratamiento con diuréticos de asa en algún momento de la evolución. En la *figura 3.27* vemos la distribución del tratamiento por sexos.



*Figura 3.27. Tratamiento diurético.*



**3.1.2.17. Inhibidores del sistema renina-angiotensina-aldosterona**

Un 41 % de pacientes recibieron tratamiento con inhibidores de la enzima convertidor de la angiotensina (IECAs) o con bloqueantes del receptor de angiotensina II (ARA-II), y un 4.8 % fueron tratados con la asociación de ambos. (Figura 3.28).

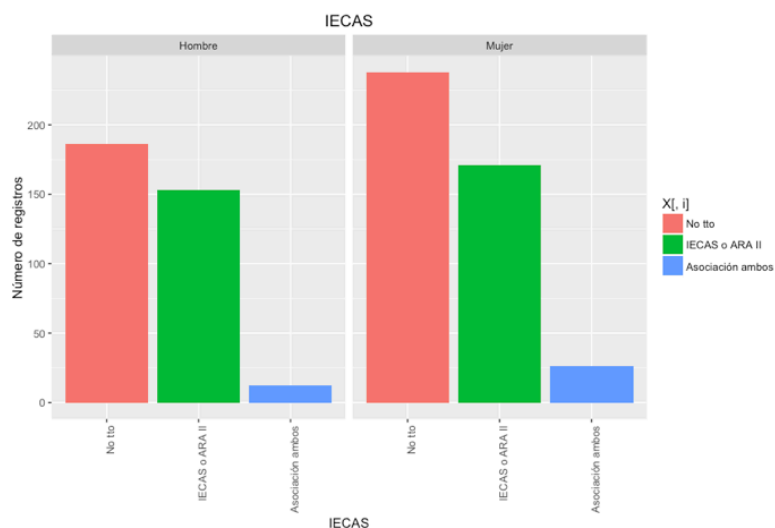


Figura 3.28. Tratamiento con inhibidores del sistema renina-angiotensina-aldosterona.

### 3.1.2.18. Hipertrofia ventricular izquierda

En 101 de los 115 pacientes se obtuvieron datos de ECG y ecocardiografía que permitieron determinar si presentaban hipertrofia ventricular izquierda. De estos pacientes, 36 (35.6% del total) cumplieron criterios de hipertrofia ventricular izquierda, siendo ligeramente más prevalente en hombres que en mujeres (41.1% versus 37.7%), *fig 3.29*.

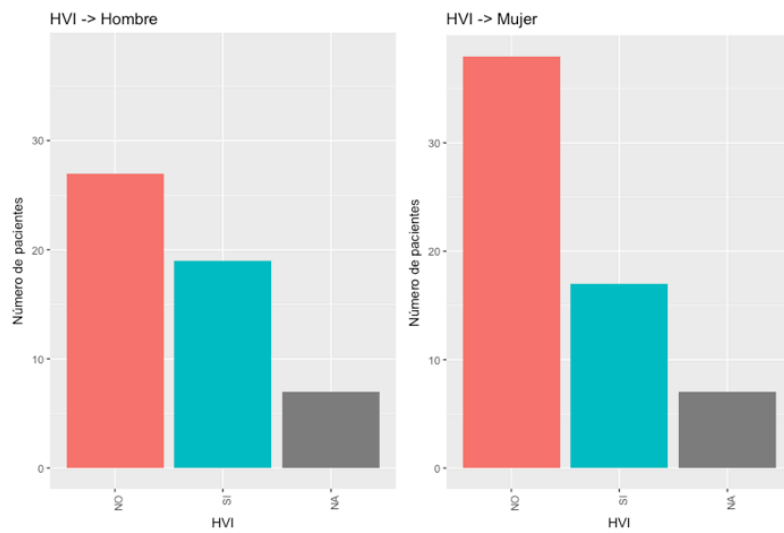


Figura 3.29. Hipertrofia ventricular izquierda.

3.1.2.19. *Insuficiencia cardíaca congestiva*

En la muestra estudiada 97 pacientes (83.6 % del total) no tenían criterios ecocardiográficos de disfunción ventricular y además estaban asintomáticos o con disnea grado 1 de la clasificación de la *NYHA*. Un paciente tenía disnea grado 2-3 (con esfuerzos moderados) y 7 pacientes presentaban disnea grado 4(en reposo). Además 10 pacientes tenían alteración de la función ventricular en ecardiografía estando totalmente asintomáticos (*Figura 3.30*).

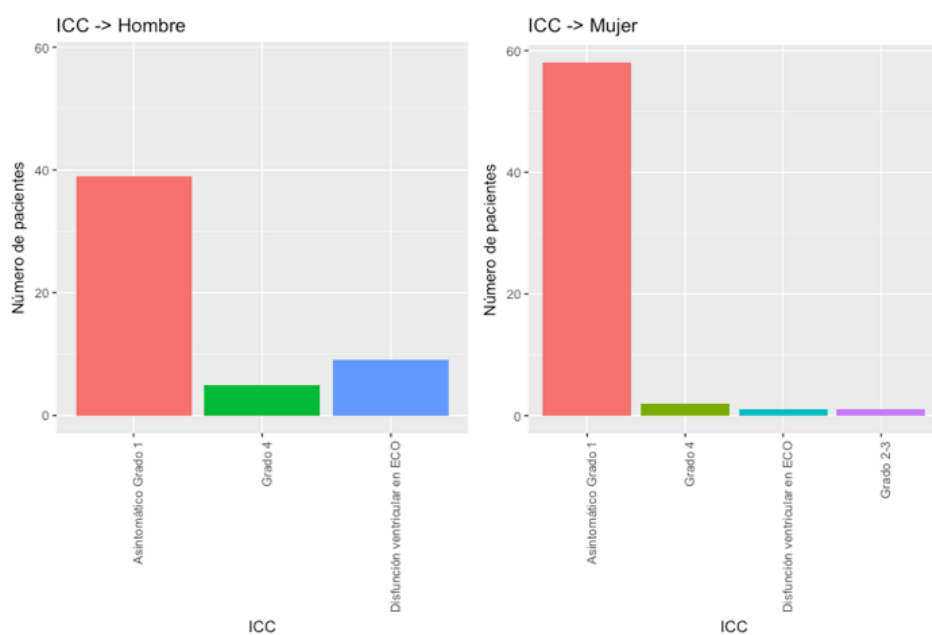


Figura 3.30. *Insuficiencia cardíaca.*

**3.1.2.20. Tratamiento antiagregante/anticoagulante**

Los pacientes han recibido tratamiento antiagregante en un 31.80 % de los registros realizados y anticoagulación oral en un 7,63 %.

#### 3.1.3. MEDICIÓN DE LA PROGRESIÓN DE LA ENFERMEDAD RENAL

Para medir la progresión de la insuficiencia renal se ha calculado la modificación del FG a lo largo del tiempo de seguimiento de cada paciente. Para ello, se ha determinado en los controles sucesivos practicados a cada paciente, y se ha realizado un ajuste lineal para expresar esta progresión en  $ml/min/1.73 m^2/año$ . De esta manera, obtenemos la tendencia de la función renal de cada paciente, es decir, en los casos en los que la función renal ha permanecido estable o con un leve deterioro, obtendremos una línea horizontal o con una ligera pendiente de descenso, y en los que ha habido mayor progresión de la enfermedad renal encontraremos una línea descendente con una mayor pendiente en la evolución del FG.

En la *figura 3.31*, a modo de ejemplo, vemos la evolución del filtrado glomerular de uno de los pacientes. En la primera gráfica se representan con puntos rojos todas las mediciones del FG realizadas a este paciente, pudiéndose apreciar una trayectoria variable en el mismo. Tras realizar una regresión lineal vemos que la tendencia en este caso es hacia una progresión lenta de la enfermedad renal, con un deterioro de  $-1.5 ml/min/1.73 m^2/año$ .

Se ha calculado la media de las pendientes de descenso del FG de todos los pacientes, siendo el resultado de  $-1.41 \pm 3.19 ml/min/1.73 m^2/año$ .

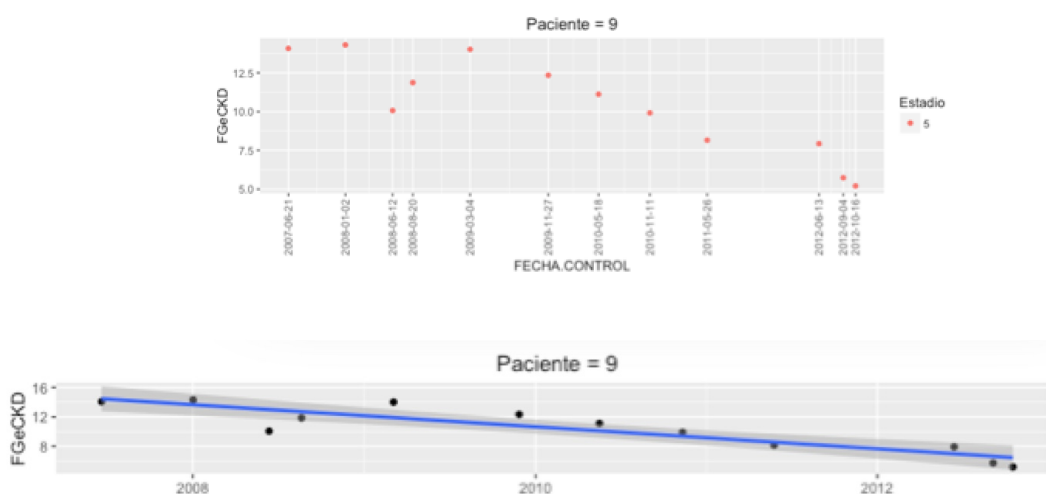


Figura 3.31. Evolución del FGe en un paciente.

#### **3.1.4. CORRELACIÓN DE LAS VARIABLES CON LA PROGRESIÓN DE LA ENFERMEDAD RENAL**

En la *tabla 3.2* se expresa la posible relación lineal (definida por el coeficiente de correlación) entre las diferentes variables continuas y la media del FG estimado por CKD-EPI. Se representan todas las variables que han mostrado correlación lineal con el FG con significación estadística ( $p < 0,05$ ), pero normalmente el criterio que se sigue es considerar una posible relación lineal si el coeficiente de correlación es, en valor absoluto, mayor o igual a 0.7. Siguiendo este criterio, dentro de esta tabla aparecen las siguientes relaciones lineales: con el FG estimado por otras fórmulas el coeficiente de correlación es prácticamente la unidad, lo que nos muestra la equivalencia de las tres fórmulas en grados avanzados de la enfermedad. También se ha encontrado correlación con el FGe basal y con la creatinina sérica.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

	Coefficiente de correlación	p valor
FGeCKD	1	0
FGeMDRD_4	0.987	1.66e-92
FGeMDRD_IDMS	0.987	1.66e-92
FG inicio	0.787	1.97e-25
Creatinina	-0.766	2e-23
Urea	-0.545	3.05e-10
PTH	-0.479	6.16e-08
PTHInicial	-0.444	3.33e-06
Fósforo	-0.404	7.46e-06
mFosforolni	-0.327	0.000452
PendienteAny	0.282	0.00225
Edad	-0.267	0.00387
AGAP_AAAlbumina	-0.246	0.0106
mFGInicio	0.223	0.0164
Albúmina	0.21	0.0241
AGAP_ACalcio	-0.211	0.0289
Hb	0.193	0.0385
Acido úrico	0.189	0.043
Bicarbonato	0.186	0.0461
Bicarbonato Ini	0.198	0.0483

Tabla 3.2. Correlación lineal entre variables continuas y el FGe por CKD\_EPI.

En la *tabla 3.3* se representan las correlaciones lineales con significación estadística entre las diferentes variables continuas con la pendiente de descenso del FG (PendienteAny). A diferencia del caso anterior, ninguna ha mostrado un coeficiente de correlación > 0,7, lo cuál quiere decir que, aunque existe significación estadística, la correlación entre las variables es débil.

	Coeficiente de correlación	p valor
PendienteAny	1	0
Proteinuria	-0.5	1.28e-08
Creatinina	-0.419	3.18e-06
mFGInicio	0.352	0.000113
FGeMDRD_IDMS	0.326	0.000383
FGeMDRD_4	0.326	0.000383
Fósforo	-0.312	0.000682
PTH.	-0.297	0.00128
FGeCKD	0.282	0.00225
mFosforolNi	-0.284	0.00255
ProteinurialNi	-0.281	0.00372
PTHlNi	-0.247	0.0129
Albumina	0.223	0.0164
COL.LDL	-0.217	0.0196
Bicarbonato	0.215	0.0213
Edad	0.213	0.0224
ColCharlsonN	0.204	0.029
mHblNi	0.202	0.0314
Urea	-0.199	0.0334
Tiempo seguim	0.194	0.0378
COL.TOT	-0.191	0.0411

*Tabla 3.3. Correlación lineal entre variables continuas y pendiente de modificación del FG.*

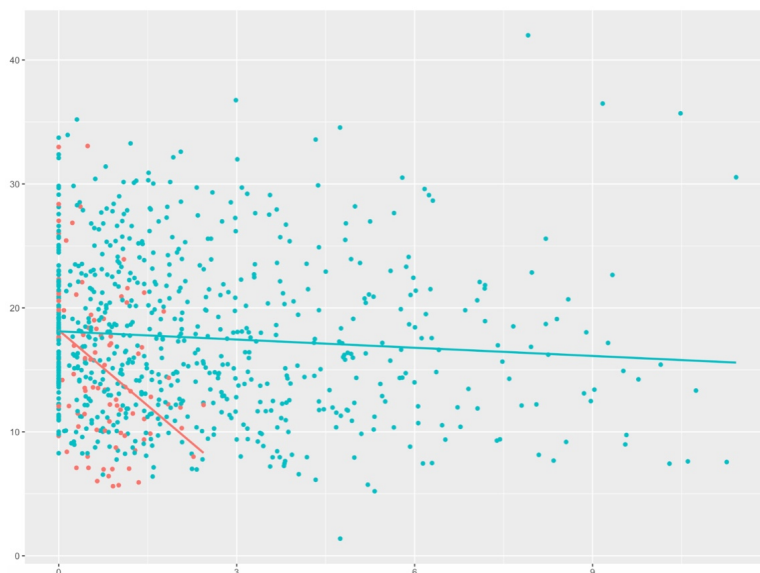
---



#### 3.1.5. CLASES DE PROGRESIÓN DE LA ENFERMEDAD RENAL

Se ha dividido la muestra en dos grupos que denominamos progresores rápidos y progresores lentos, fijando el punto de corte en una progresión de descenso del FG de  $-5 \text{ ml/min/1,73m}^2/\text{año}$ , puesto que en diferentes guías clínicas se considera una progresión rápida de la insuficiencia renal a partir de este punto (guías NICE, documento de consenso SEN).

De esta manera hemos obtenido un grupo de 18 pacientes que han evolucionado con una progresión rápida, es decir, mayor o igual a  $-5 \text{ ml/min/año}$  y otro grupo de 97 con una progresión menor. En la *figura 3.32* vemos la trayectoria de descenso del FG que han experimentado los dos grupos de pacientes y en la *tabla 3.4* se recogen las características básicas de ambos grupos (pendiente de descenso del FG, tiempo de seguimiento y FG medio durante el seguimiento y al inicio del mismo).



*Figura 3.32. Representación de la trayectoria de descenso del filtrado glomerular mediante regresión lineal. La línea roja representa la evolución de los pacientes progresores rápidos y la azul la de los progresores lentos.*

	n	Progresión (ml/min/año/1,73 m <sup>2</sup> )	Seguimiento (años)	FG inicial	FG seguimiento (ml/min/1.73 m <sup>2</sup> )
PROGRESORES RÁPIDOS	18	-7.09 ± 1.69	1.27 ± 1.44	18.12 ± 7.59	15,4 ± 5,08
PROGRESORES LENTOS	97	-0.69 ± 2.52	2.98 ± 3.81	20.18 ± 6.04	16,97 ± 8,11

Tabla 3.4. Características de los pacientes de ambos grupos. Cada paciente se ha caracterizado por el promedio/media de la variable considerada.

### 3.1.6. ANÁLISIS COMPARATIVO DE LOS DOS GRUPOS DE PROGRESIÓN

Tras dividir a los pacientes en progresores rápidos y progresores lentos, inicialmente hemos comparado en estos dos grupos de pacientes las diferentes variables analizadas, mediante contraste de hipótesis.

En algunas variables cuantitativas como (edad y talla), se ha utilizado únicamente el valor de la primera visita entre los dos grupos. En el resto de variables cuantitativas, se ha caracterizado cada una de ellas por su media o mediana en cada paciente a lo largo del seguimiento, y en algunas de ellas se han utilizado además los valores al inicio el estudio y los valores de modificación de la pendiente entre los dos primeros controles realizados al inicio del seguimiento. Las variables analizadas de estas tres formas han sido: FG estimado por CKD -EPI, proteinuria, niveles de bicarbonato en sangre, Hb, PTH, fósforo y ácido úrico.

Para caracterizar cada variable dentro del conjunto de pacientes, si la distribución de la variable es normal, se utiliza la media y su varianza, y, si es no normal, la mediana y su rango intercuartílico.

Antes de realizar el análisis comparativo se ha verificado la homogeneidad de las varianzas mediante el test de Levene.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Para realizar el análisis comparativo de las variables continuas en ambos grupos de pacientes se ha empleado el test de *Student* para aquellas que siguen una distribución normal y el de *Wilcoxon* para las que no (Anexo 1). Se ha considerado que existe significación estadística cuando  $p < 0.05$ . Por tanto, se han obtenido diferencias significativas en las variables expuestas en la *tabla 3.5*:

	p valor	Intervalo confianza
PendienteAny	1.87e-11	-7.28, -5.30
Tiempo seguim	1.15e-06	-2.890, -0.816
Proteinuria	1.21e-05	1.18, 3.00
Proteinurialni	0.000157	0.76, 2.60
mFG inicio	0.002	-11.88, -2.82
Fósforo	0.00378	0.155, 0.716
Creatinina	0.00437	0.243, 1.162
IncTiempo	0.0144	-0.2657, -0.0219
Calcio	0.0437	-0.40788, -0.00717

*Tabla 3.5. Variables continuas en las que se han encontrado diferencias significativas entre los pacientes con progresión rápida y lenta de la enfermedad renal.*

El **tiempo de seguimiento**, menor en los pacientes con progresión rápida de la enfermedad renal ( $2.98 \pm 3.78$  años en el grupo de progresión lenta versus  $1.33 \pm 0.49$  años en los progresores rápidos,  $p=1.15 \text{ e-}0.6$ ).

Grado de **proteinuria**, tanto la media de todos los valores de proteinuria que ha tenido cada paciente durante todo el seguimiento ( $3.24 \pm 1.92$  g/24 h en los progresores rápidos versus  $0.73 \pm 1.16$  g/día en los lentos,  $p = 1.21 \text{ e-}0.5$ ), como la proteinuria que presentaron al inicio del seguimiento ( $3.15 \pm 2.45$  g/24 h en el grupo de progresión rápida frente a  $0.6 \pm 1.14$  g/24 h en los progresores lentos,  $p = 0.000157$ ).

Niveles de **fósforo** en sangre, más elevados en pacientes progresores ( $4.44 \pm 0.53$  mg/dl versus  $4.01 \pm 0.53$  mg/dl,  $p = 0.00378$ ).

**Creatinina sérica** durante el seguimiento, más elevada en los pacientes progresores rápidos ( $3.97 \pm 0.87$  mg/dl versus  $3.27 \pm 0.82$ ,  $p = 0.00437$ ).

Niveles más bajos de **calcio** en los pacientes progresores rápidos ( $9.59 \pm 0.4$  versus  $9.3 \pm 0.5$ ,  $p = 0.0437$ ).

Mayor **pendiente de descenso del FG** entre los dos primeros controles realizados (mFGInicio) en el grupo de progresión rápida ( $- 6.55 \pm 12.12$  ml/min/1.73 m<sup>2</sup>/año versus  $0.17 \pm 7.32$  ml/min/1.73 m<sup>2</sup>/año,  $p = 0.002$ ).

En el **Anexo I** se pueden encontrar las medias y medianas de todas las variables cuantitativas que se han analizado en ambos grupos de pacientes y la comparación de las mismas.

En el caso de las variables categóricas, se ha caracterizado cada paciente por el valor que más se repite (moda), por ejemplo, si un paciente ha recibido tratamiento con epoetina en cuatro de seis visitas, se considera que sí que lleva este tratamiento. Para realizar el análisis de posibles dependencias entre la variable progresor rápido/progresor lento y el resto de variables discretas se ha empleado el test de Chi-cuadrado, y, en aquellos casos en que el número de elementos por celda ha sido menor de 5 se ha empleado el test de Fisher.

No se ha encontrado ninguna relación de dependencia entre ninguna de las variables discretas analizadas y la progresión rápida o lenta de la enfermedad renal. Estas variables han sido: diabetes mellitus, hipertensión arterial, hipertrofia ventricular izquierda, insuficiencia cardíaca, sexo, tabaquismo, tratamiento con alopurinol, medicación antihipertensiva, anticoagulación oral, antiagregantes, derivados de la vitamina D tanto activos, como suplementos en forma nativa, diuréticos, quelantes del fósforo, suplementos de bicarbonato, IECAs/ARA II y tratamiento con hierro.

#### 3.1.7. PROCEDIMIENTO EXPERIMENTAL EN EL AJUSTE DE MODELOS

A continuación se plantea el siguiente procedimiento experimental:

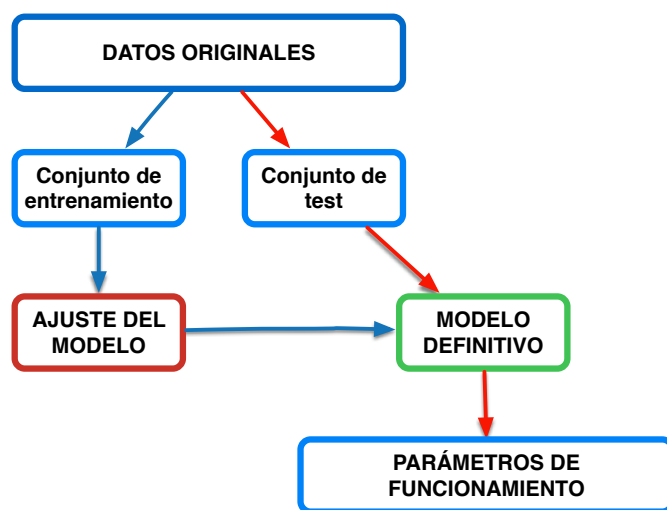
Se divide, de forma aleatoria, el conjunto inicial de datos en dos subconjuntos, uno al que llamaremos conjunto de entrenamiento que incluye el 70 % de los datos, y otro que denominaremos conjunto de test, en el que están el 30% restante.

Ajuste de los parámetros del modelo; mediante el primer subconjunto (conjunto de entrenamiento) se ajustará el modelo obteniendo sus parámetros óptimos.

Se obtendrán los valores de calidad del modelo (en nuestro caso área bajo la curva, AUC) además del umbral óptimo para discretizar la salida (recordemos que planteamos un problema de clasificación).

Con ese umbral óptimo se discretizará la salida del conjunto de test obteniendo lo que se conoce como matriz de confusión que es una tabla 2x2 donde se dan los patrones acertados/fallados por el modelo.

La *figura 3.33* describe el procedimiento experimental comentado.



*Figura 3.33. Esquema del procedimiento experimental planteado.*

### 3.1.8. MODELOS LINEALES UNIVARIANTES

Vamos a aplicar un modelo lineal a los datos con el fin de predecir si un individuo de la muestra tendrá una progresión rápida, o no, dependiendo de las variables que hemos considerado como significativas en etapas previas. La variable de salida del modelo es una variable dicotómica que define si el patrón de entrada es progresor o no progresor. En primer lugar se desarrollan una serie de regresiones logísticas para las variables que han salido significativas en análisis previos; la primera que nos encontramos es la proteinuria (valor medio de proteinuria durante el seguimiento). En este modelo se obtienen los siguientes valores para los coeficientes de la regresión:

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	2,987	0,220	2E-16
Proteinuria	-0,528	0,074	1E-12

Tabla 3.6. Coeficiente de regresión para proteinuria.

El área bajo la curva de este modelo, para el conjunto de entrenamiento, es de 0.83 por lo que se puede decir que es un buen modelo y que esta variable tiene una fuerte capacidad predictora (*Figura 3.34*).

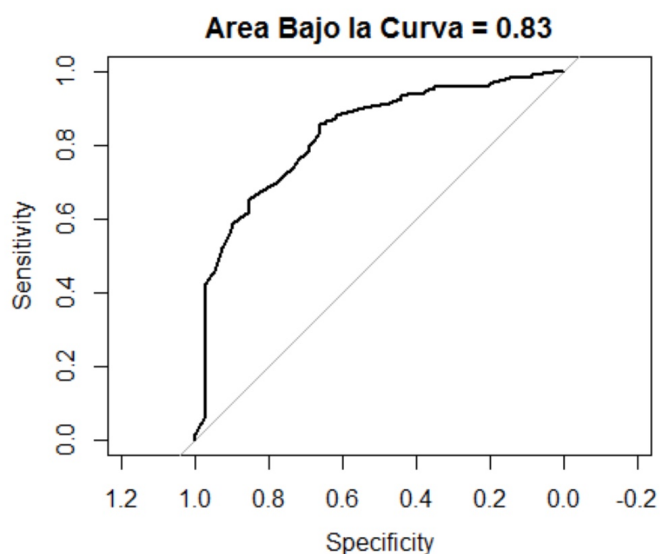


Figura 3.34. Curva sensibilidad/especificidad para el modelo lineal usando la proteinuria.

Con esta curva se determina el umbral óptimo para discretear las salidas del modelo. Con dicho umbral se determina la matriz de confusión para el conjunto de test

		PREDICCIÓN	
		0	1
REALIDAD	0	22	6
	1	41	166

Tabla 3.7. Matriz de confusión para proteinuria. (0: Progresores rápidos, 1: Progresores lentos). El modelo acertaría 22 de 28 casos en los casos de progresión rápida y 166 de 205 en los progresores lentos.

Repitiendo lo mismo para la variable **Proteinuria1ni** (proteinuria en el primer control) se obtiene lo siguiente:

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	2,987	0,161	2E-16
Proteinuria1ni	-0,168	0,046	0.0002

Tabla 3.8. Coeficiente de regresión para proteinuria basal.

El área bajo la curva de este modelo, para el conjunto de entrenamiento, es de 0.81 por lo que se puede decir, al igual que en el caso anterior, que es un buen modelo y que esta variable tiene una fuerte capacidad predictora (Figura 3.35):

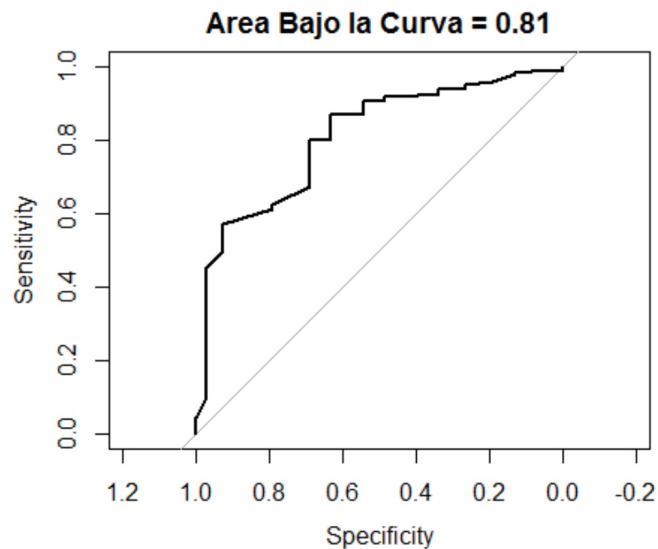


Figura 3.35. Curva sensibilidad/especificidad para el modelo lineal usando la proteinuria basal.



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Con esta curva se determina el umbral óptimo para discretizar las salidas del modelo. Con dicho umbral se determina la matriz de confusión obteniéndose los resultados mostrados en la tabla, que nos muestra que el papel predictor de la proteinuria inicial es discretamente peor que el de la proteinuria media en el seguimiento.

		PREDICCIÓN	
		0	1
REALIDAD	0	21	7
	1	44	138

*Tabla 3.9. Matriz de confusión para Proteinuria inicial (0: Progresores rápidos, 1: progresores lentos).*

Hay que destacar que aparecen diferentes valores en la tabla de confusión porque se necesitan patrones completos para poder realizar la regresión logística. Como son variables diferentes se tiene diferente número de valores completos para cada una de ellas.

Repitiendo del mismo procedimiento para el fósforo:

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	5,144	0,747	6E-12
Fósforo	-0,754	0,168	7,4E-06

Tabla 3.10. Coeficiente de regresión para fósforo.

La curva ROC muestra peores resultados que en los casos anteriores llegando a 0.66 como AUC para el conjunto de entrenamiento por lo que esta variable no tiene tanta capacidad predictora como las anteriores (Figura 3.36).

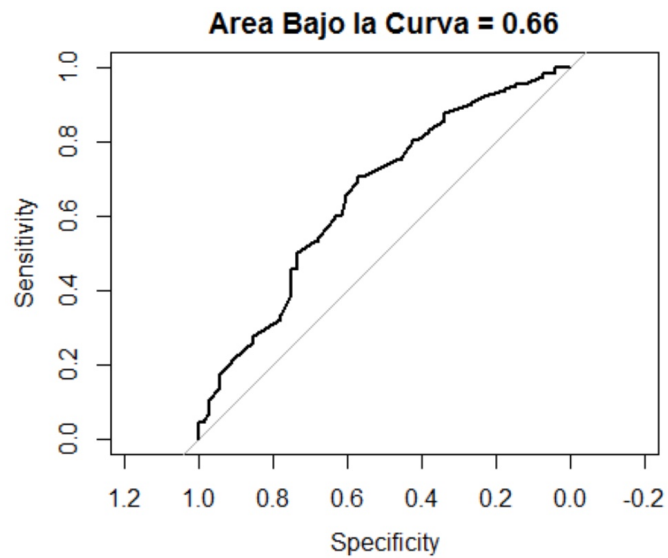


Figura 3.36. Curva sensibilidad/especificidad para el modelo lineal usando el fósforo.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Esta capacidad predictora más baja se confirma al obtener la matriz de confusión para el conjunto de test (*tabla 3.11*).

		PREDICCIÓN	
		0	1
REALIDAD	0	19	9
	1	61	146

*Tabla 3.11. Matriz de confusión para fósforo (0: Progresores 1: No progresores). Indicaría un acierto de 19 sobre 28 casos en pacientes con progresión rápida y de 146 sobre 212 no progresores.*

Ahora se repite el procedimiento para la creatinina y el calcio y se observa que no tienen una capacidad predictora alta.

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	3,074	0,416	<b>1E-13</b>
Creatinina	-0,308	0,106	<b>1E-03</b>

Tabla 3.12. Coeficiente de regresión para creatinina.

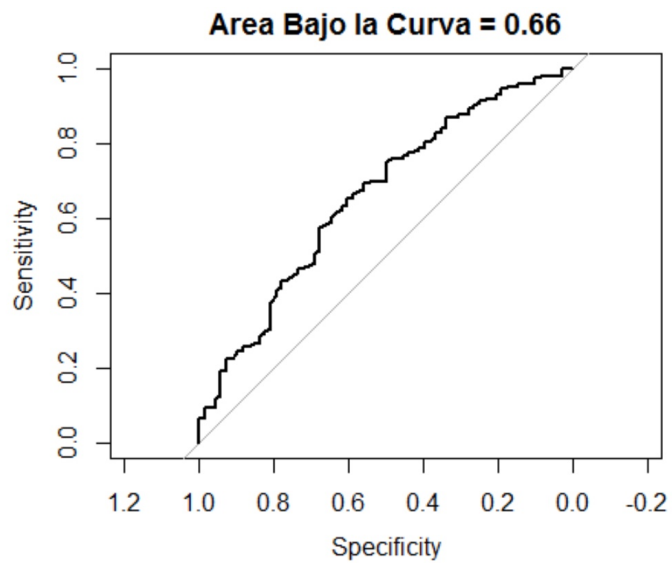


Figura 3.37. Curva sensibilidad/especificidad para el modelo lineal usando la creatinina.

		PREDICCIÓN	
		0	1
REALIDAD	0	16	12
	1	67	140

Tabla 3.13. Matriz de confusión para creatinina (0: Progresores rápidos, 1: Progresores lentos).

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

Para el calcio se obtiene lo siguiente:

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	-1,149	2,218	0,604
Calcio	0,328	0,234	0,162

Tabla 3.14. Coeficiente de regresión para calcio.

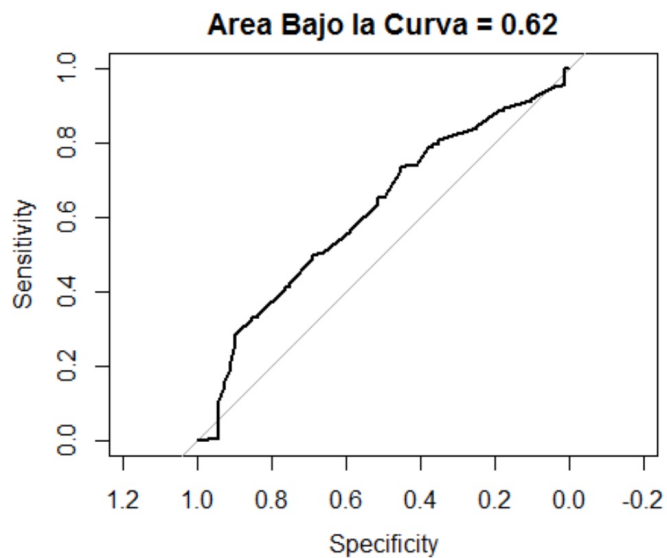


Figura 3.38. Curva sensibilidad/especificidad para el modelo lineal usando el calcio.

		PREDICCIÓN	
		0	1
REALIDAD	0	20	8
	1	110	97

Tabla 3.15. Matriz de confusión para calcio (0: Progresores rápidos, 1: Progresores lentos).

### 3.1.9. MODELOS MULTIVARIANTES

#### 3.1.9.1. Regresión logística

En este caso se combinan las variables que se obtuvieron como significativas en la etapa anterior de análisis. No se han incluido en el análisis multivariante las variables que han sido significativas anteriormente pero que se consideran indicadores *directos* de la progresión como son el tiempo de seguimiento (o DurRegAnys), que más bien es una consecuencia de la velocidad de progresión, y la pendiente de modificación del FG (PendienteAny), que ha sido en criterio a partir del cuál se han dividido los grupos.

Así pues se obtienen como coeficientes del modelo los siguientes (Tabla 3.16):

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	5,940	2,548	0,0198
Proteinuria	-0,470	0,101	3,9E-06
ProteinuriaIni	0,055	0,063	0,376
Fósforo	-0,276	0,220	0,209
Creatinina	-0,240	0,153	0,117
Calcio	-0,129	0,252	0,607

Tabla 3.16. Coeficientes de regresión de las diferentes variables en el modelo multivariante.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

Si se determina la curva ROC y el área bajo la curva se obtiene un valor de 0.8 que es igual al obtenido con el modelo univariante usando la variable Proteinuria o la ProteinuriaIni (Figura 3.39).

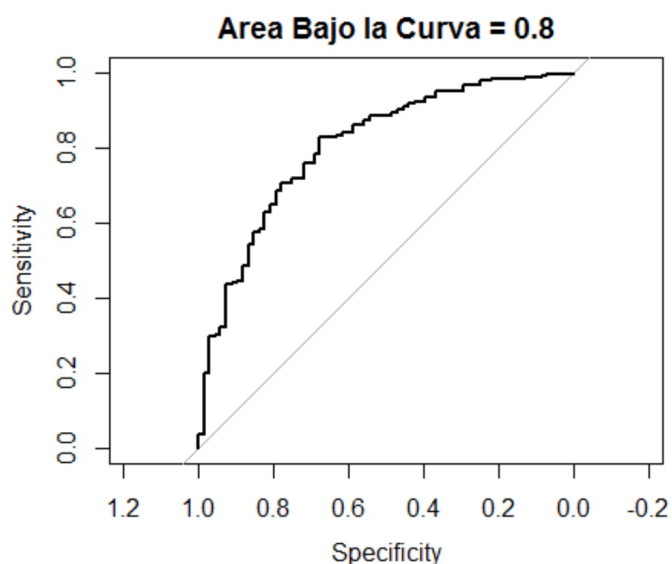


Figura 3.39. Curva sensibilidad/especificidad para el modelo lineal usando las variables estadísticamente significativas del análisis univariante.

En cuanto a la matriz de confusión para el conjunto de entrenamiento se obtiene:

		PREDICCIÓN	
		0	1
REALIDAD	0	21	7
	1	27	155

Tabla 3.17. Matriz de confusión para modelo de regresión logística.

### 3.1.9.2. Naïve Bayes

En este caso se obtiene la siguiente curva ROC (área bajo la curva de 0.827); seguimos sin superar el valor obtenido usando solamente la variable *Proteinuria* o la *ProteniuriaInl*.

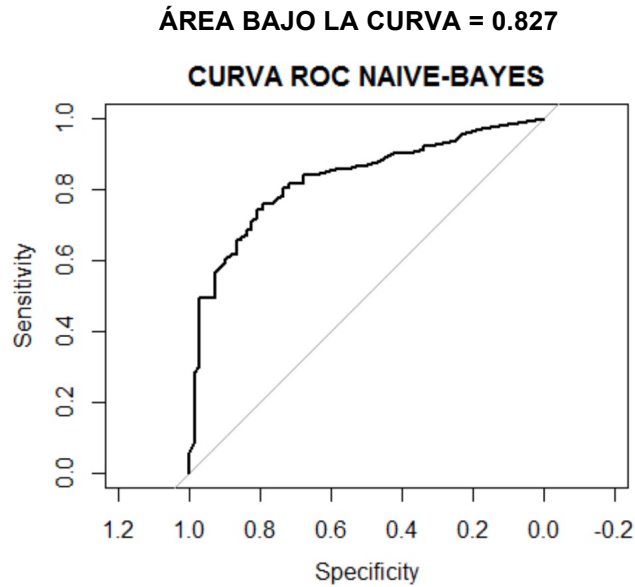


Figura 3.40. Curva sensibilidad/especificidad para el Naïve Bayes usando las variables estadísticamente significativas del análisis univariante.

En cuanto a la matriz de confusión para el conjunto de entrenamiento se obtiene:

		PREDICCIÓN	
		0	1
REALIDAD	0	22	6
	1	36	146

Tabla 3.18. Matriz de confusión para Naïve-Bayes. (0: Pacientes con progresión rápida, 1: Pacientes con progresión lenta).



**3.1.9.3. TAN (Tree augmented network)**

En este caso se obtiene la siguiente curva ROC (área bajo la curva de 0.853).

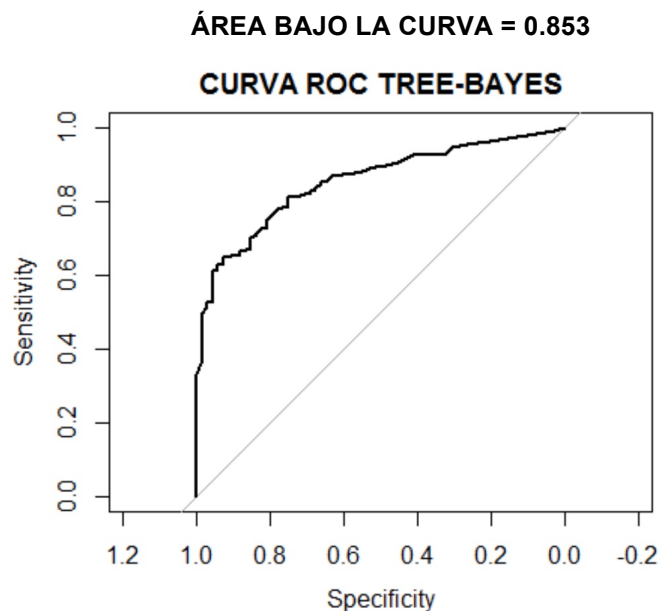


Figura 3.41. Curva sensibilidad/especificidad para el TAN usando las variables estadísticamente significativas.

En cuanto a la matriz de confusión para el conjunto de entrenamiento se obtiene:

		PREDICCIÓN	
		0	1
REALIDAD	0	23	5
	1	37	145

Tabla 3.19. Matriz de confusión para modelo TAN (0: Pacientes con progresión rápida, 1: Pacientes con progresión lenta).

En este caso la relación causal entre variables viene representada en la figura 3.42.

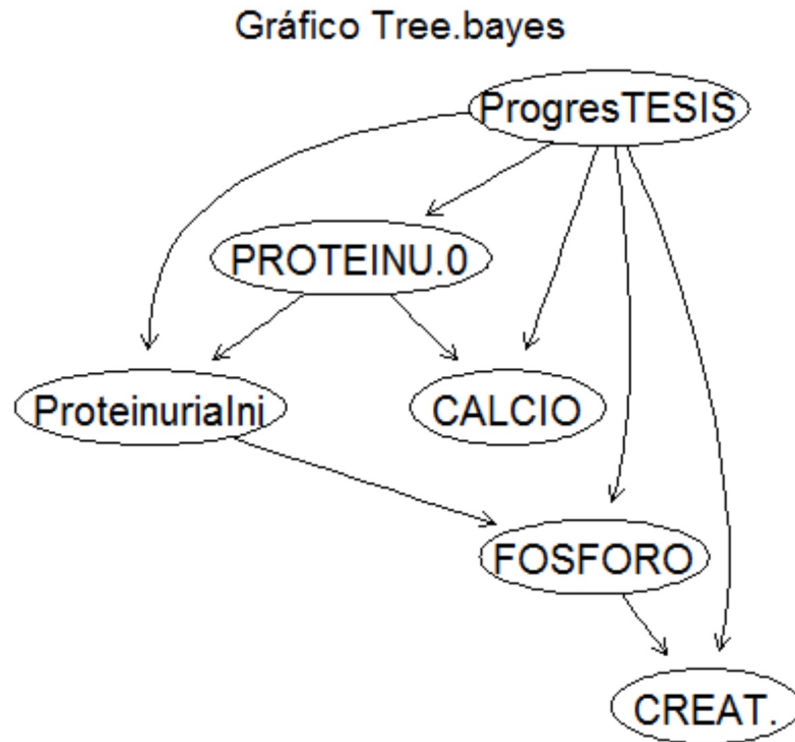


Figura 3.42. Esquema de las relaciones encontradas en el modelo TAN.

Se representan las relaciones probabilísticas entre las variables mediante el modelo TAN que nos muestra la dependencia entre ellas. Este esquema es la base del modelo predictivo mediante el que, introduciendo el valor de cada una de las variables obtenidas en un paciente determinado, podremos predecir si va a tener una progresión rápida o lenta de la enfermedad renal. El modelo ha mostrado una buena capacidad de predicción, por presentar un AUC de 0,8, pero en la matriz de confusión vemos que mejora sólo discretamente los modelos lineales (acierta 23 casos de 28 en el caso de progresión rápida).

#### 3.1.9.4. Comparación de modelos

Las diferencias entre los modelos son tan pequeñas que se variaron los conjuntos de entrenamiento/test para comprobar el funcionamiento de los diferentes modelos. Para ello se realizaron 30 divisiones diferentes de los conjuntos (todas de forma aleatoria) y se determinaron dos medidas de funcionamiento: *precisión* y *kappa*. El primer índice es el porcentaje de acierto en cada clase y, el segundo, mide el grado de concordancia entre el modelo y la realidad (Peña 08). Se suele considerar una buena concordancia cuando este valor es superior a 0.5 (0.4 - 0.6: moderado, 0.6 - 0.8: sustancial, 0.8 - 1: casi perfecto). Las *tablas 3.20* y *3.21* muestra las distribuciones de estos parámetros:

PRECISIÓN			
	PRIMER CUARTIL	SEGUNDO CUARTIL	TERCER CUARTIL
LINEAL	0,84	0,86	0,88
NAÏVE BAYES	0,78	0,81	0,84
TAN	0,82	0,85	0,87

KAPPA			
	PRIMER CUARTIL	SEGUNDO CUARTIL	TERCER CUARTIL
LINEAL	0,00	0,04	0,22
NAÏVE BAYES	0,22	0,33	0,41
TAN	0,01	0,12	0,28

*Tablas 3.20 y 3.21. Comparación de modelos multivariantes.*

Se observa que el funcionamiento de los tres sistemas es muy similar en precisión pero el índice kappa es muy bajo. Las siguientes figuras muestran la distribución de los parámetros de precisión y kappa obtenidos.

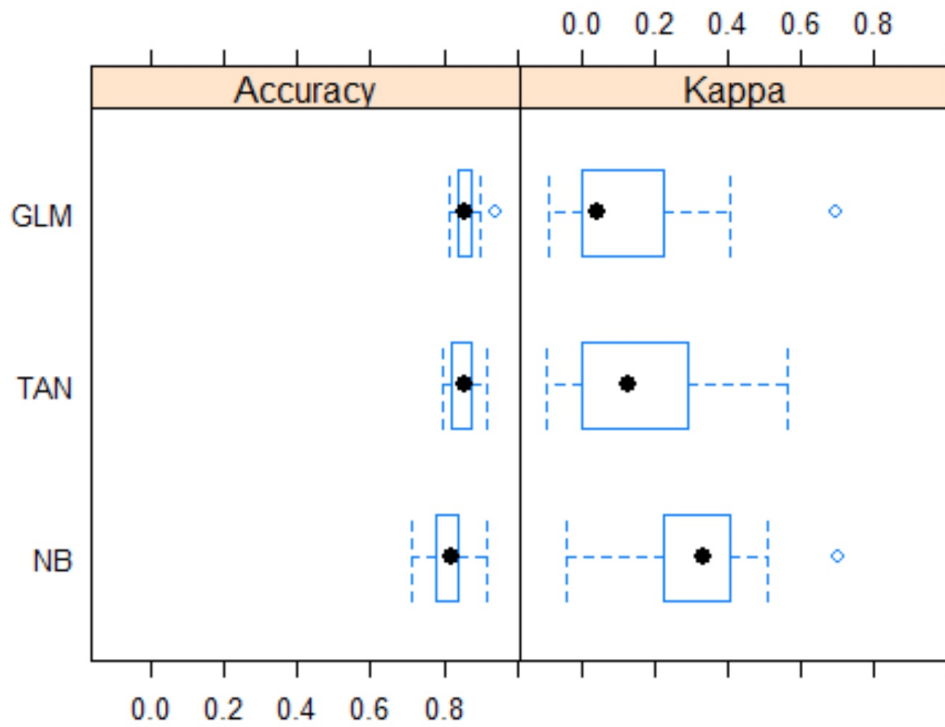


Figura 3.43. Comparación entre los diferentes modelos variando los conjuntos de entrenamiento/validación.

3.1.10. SELECCIÓN DE VARIABLES

3.1.10.1. Regresión logística

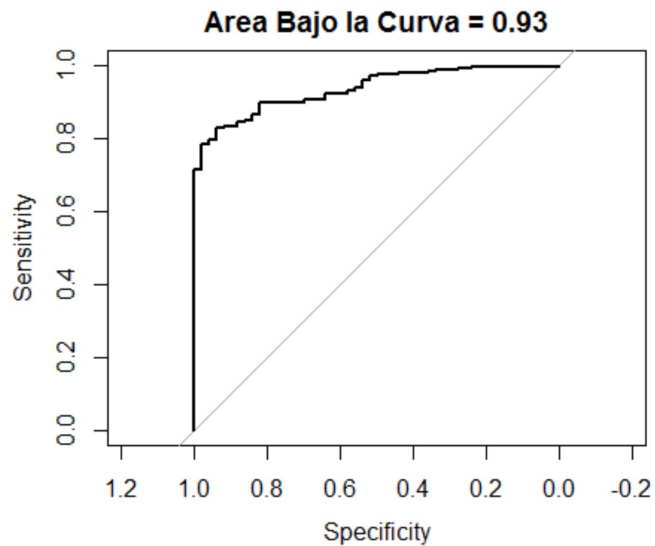
El método Boruta es un algoritmo de selección de variables basado en la idea de *Random Forest*. La base del algoritmo es utilizar variables aleatorias como entradas al modelo que se compara con el modelo original. La importancia de las variables originales se compara con la de las variables aleatorias. De manera que únicamente las variables cuya importancia es mayor que la de las variables aleatorias son consideradas importantes por el algoritmo utilizado (Kursa 10). Para mantener la coherencia del estudio se consideró el mismo número de variables que en el caso anterior, cinco, volviéndose a realizar el ajuste de los mismos modelos. Al igual que en el caso anterior se dividió de forma aleatoria el conjunto de datos en entrenamiento, 70% de los datos y test 30% restante. Obtenemos en primer lugar la regresión logística con la selección por Boruta, que incluye las variables Proteinuria, proteinuria basal o inicial, PTH basal, FG basal y pendiente de modificación de la PTH entre los dos primeros controles (*PTHIni*).

Variables	Coefficiente	Desviación estándar	Nivel de significancia
Sesgo	10,124	1,430	<b>1E-12</b>
mPTHIni	-0,006	7	0,371
ProteinuriaIni	0,212	0,097	<b>0,029</b>
PTHIni	-0,008	0,001	<b>2E-05</b>
Proteinuria	-0,671	0,148	<b>6E-06</b>
FrPorclnicio	-0,277	0,055	<b>5,4E-07</b>

Tabla 3.22. Coeficiente de regresión para regresión logística.

Si se determina la curva ROC para este modelo y se calcula el área bajo la curva se obtiene un valor de 0.93; que es un 10% superior al obtenido por el modelo lineal anterior.

Todas las variables seleccionadas por este método han alcanzado significación estadística excepto la pendiente de modificación de PTH entre los dos primeros controles. Si nos fijamos en la importancia de las variables basándonos en el valor absoluto de los coeficientes tenemos, por orden, la proteinuria media a lo largo del seguimiento, el FG basal y la proteinuria basal. La PTH basal, aunque sale significativa, tendría poco peso en el modelo predictivo de progresión de enfermedad renal por el bajo valor de su coeficiente.



*Figura 3.44. Curva sensibilidad/especificidad para el modelo lineal usando las variables seleccionadas por el Boruta.*

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Con esta curva se determina el umbral óptimo para discretizar las salidas del modelo. Con dicho umbral se determina la matriz de confusión para el conjunto de test obteniéndose los siguientes resultados:

		PREDICCIÓN	
		0	1
REALIDAD	0	14	7
	1	28	110

*Tabla 3.23. Matriz de confusión para regresión logística tras selección de variables por Boruta (0: Progresores rápidos, 1: Progresores lentos).*

### 3.1.10.2. Naïve Bayes

A continuación se ajusta un modelo de *Naïve Bayes* con las variables seleccionadas por el Boruta. El área bajo la curva de este modelo para el conjunto de test es de 0.89 (peor que el modelo lineal).

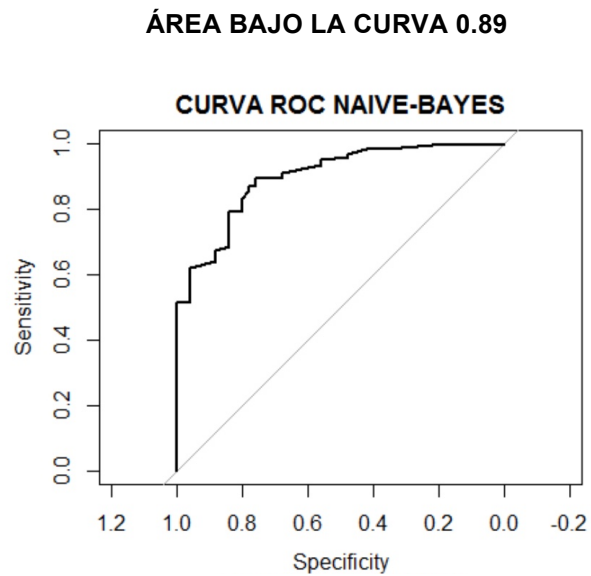


Figura 3.45. Curva sensibilidad/especificidad para el Naïve Bayes usando las variables seleccionadas por el Boruta.



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

En cuanto a la matriz de confusión para el conjunto de test se obtiene lo siguiente:

		PREDICCIÓN	
		0	1
REALIDAD	0	19	2
	1	21	117

*Tabla 3.24. Matriz de confusión para el conjunto de test (0: Progresores rápidos, 1: progresores lentos).*

Se aprecia que, teniendo un valor de AUC cercano al modelo lineal, la matriz de confusión para el conjunto de test es mejor que la del modelo lineal (mayor grado de acierto en los progresores y no progresores).

**3.1.10.3. TAN (Tree Augmented Network)**

Repitiendo el procedimiento para este modelo se llega a:

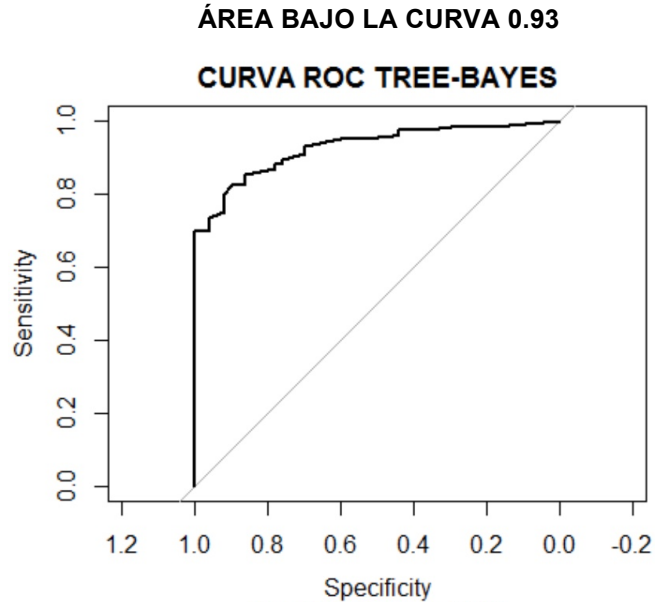


Figura 3.46. Curva sensibilidad/especificidad para el TAN usando las variables seleccionadas por el Boruta.

Repitiendo el procedimiento para el modelo TAN, obtenemos el mismo valor de AUC que el modelo lineal, 0.93 (Figura 3.46). En la tabla 3.25 se representa la matriz de confusión para el conjunto de test:

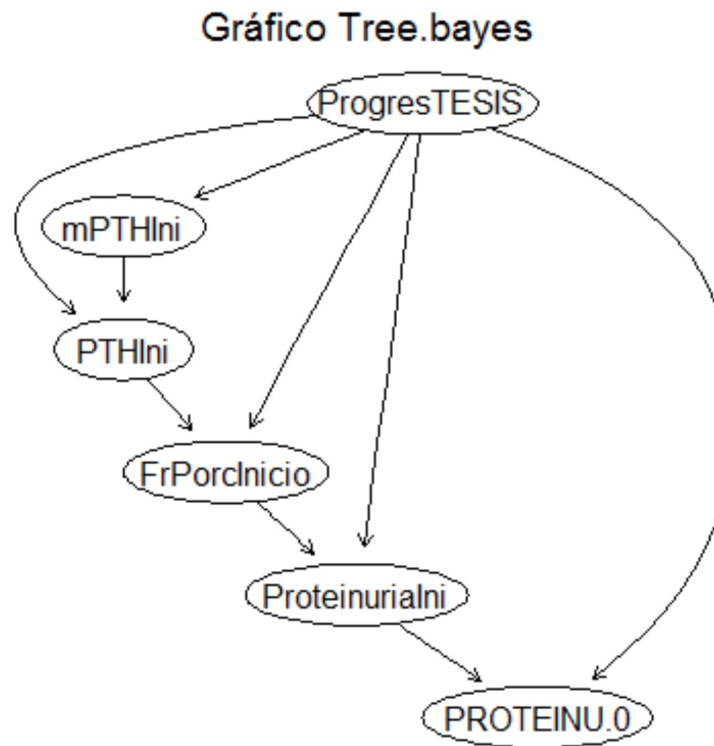
		PREDICCIÓN	
		0	1
REALIDAD	0	21	0
	1	16	122

Tabla 3.25. Matriz de confusión para el modelo TAN tras selección de variables por Boruta (0: Progresores rápidos, 1: Progresores lentos).

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Se observa que este modelo clasifica perfectamente todos los registros del grupo 0 ( progresores), con un 100% de aciertos, y, en el caso de los no progresores, acertaría 122 de 138 casos, es decir un 88 % de los no progresores. El gráfico correspondiente a este modelo es el definido en la *figura 3.47*.



*Figura 3.47. Esquema de las relaciones encontradas en el modelo TAN.*

**3.1.10.4. Comparación de modelos**

De nuevo las diferencias entre los modelos son tan pequeñas que se variaron los conjuntos de entrenamiento/test para comprobar el funcionamiento de los diferentes modelos. Se repite el mismo procedimiento del apartado 3.1.9.4 obteniéndose los siguientes resultados (*Tablas 3.26 y 3.27*):

PRECISIÓN			
	PRIMER CUARTIL	SEGUNDO CUARTIL	TERCER CUARTIL
LINEAL	0,87	0,89	0,92
NAÏVE BAYES	0,84	0,88	0,89
TAN	0,86	0,89	0,91

KAPPA			
	PRIMER CUARTIL	SEGUNDO CUARTIL	TERCER CUARTIL
LINEAL	0,30	0,44	0,62
NAÏVE BAYES	0,37	0,44	0,57
TAN	0,22	0,45	0,53

*Tablas 3.26 y 3.27. Comparación de modelos.*

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

Se observa que el funcionamiento de los tres sistemas es muy similar en precisión pero el índice kappa se ha mejorado en relación a los modelos con las variables estadísticamente significativas. La siguiente figura muestra la distribución de los parámetros de precisión y kappa obtenidos.

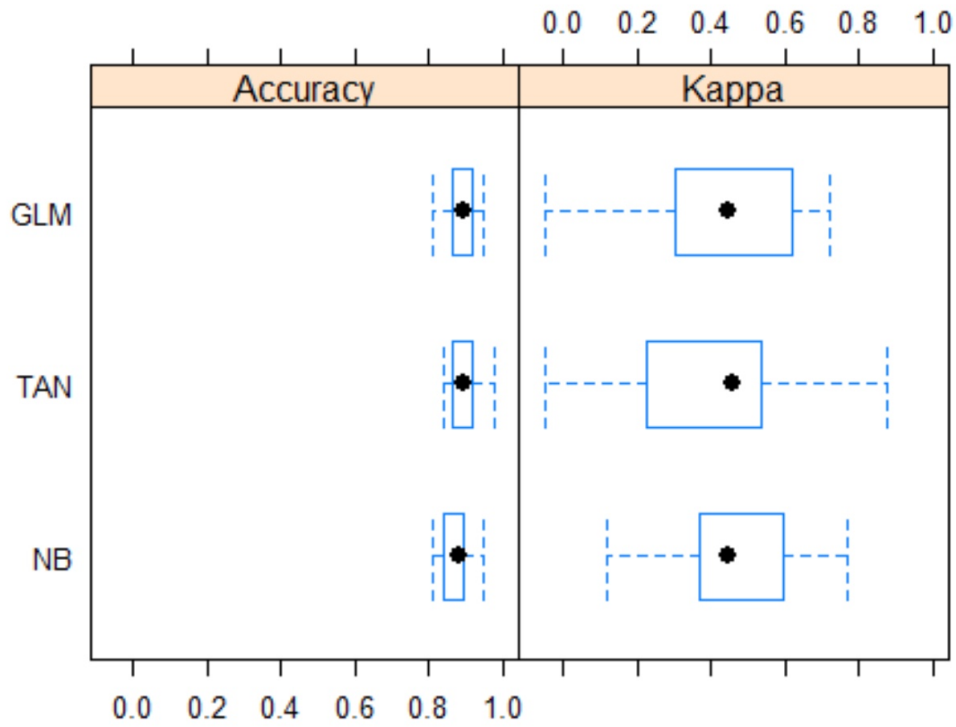


Figura 3.48. Comparación entre los diferentes modelos variando los conjuntos de entrenamiento/validación.

### 3.1.11. SELECCIÓN DE VARIABLES SIGNIFICATIVAS CON BORUTA

Se aplica Boruta al conjunto inicial de variables estadísticamente significativas repitiéndose el mismo proceso anterior.

#### 3.1.11.1. Regresión logística

Los coeficientes de la regresión logística son los siguientes:

Variables	Coficiente	Desviación estándar	Nivel de significancia
Sesgo	3,896	2,602	0,134
ProteinuriaIni	0,048	0,057	0,396
Proteinuria	-0,523	0,097	<b>7E-08</b>
Calcio	-0,116	0,267	0,663

Tabla 3.28. Coeficientes de regresión logística.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

Si se determina la curva ROC para este modelo y se determina el área bajo la curva (0.82), se comprueba que está al mismo nivel que el modelo univariante de regresión logística para las variables Proteinuria y ProteinuriaIni ( es decir, proteinuria en el seguimiento y proteinuria basal) (Figura 3.49).

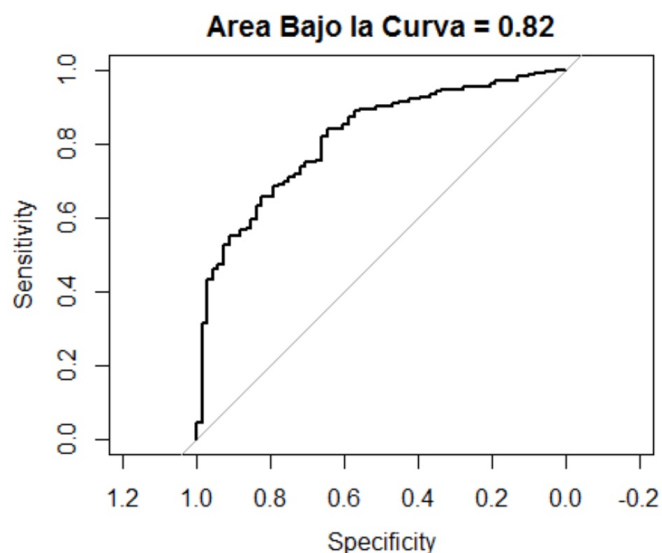


Figura 3.49. Curva sensibilidad/especificidad para el modelo lineal usando las variables seleccionadas por el Boruta a partir de las estadísticamente significativas.

Con esta curva se determina el umbral óptimo para discretizar las salidas del modelo. Con dicho umbral se determina la matriz de confusión para el conjunto de test obteniéndose los siguientes resultados:

		PREDICCIÓN	
		0	1
REALIDAD	0	24	4
	1	50	132

Tabla 3.29. Matriz de confusión para regresión logística.

**3.1.11.2. Naïve Bayes**

Repitiendo el procedimiento para el modelo Naïve Bayes con las tres variables seleccionadas se llega a:

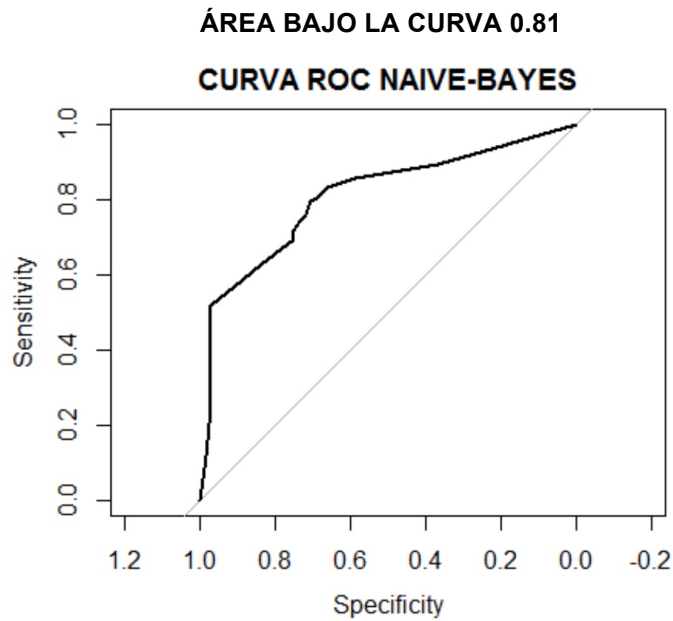


Figura 3.50. Curva sensibilidad/especificidad para Naive Bayes usando las variables seleccionadas por el Boruta a partir de las estadísticamente significativas.

El modelo presenta un área bajo la curva de 0.81. La matriz de confusión para el conjunto de test de este modelo (tabla 3.30) nos muestra que el modelo clasifica mejor el grupo 1 y peor el grupo 0 respecto al modelo lineal anterior.

		PREDICCIÓN	
		0	1
REALIDAD	0	21	7
	1	33	149

Tabla 3.30. Matriz de confusión para regresión logística tras selección mixta de variables.



3.1.11.3. TAN (Tree Augmented Network)

Al igual que en los casos anteriores se obtiene el modelo TAN para las variables seleccionadas.

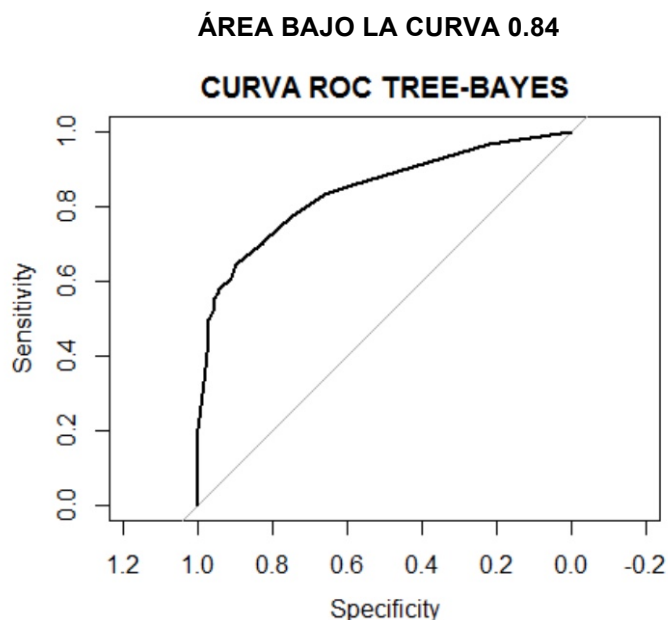


Figura 3.51. Curva sensibilidad/especificidad para TAN usando las variables seleccionadas por el Boruta a partir de las estadísticamente significativas.

El área bajo la curva es de 0.84; la matriz de confusión para el conjunto de test viene representada en la *tabla 3.31*, y nos muestra que este test presenta un comportamiento muy similar al del Naïve-Bayes.

		PREDICCIÓN	
		0	1
REALIDAD	0	21	7
	1	34	148

Tabla 3.31. Matriz de confusión para el modelo TAN.

Gráfico Tree.bayes

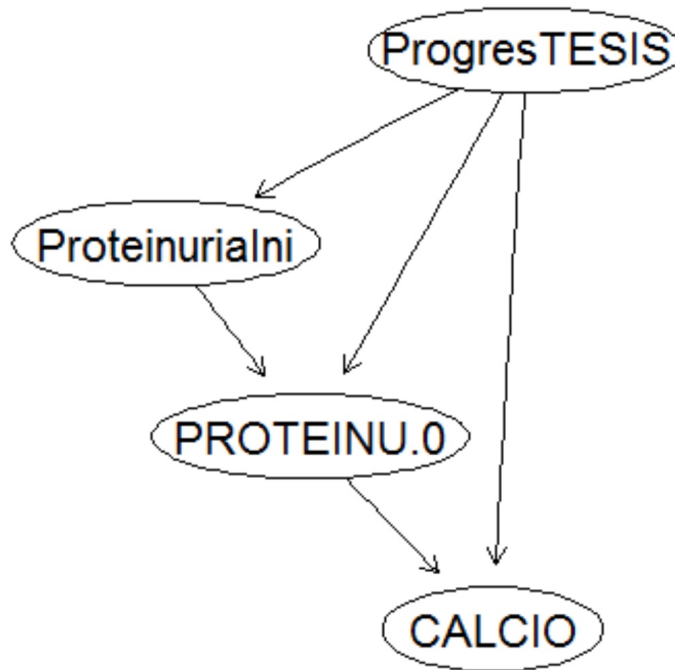


Figura 3.52. Esquema de las relaciones encontradas en el modelo TAN tras aplicar Boruta a las variables significativas.

3.1.12. ÁRBOL DE DECISIÓN

De los análisis anteriores se observa como variables significativas e importantes para el modelo de clasificación final las variables ProteinuriaIni, Proteinuria, Calcio y Creatinina. Con ellas se construye un árbol de decisión que queda reflejado en la *figura 3.53*.

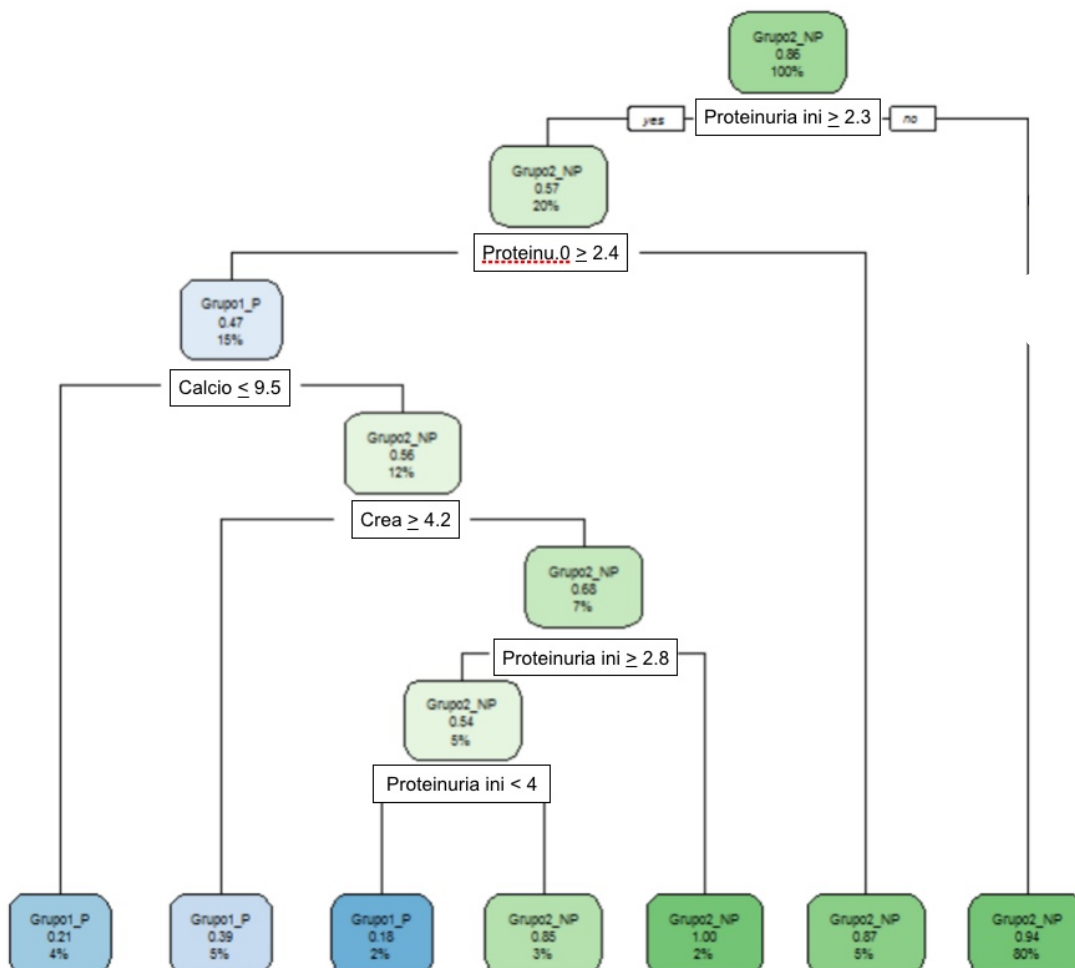


Figura 3.53. Árbol de decisión construido con las variables más importantes obtenidas en los modelos anteriores.

Este modelo tiene un área bajo la curva de 0.79 mostrándose la curva ROC en la figura 3.54.

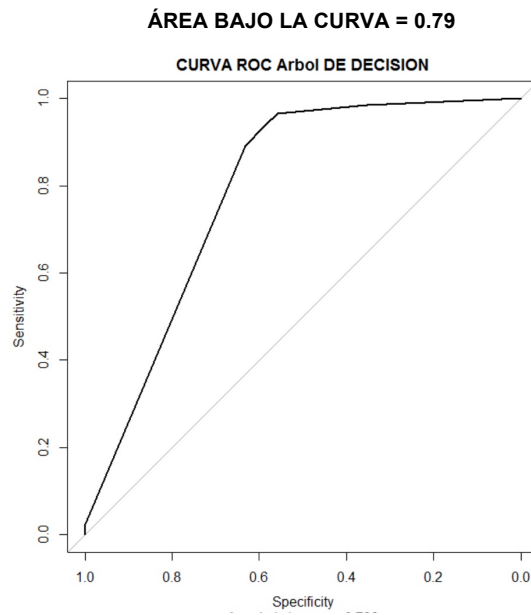


Figura 3.54. Curva sensibilidad/especificidad para el árbol de decisión desarrollado.

En cuanto a la matriz de confusión para el conjunto de test se tiene lo siguiente:

		PREDICCIÓN	
		0	1
REALIDAD	0	20	8
	1	27	155

Tabla 3.32. Matriz de confusión del árbol de decisión (0: Progresores rápidos, 1: Progresores lentos)

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Se observa un comportamiento similar a modelos anteriores; la ventaja de este modelo es que es explicativo y visual pero no mejora la precisión de otros modelos vistos anteriormente.

Este modelo nos proporcionaría una herramienta visual de fácil manejo en la clínica diaria, sin necesidad de realizar cálculos ni de disponer de una aplicación informática. Sólo conociendo estos datos básicos del paciente concreto (proteinuria, proteinuria basal, calcio y creatinina), y siguiendo el algoritmo, podemos predecir con bastante seguridad (AUC de 0.79) la probabilidad de que este paciente vaya a presentar una progresión rápida o no.

### 3.1.13. MAPAS AUTOORGANIZADOS

#### 3.1.13.1. Análisis global

Los mapas autoorganizados son un tipo de red neuronal que permite representar en un plano bidimensional, múltiples relaciones de asociación entre diferentes variables de forma simultánea. En la *figura 3.55* se han representado, mediante un mapa autoorganizado, las variables continuas analizadas, lo cuál nos permite establecer asociaciones entre todas ellas de forma simultánea. Este mapa consta de dos apartados, uno denominado de “ganadoras”, (*figura 3.55 a*) en el que se representa la distribución de los pacientes en las denominadas celdas o neuronas, que son cada una de las subdivisiones del mapa, y el apartado de componentes (*figura 3.55 b*). En cada una de estas celdas se pueden incluir varios pacientes o registros en un número variable, y todos ellos van a compartir las mismas características en cuanto a las diferentes variables analizadas, por lo que todas las celdas no están ocupadas por el mismo número de pacientes. También hay celdas vacías, lo que significa que en la población que manejamos no hay casos que presenten las características que teóricamente tendría un paciente que ocupase ese lugar. Por este motivo se deben evaluar ambos mapas de forma conjunta. El mapa de “ganadoras” se puede colorear para representar una variable concreta.

En la *figura 3.55 a* se ha representado el sexo de los pacientes, de forma que los coloreados en rojo son los hombres y en verde las mujeres. En algunas celdas hay pacientes de ambos sexos, por lo que podemos ver los dos colores de forma simultánea.

El otro apartado, *figura 3.55 b*, es el mapa de “componentes”, en el que se han representado las 16 variables continuas analizadas más relevantes. Cada paciente o registro representado ocupa la misma localización geográfica en cada una de las componentes, es decir, va a presentar las características que definirán a la celda que ocupa en cada uno de los componentes del mapa.

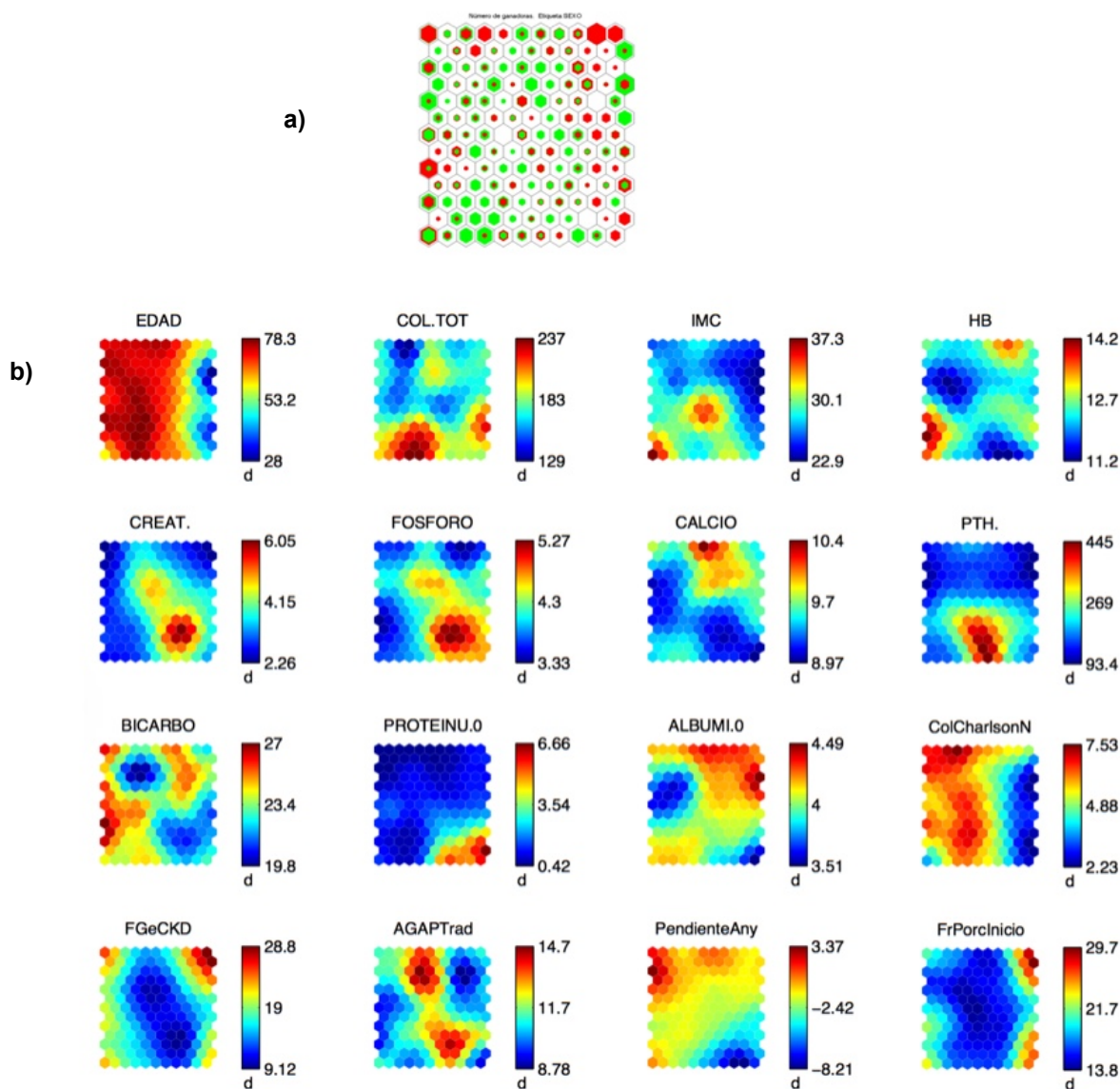


Figura 3.55. Representación mediante SOM. a) Mapa de "ganadoras", en el que se representa el sexo de los pacientes (Verde= mujeres, Rojo = hombres).

b) Mapa de "componentes", que permite establecer relaciones entre las diferentes variables en función de la posición geográfica. En cada una de las componentes en las celdas coloreadas en rojo representan el lugar que ocupan los pacientes con valores más elevados de la variable que se representa y que se indica en cada componente.

El componente titulado “pendienteAny” (Figura 3.56) representa la pendiente de descenso del FG. En esta imagen, los pacientes que ocupan las celdas de color azul, situadas en la esquina inferior derecha son los que han presentado un deterioro más rápido de la función renal.

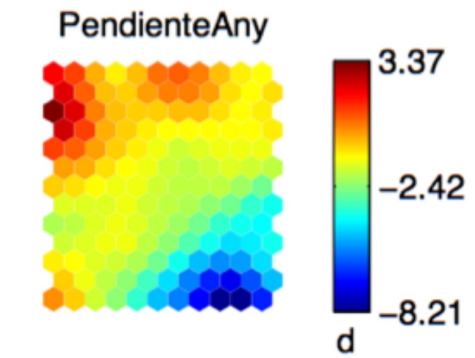


Figura 3.56. Representación mediante SOM de la pendiente de descenso del FG. Ampliación del componente del mapa que representa la pendiente de descenso del FG. La zona coloreada en azul corresponde a los progresores rápidos y se relaciona con la misma zona del resto de mapas de componentes.



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

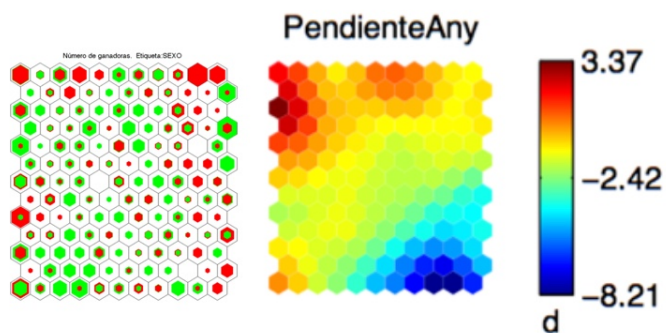
Dado que estos pacientes ocupan la misma posición geográfica en el resto de componentes del mapa, podemos deducir que los pacientes con una mayor progresión de la enfermedad renal:

- Tienen una edad comprendida entre 50 y 60 años
- Presentan un IMC de alrededor de 30.
- Presentan niveles más bajos de Hb.
- Cifras más elevadas de creatinina sérica.
- Tienen niveles más elevados de fósforo y PTH y más bajos de calcio.
- Niveles más bajos de bicarbonato en sangre y anión GAP más elevado.
- Proteinuria más severa.
- Filtrado glomerular al inicio del seguimiento menor de  $25 \text{ ml/min/1.73 m}^2$  aproximadamente , y más bajos durante la evolución (por debajo de  $20 \text{ ml/min/1.73 m}^2$ ).

Los pacientes con mayor patología asociada, es decir, con valores mayores en el índice de comorbilidad de Charlson, son pacientes más mayores con peores FG de inicio, pero presentan una menor progresión de la enfermedad renal.

Las variables categóricas se han representado de forma individual como mapas de “ganadoras”. Si las comparamos con el componente *Pendiente Any*, podemos deducir:

Los pacientes con una progresión más rápida de la enfermedad, es decir, los que ocupan la esquina inferior derecha, son de ambos sexos, puesto que en esa zona del mapa, las casillas están coloreadas en rojo (hombres) y verde (mujeres) en una proporción similar, por lo que se puede afirmar que en nuestra muestra, la progresión de la enfermedad renal no se ha visto influenciada por el sexo de los pacientes (*Figura 3.57*).



*Figura 3.57. Representación mediante SOM de la relación entre el sexo de los pacientes y la progresión de la enfermedad renal.*

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

En cuanto a la enfermedad de base, se objetiva que las nefropatías glomerulares se agrupan en las celdas que corresponden a un descenso más rápido del FG, y los pacientes con nefropatías vasculares, por el contrario, ocupan las celdas que corresponden con una función renal más estable. No hemos constatado mayor progresión de enfermedad renal en pacientes con nefropatía diabética, ni en los que presentan poliquistosis renal. Los pacientes con nefropatías glomerulares y progresión rápida, presentan además proteinuria más severa (Figura 3.58).

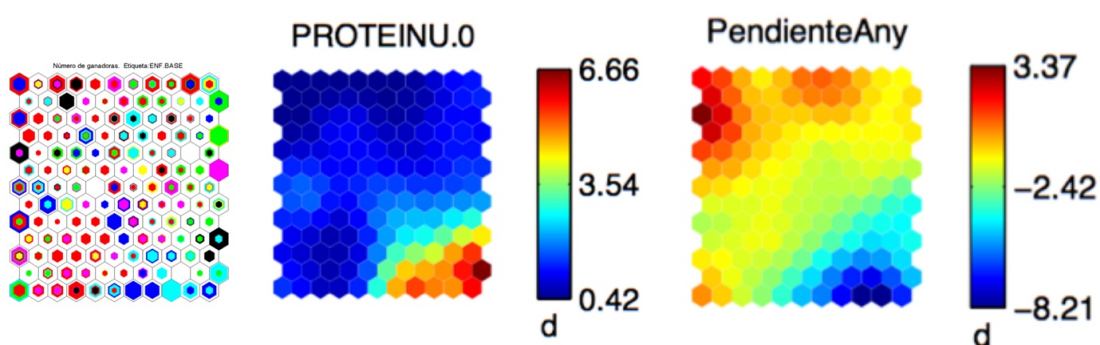


Figura 3.58. Representación de la relación entre el descenso del FG y la nefropatía de base. Codificación enfermedad de base: rojo: nefropatía vascular, verde: diabetes mellitus, azul: poliquistosis, cyan: nefropatías glomerulares, fucsia: nefropatía intersticial, negro: no filiadas, amarillo: otras.

Los pacientes con mayor progresión presentaban en su mayoría hipertensión arterial (casillas coloreadas en verde), aunque no mayor prevalencia de hipertensión refractaria (color azul) (Figura 3.59).

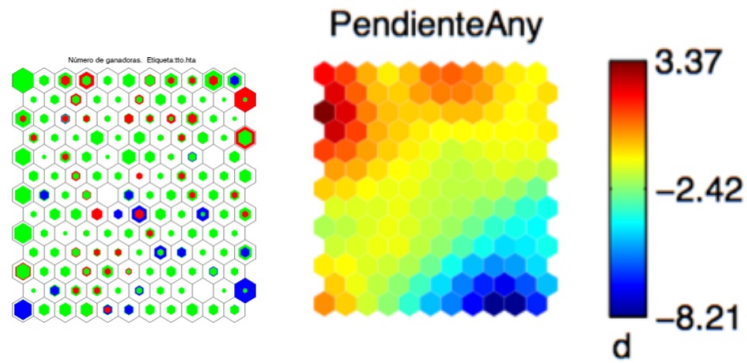


Figura 3.59. Progresión de enfermedad renal e HTA. Rojo: no HTA; verde: tratamiento con menos de 3 fármacos; azul: HTA refractaria).

No se aprecia relación entre la progresión de la enfermedad renal con el hábito tabáquico (Figura 3.60), ni tampoco con DM (Figura 3.61).

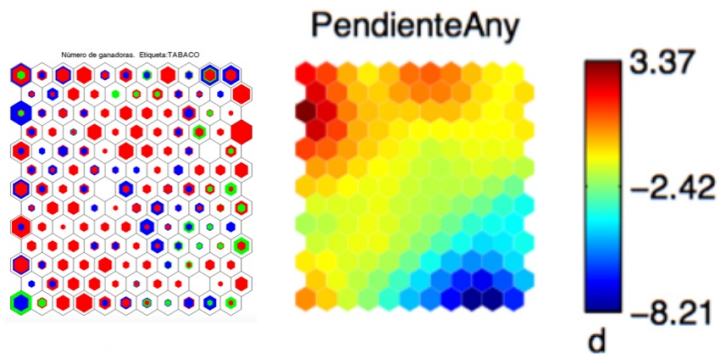


Figura 3.60. Tabaco y progresión. Rojo no fumador, verde exfumadores, azul fumadores.

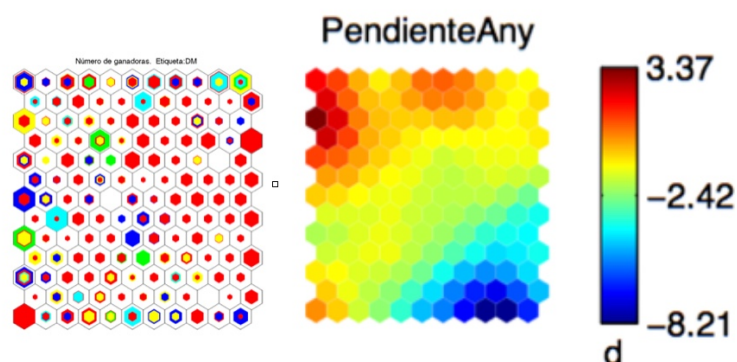


Figura 3.61. Relación entre DM y progresión. Rojo: no DM, cyan: controlados con dieta, amarillo: tratamiento con ADO, azul: insulina, verde: asociación de ambos tratamientos.

Los pacientes con una función renal más estable, es decir, con menor pendiente de descenso del FG, ocupan las celdas de la esquina superior izquierda del mapa en la *figura 3.56*, que están coloreadas en rojo. Estos pacientes son mayores de 60 años, sin diferencias en cuanto al sexo, con niveles más bajos de colesterol, tienen niveles medios de Hb y normales de bicarbonato en sangre, y más bajos de creatinina, fósforo, PTH y proteinuria. Los niveles de calcio han sido variables. Han presentado un índice de comorbilidad de Charlson más elevado. No hemos encontrado diferencias en cuanto a la presencia de diabetes mellitus ni en el hábito tabáquico, y han presentado un IMC por debajo de 27. Este grupo de pacientes presentaba unos valores de FG variables tanto al inicio del seguimiento como durante el mismo.

### **3.1.13.2. Lípidos y progresión de enfermedad renal**

Se ha generado un nuevo mapa para analizar la relación entre el perfil lipídico de los pacientes estudiados y la progresión de la enfermedad renal (*Figura 3.62*). Al crear un nuevo mapa cambia la distribución espacial en las diferentes celdas, por lo que el mapa de ganadoras en el que se representa el sexo de los pacientes adquiere un formato diferente al obtenido en el mapa del apartado anterior. En este caso, los pacientes que han presentado una progresión más rápida de la enfermedad renal, coloreados en azul, ocupan la parte lateral izquierda del componente titulado "PendienteAny".

Del análisis del mapa podemos concluir:

- Existe una relación directa entre los niveles de colesterol total y colesterol LDL y HDL en este grupo de pacientes. Estos niveles han sido más elevados en mujeres de edad más avanzada.

- Los pacientes con niveles elevados de TG han presentado también cifras elevadas de colesterol total y LDL, pero niveles bajos de colesterol HDL.

- Los pacientes con colesterol total y LDL más elevados tienen FG más bajos, sin que esta relación se cumpla a la inversa, es decir, no todos los pacientes con FG bajos tienen niveles bajos de colesterol, puesto que presentan cifras variables.

- La progresión más rápida de la insuficiencia renal se ha producido en pacientes con niveles elevados de colesterol total y colesterol LDL, cifras más bajas de triglicéridos y colesterol HDL.

- No encontramos relación entre la obesidad y la progresión de la enfermedad renal, puesto que los pacientes con IMC más elevados presentan gran variabilidad en cuanto al FG y pendiente de descenso del mismo, presentando además niveles medios de colesterol y triglicéridos. En cambio los IMC más bajos se relacionan con triglicéridos más bajos, FG más elevados y poca progresión de enfermedad renal.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

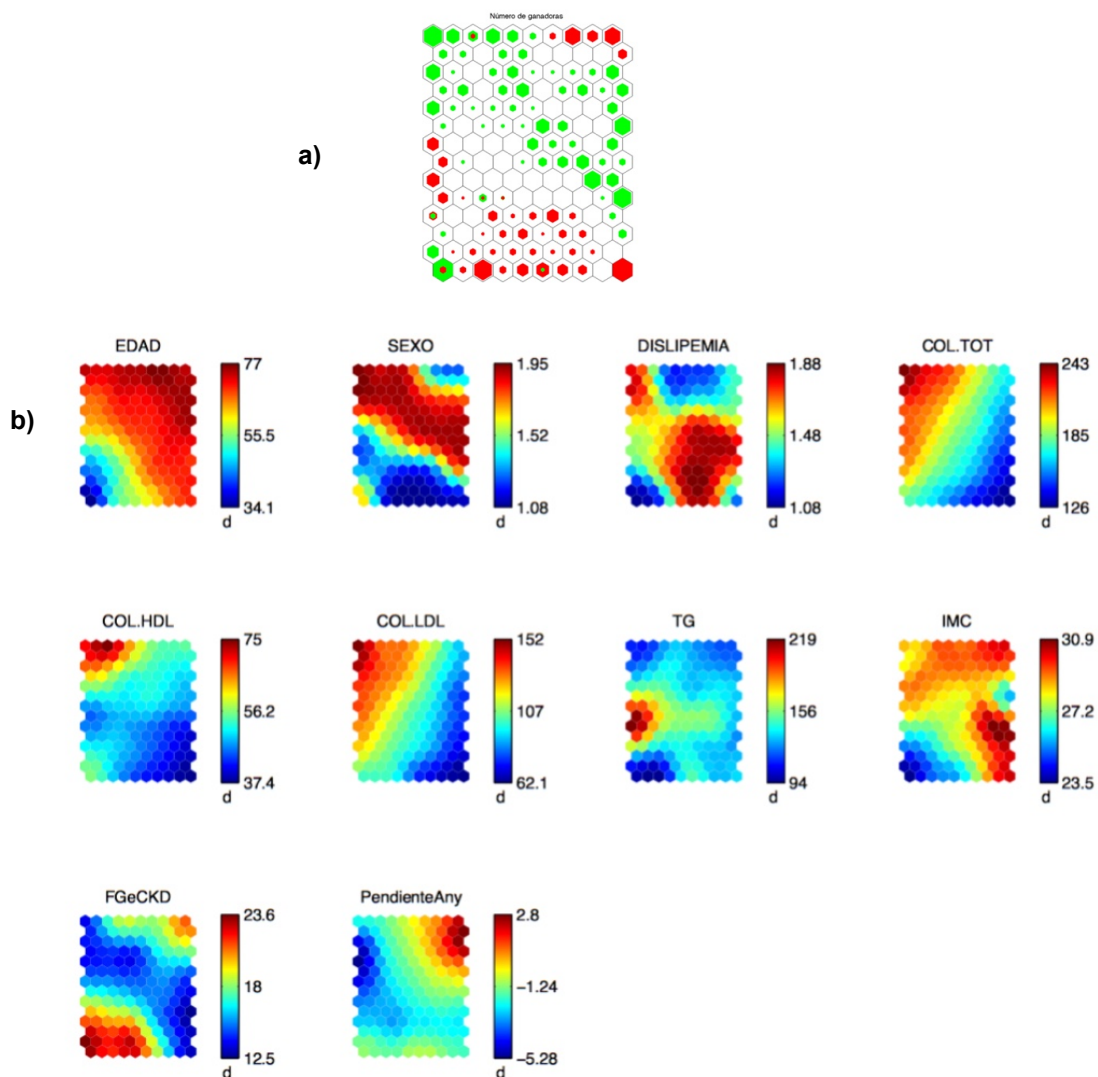


Figura 3.62. Relación entre alteraciones en perfil lipídico y progresión. a) Mapa de “ganadoras” coloreadas en cuanto al sexo (verde: mujer, rojo: hombre). b) Mapa de componentes: Edad (azul los pacientes más jóvenes y en rojo los más mayores), sexo (rojo: mujeres, azul: hombres, amarillo: celdas ocupadas por pacientes de ambos sexos), tratamiento con estatinas (dislipemia) (azul: No llevan, rojo: sí llevan, amarillo: celdas mixtas), niveles de colesterol total, HDL, LDL y triglicéridos, IMC, FG estimado por CKD-EPI y pendiente de descenso del FG.

- En cuanto al tratamiento con estatinas (etiqueta *Dislipemia* en la *Figura 3.62*) los pacientes que las reciben son mayoritariamente mujeres, de mayor edad, con IMC mayor de 25 y presentan niveles variables de colesterol total y LDL, mientras que muestran unos niveles más elevados de colesterol HDL. Estos pacientes han mostrado una mayor progresión de la enfermedad renal y mantenido unos FG bajos durante el seguimiento.

### **3.1.13.3. Anemia y progresión de enfermedad renal**

En la *figura 3.63* tenemos un nuevo mapa que representa la relación entre los niveles de hemoglobina en sangre, el metabolismo del hierro y su tratamiento.

- El grupo de pacientes con mayor progresión de enfermedad renal (esquina inferior derecha del mapa de componentes), presentó cifras más bajas de Hb a pesar de mantener valores de IST y ferritina dentro del rango de normalidad. El tratamiento con hierro no ha sido constante en este grupo de pacientes, y los que lo llevaban lo recibían en su mayoría por vía oral. Tampoco precisaban dosis muy elevadas de agentes eritropoyéticos.

- Los pacientes que han cursado con una progresión más lenta de la insuficiencia renal, presentaban niveles más elevados de Hb, tenían cifras de ferritina e IST dentro de la normalidad, precisaban dosis más bajas de epoetina, y no llevaban tratamiento con hierro, o bien lo recibían en forma de suplementos orales.

- Un pequeño número de pacientes recibía hierro endovenoso. Estos pacientes tenían un FG más estable, con una pendiente de descenso aproximada de  $-1,5 \text{ ml/min/año}/1,73 \text{ m}^2$ , y son los que presentaban mayor depleción en los depósitos de hierro, puesto que se corresponde con aquellos pacientes con niveles de IST y ferritina más bajos. También tenían cifras más bajas de Hb.



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

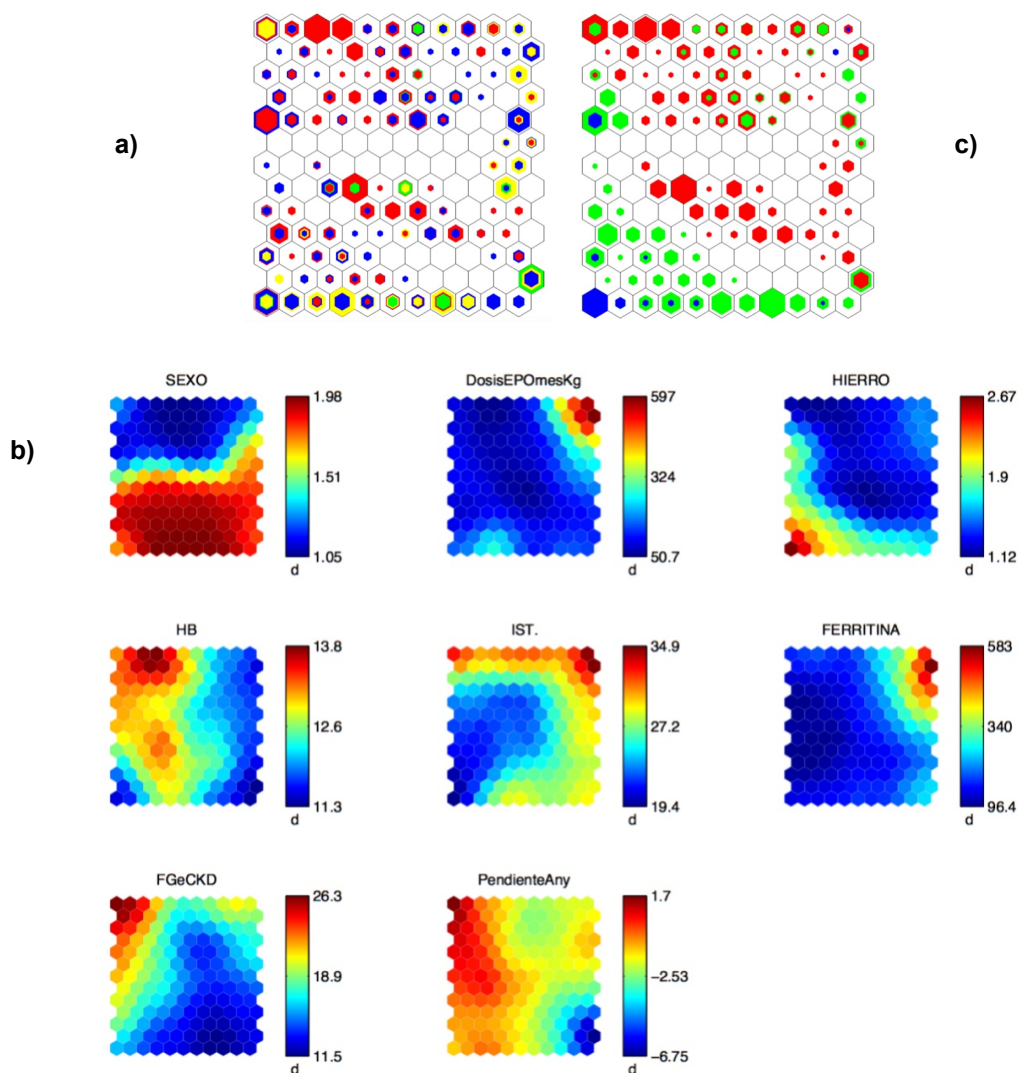


Figura 3.63. Relación entre anemia, su tratamiento y progresión. a) Distribución del tratamiento con derivados eritropoyéticos (Rojo: no tratamiento, Verde: epoetina alfa o beta, Azul: CERA, Amarillo: darbepoetina). b) Mapas de “componentes”: Sexo (Rojo: Mujeres, Azul: Hombres), DosisEPOmeskg, (Azul: dosis más bajas, Rojo: dosis más elevadas), Tratamiento con hierro (Azul: No reciben tratamiento, Rojo: Sí lo reciben), Niveles de Hb en sangre, IST, niveles de ferritina en sangre, FG estimado por CKD-EPI y pendiente anual de descenso del FG. c) Distribución del tratamiento con hierro (Rojo: No tratamiento, verde: Hierro oral, Azul: Hierro endovenoso).

- Los pacientes con dosis más elevadas de epoetina, en los que podríamos sospechar una resistencia a la acción de la misma, hemos encontrado que no presentaban ferropenia, incluso hallamos cifras de ferritina elevadas, lo que puede sugerir la presencia de un estado inflamatorio o de otras causas de resistencia a la acción de los agentes eritropoyéticos. Estos pacientes presentan cifras más bajas de Hb, y, no son los que sufren mayor progresión de enfermedad renal, sino que son pacientes con un FG durante el seguimiento de unos 18-20 ml/min y con un descenso del FG de unos 2 ml/min/año. Se aprecia además una relación directa entre el tratamiento con hierro y derivados eritropoyéticos, de manera que, en general, los pacientes que precisan tratamiento para la anemia secundaria a insuficiencia renal reciben los dos tratamientos de forma simultánea.

#### **3.1.13.4. Hipertensión arterial y progresión de enfermedad renal**

En el mapa de la *figura 3.64* podemos observar con más detalle la relación entre la hipertensión arterial y la progresión de la enfermedad renal en los casos estudiados.

- Los pacientes con un descenso más rápido del FG (celdas azules del componente *PendienteAny*) precisaban tratamiento para el control de la TA, recibían tratamiento con IECA-ARA II pero no con la asociación de ambos. No tuvieron más hipertensión refractaria, ni recibieron más tratamiento con diuréticos que el resto.

- Los pacientes con menor progresión de enfermedad renal (celdas rojas del componente *PendienteAny*), han presentado una gran variabilidad en cuanto al tratamiento antihipertensivo. Dentro de este grupo de pacientes se encuentran los que no han precisado medicación antihipertensiva ni diuréticos.

Los pacientes que han presentado hipertensión refractaria, es decir, que han precisado más de tres fármacos para su control, o que no se ha controlado con tratamiento, son pacientes que han presentado una caída del FG de unos -2 ml/min/1.73 m<sup>2</sup>/año, presentaban un IMC > 30 y recibían como parte del tratamiento antihipertensivo asociación de IECAs y ARA II. En este grupo de pacientes destacaba la nefropatía glomerular como la más frecuente en la etiología.

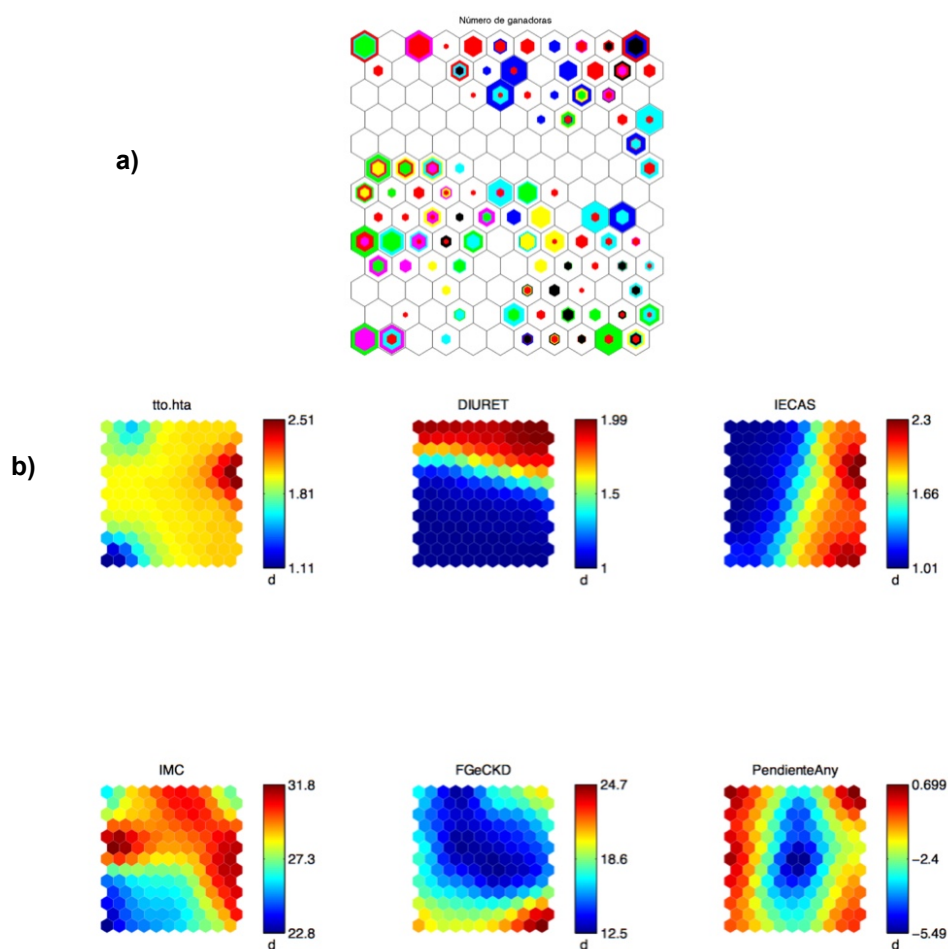


Figura 3.64. Relación entre HTA, su tratamiento y progresión de la enfermedad renal. a) Mapa de ganadoras que representa la distribución de los pacientes en función de la nefropatía de base (Rojo: nefropatía vascular, Verde: DM, azul: poliquistosis, cyan: nefropatías glomerulares, Fucsia: nefropatía intersticial, Negro: No filiadas, amarillo:Otras). b) Mapa de componentes: Tratamiento HTA (Azul: No tratamiento, amarillo: tratamiento con 3 o menos fármacos, Rojo: hipertensión refractaria), Diuréticos (Azul: No llevan, Rojo: Si llevan), IECAs (Azul: No llevan, amarillo-naranja: Tratamiento con IECAs o ARA II, Rojo oscuro: Asociación de ambos tratamientos, IMC, FG estimado por CKD-EPI, pendiente de descenso del FG en ml/min/1,73 m<sup>2</sup>/año

**3.1.13.5. Equilibrio acido-base y progresión de enfermedad renal**

Al analizar la relación entre la progresión de la función renal y el estado ácido-base de los pacientes, se ha encontrado asociación entre la presencia de niveles bajos de bicarbonato en sangre, los niveles más bajos de FG y mayor progresión de la insuficiencia renal, como se representa en la *figura 3.65*. Esta relación también existe en sentido inverso, es decir, los pacientes que han presentado una menor progresión de la enfermedad renal, han presentado niveles más elevados de bicarbonato. Estos pacientes, como se ha descrito previamente, tenían niveles variables de FG. También se ha encontrado una relación directa entre los niveles de bicarbonato en sangre y el anión GAP, de forma que los valores de anión GAP más elevados se han correspondido con niveles bajos de bicarbonato, y también con cifras bajas de FG y mayor pendiente de descenso del mismo.

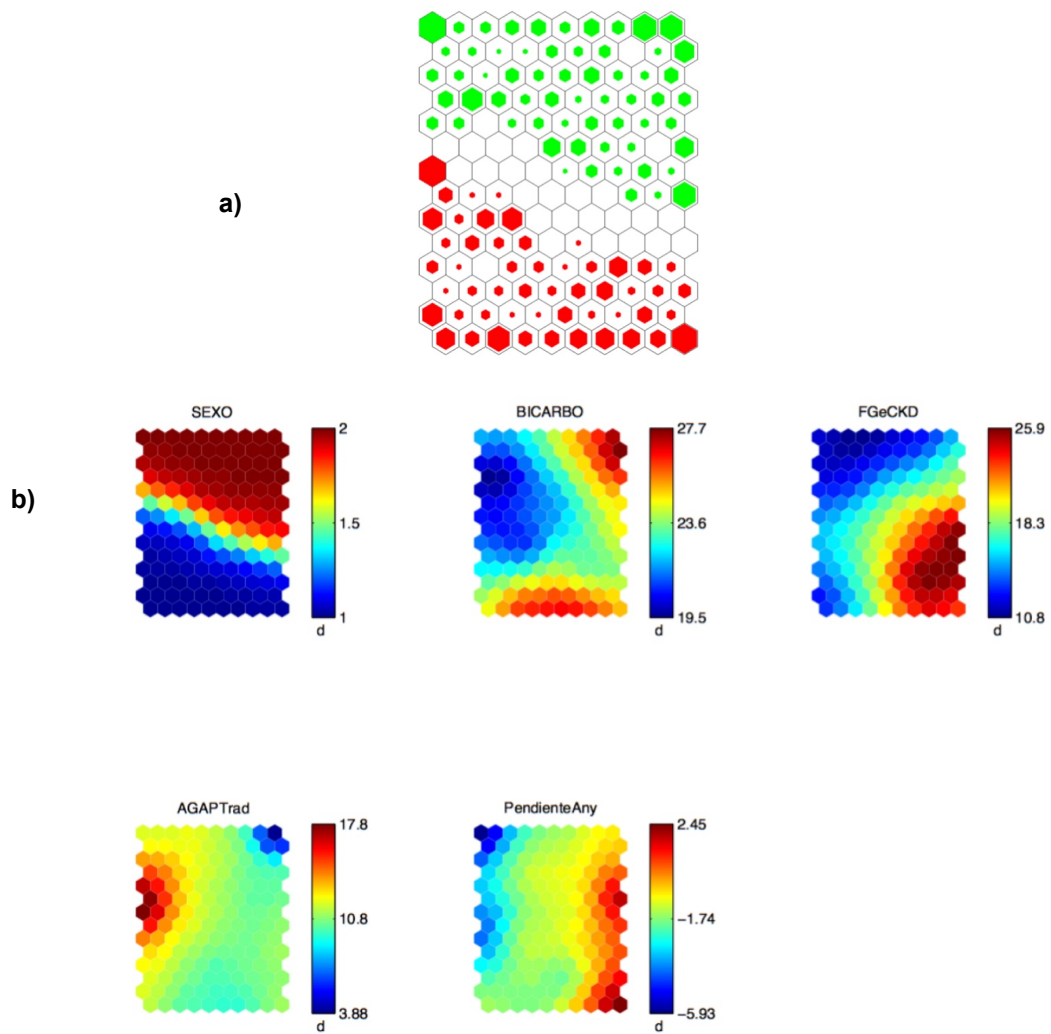


Figura 3.65. Relación entre equilibrio ácido-base y progresión de la enfermedad renal. **a)** Mapa de ganadoras con representación del sexo de los pacientes (Rojo: Hombres, Azul: Mujeres). **b)** Mapa de componentes: Sexo (Azul: hombres, Rojo: Mujeres), Niveles de bicarbonato en sangre, FG estimado por CKD, niveles de anionGAP y pendiente de descenso del FG.

**3.1.13.6. Metabolismo óseo-mineral y progresión de enfermedad renal**

En la *figura 3.66* se representa la evolución de la insuficiencia renal y su relación con el metabolismo óseo-mineral en nuestros pacientes.

- Se aprecia que, los que han presentado una progresión más rápida que además son los que han mantenido FG más bajos, como se ha descrito previamente, tienen los niveles más elevados de PTH sérica, así como niveles más bajos de vitamina D nativa.

- También este grupo de pacientes ha presentado los niveles más elevados de fósforo y las cifras más bajas de calcio en sangre.

- En este caso sí se confirma la relación inversa, es decir, los pacientes que han permanecido más estables en su evolución, son los que han presentado las cifras más bajas de PTH. Este grupo de pacientes ha presentado las cifras de calcio más elevadas y las más bajas de fósforo.

- Los pacientes que han recibido tratamiento con derivados activos de la vitamina D son los que presentan FG más descendidos, y mayor progresión de la enfermedad renal.

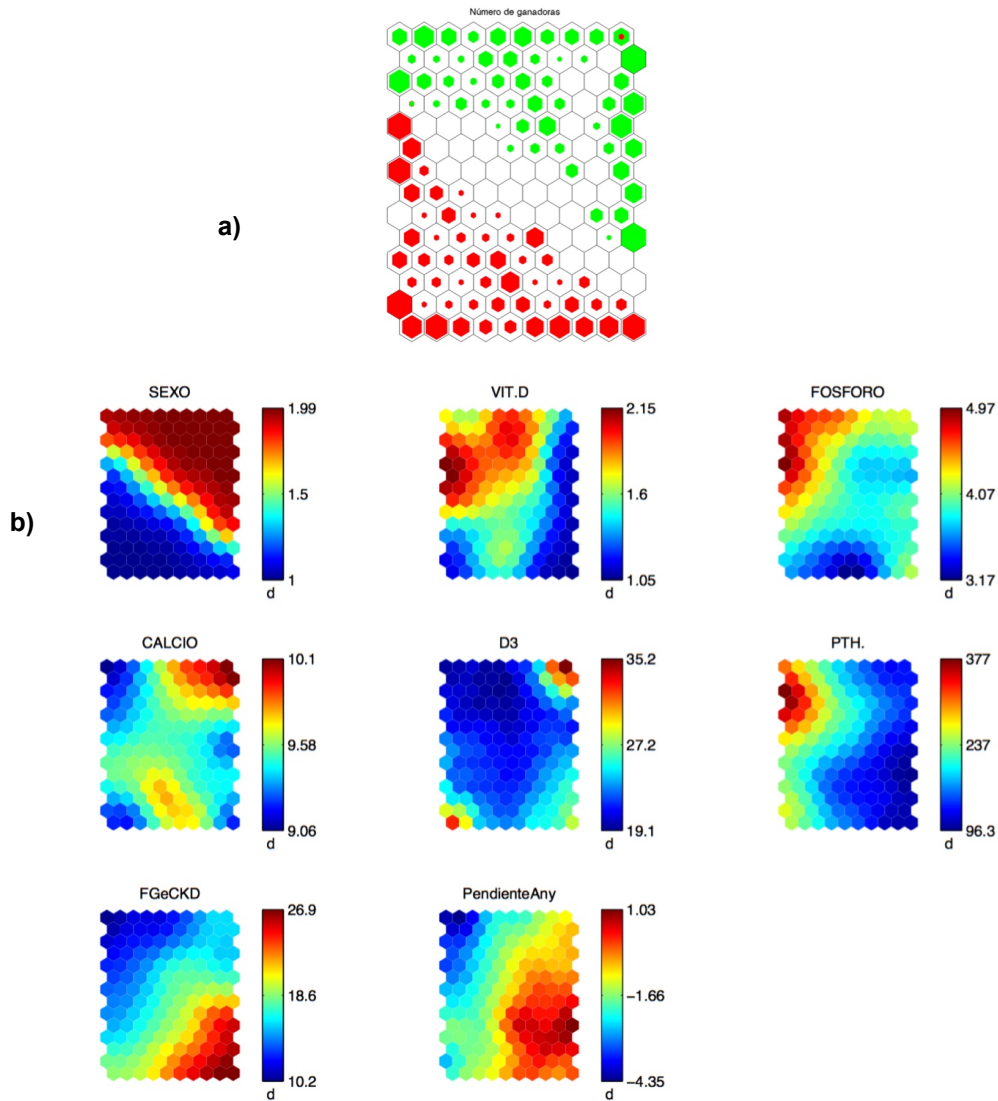


Figura 3.66. Relación entre alteraciones del metabolismo óseo-mineral y progresión de la enfermedad renal. **a)** Mapa de ganadoras en el que se ha representado el sexo de los pacientes (Rojo: Hombres, Verde: Mujeres). **b)** Mapa de componentes: Sexo (Azul: Hombres, Verde: Mujeres), Vit D (Tratamiento con derivados activos de la vitamina D. Azul: No tratamiento, Amarillo: calcitriol, Rojo: Paricalcitol), niveles de fósforo en sangre, calcemia, niveles de 25-OH-vit D, cifras de PTH, FG estimado por CKD-EPI, pendiente de descenso del FG.

### 3.2. DESARROLLO ARGUMENTAL

En el presente estudio se ha realizado un análisis retrospectivo sobre una población de pacientes que han sido controlados en la unidad de Enfermedad Renal Crónica Avanzada del Hospital La Fe. En estos pacientes se ha analizado la velocidad de progresión de la enfermedad renal y la influencia que han ejercido sobre ella diferentes factores tanto clínicos como analíticos considerados rutinarios en el manejo de este tipo de enfermos. A partir de ciertas variables que han mostrado su influencia sobre la progresión de la enfermedad renal en la parte inicial de este trabajo, se han desarrollado modelos predictivos que nos permitirán saber cuál va a ser la evolución que presentará un paciente concreto a partir de ciertos valores clínico-analíticos. Estos modelos predictivos se han realizado mediante procedimientos de aprendizaje máquina, que aportan más precisión que las técnicas estadísticas clásicas.

Asimismo, otro aspecto novedoso y que supone una aportación de esta tesis es la aplicación de mapas autoorganizados a nuestra base de datos, que ha permitido una caracterización minuciosa de los pacientes mediante el establecimiento de relaciones de asociación entre diferentes variables de forma simultánea. Estas asociaciones entre las diferentes variables se pueden hacer extensibles a otros pacientes con similares características. La caracterización de estos pacientes y la aplicación de modelos predictivos en la práctica clínica puede suponer una gran ventaja en cuanto que nos va a permitir incidir en ciertos aspectos propios de la enfermedad renal crónica, sobre todo en aquellos pacientes en los que se prevea una evolución más rápida, para intentar frenarla y, por tanto, retrasar el inicio del tratamiento sustitutivo. Además, vamos a poder planificar de una forma más adecuada el inicio de diálisis o incluso el momento de realización de un trasplante renal anticipado de donante vivo. Todo ello podría conllevar a un ahorro económico considerable, tratándose de medidas coste-efectivas. De hecho, las últimas guías KDIGO recomiendan el uso de modelos predictivos, pero la falta de modelos fácilmente aplicables y con una validación externa adecuada ha retrasado su integración en la práctica clínica, tanto en enfermedad renal crónica como en otros ámbitos de la Medicina.



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

El actual trabajo se ha basado en un estudio retrospectivo realizado sobre 115 pacientes. El período de tiempo de estudio ha sido diferente para cada paciente en función de la velocidad de progresión de la enfermedad, pero siempre ha sido mayor de seis meses para poder determinar la trayectoria de modificación del FG. También el número de registros (visitas) realizado para cada paciente ha sido variable, individualizándose en función de sus necesidades, tal como se realiza en la práctica clínica diaria. Otros trabajos, se han diseñado también de este modo (Tangri 2016, Rucci-2013)

En nuestra muestra se ha encontrado una media de descenso del FG global de  $-1,4 \text{ ml/min/1,73 m}^2/\text{año}$ . Para determinar esta pendiente de descenso, en cada caso se ha empleado un modelo de regresión lineal, como en otros trabajos publicados, por ejemplo el de Yuste *et al.* quienes también expresan esta velocidad de progresión como  $\text{ml/min/1,73 m}^2/\text{año}$  (Yuste 2013).

En publicaciones recientes, se ha descrito cómo la variación de la función renal en cada paciente sigue diferentes trayectorias, con una amplia variación inter e intraindividual, que dependerá de diferentes factores. En este sentido, Hu *et al.* describieron una progresión más rápida a la fase terminal de la enfermedad renal crónica en aquellos pacientes que habían padecido un episodio de fracaso renal agudo (Hu 2009). En grandes series de pacientes se han realizado estudios distinguiendo diferentes trayectorias de progresión de la enfermedad renal, así O'Hare *et al.* realizaron un estudio retrospectivo sobre más de 6000 pacientes y los dividieron en cuatro grupos en función de estas trayectorias, para analizar la influencia de los diferentes factores de riesgo en esta evolución. También Li *et al.* estudiaron en un grupo de 846 pacientes las diferentes trayectorias de progresión de la enfermedad renal y la posible influencia de diferentes factores en esta evolución. Este último autor publica posteriormente otro trabajo en el que estudia a 74 pacientes del grupo anterior que compartían un mismo modelo de progresión de la enfermedad renal que consistía en un período de rápido descenso seguido de una fase de estabilización del filtrado glomerular y comparaba factores de riesgo entre estas dos fases en un mismo paciente.

El presente estudio se basa en una muestra más reducida, por lo que se ha preferido estimar una trayectoria lineal para definir la evolución en cada paciente y poder realizar comparaciones.

Una vez definida esta trayectoria de descenso del FG para cada paciente, éstos se han dividido en progresores rápidos y progresores lentos en función de si esta trayectoria de descenso del FG ha sido mayor o menor de  $-5 \text{ ml/min}/1.73 \text{ m}^2/\text{año}$ , obteniendo 18 pacientes que han presentado una progresión más rápida frente a 97 pacientes que se han mantenido más estables. Por tanto, sólo un 15,6 % de pacientes ha presentado una progresión rápida de la insuficiencia renal, es decir, más de un 80 % de pacientes con enfermedad renal crónica con un FG basal menor de  $45 \text{ ml/min}$  se va a mantener estable o con un lento descenso del mismo. En otros estudios en los que se ha clasificado a los pacientes en función de la pendiente de descenso de la función renal, se ha empleado un criterio de corte diferente, así, en el trabajo publicado por Yuste *et al.* se clasifica a los pacientes en dos grupos de progresión en función de si la pendiente de evolución del FG es positiva o negativa, es decir, si existe una mejoría o un deterioro del mismo (Yuste 2013). En otro estudio retrospectivo realizado con pacientes con ERC estadio 3 se consideran progresores los que muestran una pendiente de descenso del FG menor de  $< 3 \text{ ml/min}/1,73 \text{ m}^2/\text{año}$ , y no progresores si dicha pendiente estaba entre  $-1$  y  $1 \text{ ml/min}/1,73 \text{ m}^2/\text{año}$  (Chase 2014). Marks *et al.* publican un trabajo en el que también analizan factores de progresión de enfermedad renal en 3322 pacientes con ERC grados 3 y 4, y definen progresión como el descenso del FG en  $15 \text{ ml/min}/1.73 \text{ m}^2$  o caída del mismo por debajo de  $10 \text{ ml/min}/1.73 \text{ m}^2$ . Estos autores también encuentran un porcentaje de pacientes progresores rápidos de un 15 %.

En el presente trabajo, se ha considerado el punto de corte entre ambos grupos de pacientes en una pendiente de  $-5 \text{ ml/min/año}/1,73 \text{ m}^2$ , de acuerdo con el criterio establecido por las guías KDIGO de 2012 y que posteriormente confirma la Sociedad Española de Nefrología (Documento de consenso de ERC).

Una vez definidos los grupos de progresión, se ha analizado la influencia de las distintas variables sobre la progresión de la enfermedad renal. Este análisis es complejo porque, entre otros factores, hay muchos menos ensayos clínicos

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

realizados en pacientes con enfermedad renal crónica que con otras patologías (Coresh 2016), siendo la propia insuficiencia renal un factor de exclusión en muchos de ellos, por lo que, al revisar la literatura, encontramos que las conclusiones no son definitivas en cuanto al efecto de muchos factores sobre la función renal en pacientes con ERC.

Para analizar la influencia de las diferentes variables en la evolución de la insuficiencia renal se han comparado las diferentes variables expuestas en el apartado de Resultados entre el grupo de pacientes progresores rápidos y los no progresores mediante contraste de hipótesis (*Wilcoxon*, *t de Student* para las variables continuas y Chi cuadrado y F de Fisher para las discretas). Se han identificado como factores de progresión: el menor tiempo de seguimiento en la consulta ( $p=1.15 \times 10^{-6}$ ), el grado de proteinuria mantenido a lo largo de la evolución ( $p=1.21 \times 10^{-5}$ ), la proteinuria basal, ( $p= 0.000157$ ), los niveles elevados de fósforo ( $p=0,00378$ ), los niveles bajos de calcio ( $p=0,0437$ ), la creatinina elevada ( $p=0,00437$ ) y la mayor pendiente inicial de descenso del FG ( $p=0.02$ ).

En otras variables se han encontrado diferencias que, aunque estrictamente no han alcanzado significación estadística, han quedado cercanas, como la edad (menor en el grupo de progresión rápida  $p=0.08$ ), niveles más elevados de colesterol LDL en el grupo de progresores rápidos ( $p=0.06$ ), PTH elevada ( $p=0.07$ ), bicarbonato bajo en sangre ( $p=0.06$ ), albúmina baja ( $p=0.06$ ), FG medio estimado por MDRD (0.05) y por CKD-EPI ( $p=0.09$ ).

El menor tiempo de seguimiento en la consulta, en este caso, más que un factor de progresión, debería ser considerado como una consecuencia de una evolución más rápida de la insuficiencia renal, que conllevaría un inicio más temprano de la terapia sustitutiva de la función renal y, por tanto, una menor duración del seguimiento en los pacientes con una mayor progresión. En nuestra muestra, no se puede atribuir este menor tiempo de seguimiento a una peor función renal al principio del seguimiento, puesto que no se han encontrado diferencias en el FG inicial ( $18.12 \pm 7.59$  ml/min en el grupo de progresores rápidos versus  $20.18 \pm 6.04$  ml/min en los progresores lentos,  $p=0.288$ ).

En el presente estudio se ha encontrado que los pacientes que presentan una mayor progresión de la enfermedad renal son más jóvenes, con una media de edad de 59 años en el grupo de progresores, frente a 70 años en los no progresores, aunque la diferencia no ha sido significativa ( $p=0.08$ ). Este hallazgo se ha descrito en otros trabajos (Tangri 2011, O'Hare 2012), y quizá podría explicarse por la mayor prevalencia de nefropatías vasculares en pacientes de mayor edad con una evolución más lenta, frente a otras nefropatías de evolución más rápida, hecho que no ocurre en pacientes más jóvenes.

No se han encontrado diferencias en cuanto al sexo de los pacientes. En este sentido, los resultados de otros trabajos son controvertidos, aunque en un metaanálisis reciente publicado por Tsai *et al.* en 2016, en el que se recogen datos de 319.000 pacientes en 40 estudios, se concluye que el sexo masculino constituye un factor de progresión en estadios avanzados de la enfermedad renal crónica. Estos resultados se han descrito previamente en otros metaanálisis, como el realizado por Neugarten *et al.* sobre 68 estudios y 11.345 pacientes, que mostró que el sexo masculino se asociaba con una progresión más rápida de la enfermedad renal en pacientes con ERC por nefropatía no diabética, en nefropatía IgA, nefropatía membranosa o poliquistosis hepatorenal. Sin embargo, en otros trabajos, las mujeres han presentado una mayor velocidad de progresión de la enfermedad, como el metaanálisis realizado por Jafar *et al.* sobre 1860 pacientes en 11 estudios, si bien, se trataba de mujeres postmenopáusicas, sin el posible potencial protector de los estrógenos (Jafar 2003).

Otra de las diferencias encontradas entre ambos grupos de pacientes son las cifras medias de creatinina para cada paciente, más elevadas en el grupo de pacientes con progresión rápida, siendo en estos de  $3.97 \pm 0,87$  mg/dl, frente a  $3.27 \pm 0.82$  en los no progresores ( $p = 0,04$ ). Este hallazgo se ha descrito en otros trabajos (Hsu 2009, Keane 2006), aunque más que un factor de progresión, podría considerarse una consecuencia del descenso del FG.

La causa más frecuente de enfermedad renal en estos pacientes han sido en primer lugar las nefropatías vasculares, seguidas de la diabetes y, en tercer lugar de las nefropatías glomerulares. A diferencia de los datos publicados por el Registro

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Español de Enfermedades Renales, la nefropatía diabética no es la causa más frecuente de enfermedad renal en nuestra muestra, ocupando el segundo lugar. No hemos encontrado relación entre la nefropatía de base y la progresión de la enfermedad renal, en contra de otros trabajos publicados (Hsu 2009, Rucci 2014, Haynes 2014). El estudio MDRD demostró una mayor progresión de enfermedad renal en pacientes con poliquistosis renal (Hunsiker 1997). Aunque en este trabajo no se incluyeron pacientes diabéticos, este dato se ha confirmado en estudios posteriores, como en el realizado por Haynes *et al.* a partir del estudio SHARP (Haynes 2014).

En cuanto a la patología asociada, los pacientes han presentado un índice de comorbilidad de Charlson de  $5.3 \pm 2.0$  (media  $\pm$  ds), sin encontrar diferencias entre los progresores rápidos y lentos.

No hemos encontrado diferencias en ambos grupos de pacientes en cuanto a la presencia de DM, ni tampoco en los niveles medios de glucemia ni de Hb glicosilada. Aunque en la mayoría de trabajos realizados se confirma que la diabetes mellitus ejerce un papel importante en el desarrollo de la enfermedad renal (Branchati 97, Lorenzo 2007, Stanton 2014, DCCT/EDIC 11, Adler 2003), su efecto como factor de progresión una vez que existe deterioro de función renal no está tan claro (Yuste 2013, Tsai 2016), por lo que los resultados de nuestro trabajo no difieren de los publicados en la literatura.

No hemos podido demostrar relación de dependencia entre la progresión de la enfermedad renal y la presencia de hipertensión arterial, ni tampoco con el tratamiento de la misma, puesto que no se han encontrado diferencias en este sentido entre los pacientes progresores rápidos y los progresores lentos. A pesar del conocido papel de la hipertensión arterial en la progresión de la enfermedad renal, en algunos estudios y revisiones tampoco se ha confirmado esta asociación (Yuste 2013, Tsai 2016). En el presente trabajo podría justificarse este hecho porque la gran mayoría de pacientes son hipertensos (88,5 %), lo que dificulta la valoración de este factor en la progresión.

Los niveles de proteinuria, el factor de progresión de enfermedad renal más reconocido en la actualidad, han sido muy diferentes en los dos grupos de pacientes

(0,73 +/- 1,13 g/día en el grupo de no progresores, frente a 3,28 +/-1,98 g/día en los progresores,  $p = 2,57 \times 10^{-0,5}$ ), por lo que en el presente trabajo se confirma su papel como factor de progresión.

Se ha encontrado diferencia significativa en los niveles de fósforo de ambos grupos de pacientes, siendo más elevados en el grupo de pacientes con progresión rápida de la enfermedad renal ( $4,44 \pm 0,53$  vs  $4,01 \pm 0,53$ ,  $p = 0,00378$ ). Este resultado es congruente con los resultados de múltiples trabajos publicados en los que se sugiere que la hiperfosforemia es un factor de progresión de la enfermedad renal, aunque queda por determinar su relación de causalidad, puesto que hasta la fecha, no se han publicado estudios controlados que hayan demostrado un enlentecimiento en la progresión de la insuficiencia renal tras reducir los niveles de fósforo mediante quelantes (Collister 2016).

También se han encontrado niveles de calcio más bajos en el grupo de pacientes progresores rápidos ( $9,3 \pm 0,5$  vs  $9,59 \pm 0,4$  mg/dl,  $p = 0,04$ ). Son pocos los trabajos que han asignado al calcio sérico un papel en la progresión de la enfermedad renal. Tangri *et al.* crearon un modelo predictivo de evolución hacia insuficiencia renal terminal mediante el desarrollo de varias fórmulas y afirman que el calcio bajo, junto con otros factores, puede predecir un mayor riesgo de fallo renal en pacientes con ERC estadios 4 y 5 (Tangri 2011). Posteriormente, Chase *et al.* realizan un estudio retrospectivo en pacientes con ERC estadio 3 en el que comparan pacientes con progresión rápida y lenta de la enfermedad renal, y comprueban que los pacientes con peor evolución tenían cifras de calcio más bajas al inicio del estudio (Chase 2014). Sin embargo, en estudios y metaanálisis publicados posteriormente, no se ha encontrado esta asociación, y mucho menos se ha podido demostrar una relación de causalidad.

En cuanto a los niveles de PTH no se ha llegado a encontrar diferencia significativa en los niveles en los pacientes progresores y los no progresores ( $p=0,07$ ), al contrario de lo publicado por Yuste *et al.*, quienes encontraron asociación entre los niveles elevados de PTH y la mayor progresión de enfermedad renal (Yuste 2013). En realidad son pocos los trabajos que hayan analizado la relación entre los niveles de PTH y la progresión de la enfermedad renal.

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

No hemos encontrado asociación entre los niveles de 25-OH-vit D3 con el FG ni con la pendiente de descenso del mismo, pues no han habido diferencias en ambos grupos de pacientes. Tampoco el tratamiento con derivados activos de la vitamina D ni con suplementos de vit D nativa ha mostrado diferencias en ambos grupos.

Los niveles de bicarbonato en sangre obtenidos han sido menores en el grupo de progresión más rápida de la insuficiencia renal, aunque no se ha llegado a alcanzar significación estadística ( $22.53 \pm 2.62$  frente a  $23.83 \pm 2.33$  mmol/l, p 0,06). Probablemente, este hecho puede justificarse porque muchos de los pacientes con enfermedad renal crónica reciben suplementos de bicarbonato sódico (en el presente estudio llevaban este tratamiento en un 41 % de los registros realizados).

En cambio, en contra de otros estudios publicados (Togawa 2013), en los que se ha encontrado relación entre el incremento del anión GAP y la progresión de la enfermedad renal, sugiriendo incluso que este incremento supone un factor de progresión de la enfermedad, no hemos encontrado relación entre el anión GAP calculado mediante cada una de las fórmulas indicadas entre los pacientes progresores rápidos y los progresores lentos. En cuanto al tratamiento de la acidosis, no se ha encontrado relación entre el tratamiento con bicarbonato y la progresión de la enfermedad renal, puesto que no ha habido diferencia entre ambos grupos de pacientes.

Si comparamos los dos grupos de pacientes (progresores rápidos y lentos), no se han encontrado diferencias en la cifra de Hb, ni tampoco en las dosis de epoetina utilizadas, ni en los parámetros de metabolismo férrico. La Hb durante el seguimiento ha sido de  $12,3 \pm 0.1$  g/dl para los pacientes con una progresión más rápida frente a  $12,06 \pm 1.07$  en los pacientes no progresores (p 0,3). En cuanto al tratamiento con agentes eritropoyéticos, los pacientes lo han recibido en un 46,3 % de los registros realizados. Las dosis de epoetina administradas han sido:  $205.31 \pm 134.59$  UI/kg/mes en los pacientes más estables, frente a  $258.32 \pm 139.3$  UI/kg/mes en los pacientes con progresión más rápida de la enfermedad renal (p 0,2).

Estos resultados no nos permiten confirmar en nuestros pacientes el papel de la anemia en la progresión de la enfermedad renal. Este hecho podría justificarse

porque durante el seguimiento los pacientes han recibido tratamiento con hierro y derivados eritropoyéticos, que han impedido llegar a alcanzar cifras de Hb excesivamente bajas, pero tampoco ha habido diferencias al considerar los niveles de Hb al inicio del seguimiento, cuando la mayoría de pacientes no habían sido tratados. Tampoco la literatura ha clarificado el papel de la anemia en la progresión de la enfermedad renal, pues aunque estudios iniciales parecían demostrar su papel en la progresión de la enfermedad renal, metaanálisis recientes y estudios prospectivos comparativos entre pacientes progresores y no progresores, no han confirmado a la anemia como factor de progresión de enfermedad renal (Tsai 2016, Li 2014, Chase 2014).

En el presente estudio los pacientes han presentado unos niveles medios de ácido úrico en sangre de  $7.13 \pm 1.39$  mg/dl, y han recibido tratamiento con alopurinol en un 31.9 % de los registros realizados, pero no se han encontrado diferencias en ambos grupos de pacientes en cuanto a los niveles de ácido úrico ( $p=0.7$ ) ni al tratamiento con alopurinol.

En el presente estudio no se han encontrado diferencias en los niveles de colesterol total, HDL y triglicéridos entre ambos grupos de pacientes, aunque en el caso del colesterol LDL se ha rozado la significación estadística ( $p = 0,06$ ). El tratamiento con estatinas tampoco ha mostrado ninguna asociación con la progresión de la enfermedad renal, ni se han encontrado diferencias entre los dos grupos de pacientes. Estos resultados son congruentes con los publicados en la literatura, que no han podido confirmar el papel de la dislipemia y su tratamiento en la progresión de la enfermedad renal cuando se parte de un  $FG < 60$  ml/min (Kalaitzidis 2011, Baigent 2011, Rahman 2014).

En nuestra muestra hemos encontrado diferencias en los niveles de albúmina sérica entre el grupo de pacientes progresores rápidos ( $4.1 \pm 0,28$  g/dl) y progresores lentos ( $4.06 \pm 0.39$ ), pero sin alcanzar significación estadística ( $p=0.06$ ). En algunos trabajos se ha asignado a los niveles bajos de albúmina un factor pronóstico en la progresión de la enfermedad renal (Tangri 2011, Chase 2014), pero en otros no se ha podido demostrar esta asociación.



### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

Tampoco hemos encontrado asociación en cuanto al uso de diuréticos o de bloqueantes del sistema renina-angiotensina-aldosterona (BSRAA) con la progresión de la enfermedad renal.

Una vez analizados los resultados obtenidos en el contraste de hipótesis, mediante un modelo de regresión logística univariante se ha analizado el valor predictivo de las variables que han resultado significativas en el contraste de hipótesis. Se ha encontrado, de acuerdo con las publicaciones existentes, que la proteinuria (tanto basal como la media obtenida en el seguimiento) es un buen predictor de progresión de la enfermedad renal, con un AUC de 0.83. Los valores de creatinina y fósforo en sangre también han sido significativos, pero en cambio, ninguno de los dos se ha considerado un buen factor predictor de progresión, pues el AUC para cada uno de ellos es de 0.6.

Después del análisis univariante se ha realizado un análisis multivariante de los datos, con diferentes técnicas, con el objeto de desarrollar un modelo predictivo de evolución de la enfermedad basal, a partir de algunas variables consideradas significativas. Este análisis multivariante se ha realizado mediante modelos clásicos (regresión logística) y mediante modelos de aprendizaje máquina.

Hasta la fecha, la mayoría de modelos predictivos de progresión de enfermedad renal publicados están basados en modelos de regresión multivariante de Cox (Keane 2006, Johnson 2008, Hallan 2009, Tangri 2011, Coresh 2014), y únicamente se ha validado en otras poblaciones demostrando su validez el de Tangri (Tangri 2017). Coresh *et al.* publican en 2014 un modelo predictivo de progresión de enfermedad renal para el que utilizan *Naive-Bayes*, además de regresión logística, pero no se han desarrollado más modelos predictivos con estas técnicas.

El modelo de regresión logística multivariante se ha realizado a partir de todas las variables significativas en el contraste de hipótesis, excepto el tiempo de seguimiento, que no hemos incluido por considerarlo un equivalente a la progresión. Se ha concluido que es un buen modelo predictor de progresión de enfermedad renal, pues ha mostrado un AUC de 0,8 y la matriz de confusión nos muestra que este modelo tendría capacidad para acertar un 75 % de casos en pacientes con progresión rápida y un 85 % en progresores lentos.

---

Posteriormente se han realizado el análisis multivariante utilizando las mismas variables pero mediante técnicas de aprendizaje-máquina (Naive-Bayes y TAN) que, como se ha explicado previamente, son modelos probabilísticos, en los que, a partir de unas variables y, mediante una serie de cálculos matemáticos, se puede determinar la probabilidad de un suceso concreto, que en nuestro caso sería la progresión de la enfermedad renal. La diferencia principal entre estos dos modelos, es que el Naive-Bayes considera que todas las variables son independientes entre sí, y el TAN asume posibles relaciones de dependencia entre las diferentes variables. En esta parte del trabajo los dos modelos han mostrado un AUC ligeramente superior al modelo de regresión logística. El TAN mejora la capacidad de predicción en los progresores rápidos que sería de un 82 % de aciertos, pero empeora ligeramente la probabilidad de acierto en los no progresores, que pasa a ser de un 79%. Las tres técnicas han mostrado un índice *kappa* reducido, lo cuál nos dice que el modelo no presenta un buen funcionamiento.

Para intentar mejorar la capacidad predictiva de los modelos, se ha realizado una nueva selección de variables por una técnica de aprendizaje-máquina (Boruta). Entre las variables seleccionadas se han obtenido algunas coincidentes con las previas (elegidas por mostrar significación estadística en el contraste de hipótesis), como la proteinuria media y la basal, y otras nuevas, como el FG basal y la PTH basal. Llama la atención la elección de esta última variable por el sistema, pues no se encuentra entre las que se han asociado clásicamente con la progresión de la enfermedad, aunque está asociada a grados avanzados de la ERC. El hecho de que se haya dado más importancia al valor basal que al medio en el seguimiento, obedece probablemente a que es en los primeros controles donde la PTH suele estar más elevada, y posteriormente sus niveles pueden estar más controlados con las medidas terapéuticas que se instauran. En cualquier caso, la PTH tiene menos peso en el modelo predictivo pues tiene un coeficiente más bajo que el resto de variables.

Con las variables obtenidas por este método se han vuelto a repetir los modelos predictivos descritos, encontrando un mejor funcionamiento de los tres. En el modelo de regresión logística se ha obtenido un AUC de 0.93, pero sin embargo la matriz de confusión nos muestra que la capacidad de acierto en los pacientes

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

progresores rápidos sería sólo de un 66 %. El *Naïve-Bayes* ha logrado un AUC de 0,89 y una mejoría en la matriz de confusión, con un 84.7 % de probabilidad de acierto en el caso de los progresores y de un 90.4 % en los no progresores. El TAN todavía ha mejorado estos parámetros, con un AUC de 0.93 y una probabilidad de acierto en los pacientes de progresión rápida de un 100 %.

De todos estos modelos ensayados, vemos que los mejores resultados se obtienen con el TAN tras selección de variables por Boruta. Con ello generamos un modelo, en el que, a partir del valor en un paciente concreto de las variables proteinuria al inicio del seguimiento, proteinuria media, FG basal, PTH basal y variación de la PTH en los dos primeros controles, podremos saber si el paciente va a ser progresor rápido o lento. Para que el modelo sea aplicable en la clínica diaria, una vez creado y aprobada su validez, como es el caso, se puede diseñar una aplicación informática para su utilización, de manera que sólo tendríamos que introducir estos valores que obtenemos en la analítica rutinaria en estos pacientes.

Tras comprobar el mejor funcionamiento de los modelos utilizando la selección de variables por Boruta, parece lógico pensar que estas herramientas pueden ser de gran utilidad en la selección de variables para el diseño de modelos predictivos.

Con estas variables, se ha construido además un árbol de decisión fácilmente aplicable en la clínica diaria y se ha representado en el capítulo de Resultados. Rucci *et al.* diseñaron un árbol similar a partir de la proteinuria, FG basal, fósforo, edad, género y presencia de DM, pero, a diferencia del presente trabajo, recoge datos de pacientes en todas las categorías de ERC y, además, está basado en el análisis de los datos de un registro en el que se recogen todos los pacientes que se remiten a Nefrología en una región de Italia, recogándose aquellos con más de cuatro determinaciones analíticas en un período de tiempo, lo que puede suponer una situación de inestabilidad.

El empleo de estos árboles de decisión pueden ser una herramienta muy útil en la práctica diaria, pues son muy fáciles y rápidos de interpretar, y nos pueden ayudar en la toma de decisiones.

Además de desarrollar estos modelos predictivos, hemos utilizado mapas autoorganizados para caracterizar nuestra muestra de pacientes.

Los mapas autoorganizados (*SOM*) se están empezando a utilizar en la actualidad en diferentes campos de la Medicina, sobre todo a nivel experimental, aunque su uso todavía no está extendido. Existen pocos trabajos sobre su aplicación a nivel clínico, pero pueden abrir un campo muy interesante por su capacidad para establecer relaciones de asociación de forma simultánea entre múltiples variables. Por ejemplo, recientemente se ha publicado un trabajo en el que se desarrollan unos *SOM* para predecir y diagnosticar la recidiva precoz de carcinoma de células renales tras nefrectomía, basándose en niveles en suero de diferentes marcadores metabólicos (Zheng 2016).

En el presente trabajo, se han creado varios *SOM* para analizar la asociación simultánea de diferentes variables sobre la progresión de la enfermedad renal y, aunque no se puede establecer una relación de causalidad entre las diferentes variables, sí que es posible extraer conclusiones que no nos permiten otras técnicas de estadística más clásicas. Del mapa general, en el que se representan 16 variables, concluimos que los pacientes con mayor progresión de enfermedad renal, tienen una edad comprendida entre 50 y 60 años, un IMC de alrededor de 30, cifras más elevadas de creatinina, fósforo, PTH y anión GAP, y más bajas de Hb, calcio y bicarbonato. También han presentado proteinuria más severa y albúmina más baja. No han mostrado diferencias en cuanto al sexo. Hasta aquí los resultados son congruentes con los obtenidos en la estadística clásica, lo que apoya la validez de estas técnicas. Además, extraemos nuevas conclusiones procedentes de las relaciones de asociación establecidas mediante estos mapas.

Al considerar la enfermedad de base del paciente, con la estadística clásica, no hemos encontrado relación con la progresión de la enfermedad renal, sin embargo, mediante las redes neuronales encontramos que los pacientes con nefropatías glomerulares como enfermedad de base, han presentado una progresión más rápida de la enfermedad renal y, en cambio, las nefropatías vasculares e intersticiales son las que han presentado una evolución más lenta. Los pacientes con nefropatía diabética y poliquistosis renal como enfermedades de base han mostrado

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

una evolución variable de la enfermedad renal. La progresión más rápida de la insuficiencia renal en los pacientes con nefropatías glomerulares, viene justificada por su mayor grado de proteinuria.

En cuanto a la patología asociada, no encontramos ninguna asociación con la progresión de la enfermedad mediante las técnicas de estadística clásica, sin embargo, en el mapa autoorganizado vemos que los pacientes con mayor índice de comorbilidad de Charlson han presentado una progresión más lenta de la enfermedad renal, a pesar de que tienen filtrados glomerulares más bajos al inicio del seguimiento. Son pacientes de edad más avanzada y con menor grado de proteinuria.

Del análisis de la relación entre el metabolismo lipídico y la progresión de la enfermedad renal crónica, no pudimos encontrar ninguna relación en el análisis univariante como se ha comentado anteriormente, salvo en el caso del colesterol LDL en que se rozaba la significación estadística ( $p=0.06$ ). Sin embargo, los mapas SOM nos aportan información añadida, de la que destaca la relación directa existente entre los niveles de colesterol total, HDL y LDL, así como la asociación entre los niveles de colesterol total y LDL con la progresión rápida de la enfermedad renal, aunque sin poder establecer relación de causalidad. En cuanto al tratamiento con estatinas, sí que encontramos una clara asociación entre la presencia de este tratamiento con cifras bajas de filtrado glomerular en el seguimiento así como con la progresión más rápida de la enfermedad renal, lo cuál debe interpretarse con cautela, puesto que lo publicado en la literatura defiende un papel neutro o protector de estos fármacos sobre la progresión de la función renal (Baigent 2011, Cardoso 2012). Probablemente más que un efecto directo de las estatinas sobre la progresión de la enfermedad renal, esta asociación puede indicar que se trata de un grupo de pacientes que podría presentar más patología cardiovascular, puesto que, como se ha visto este tratamiento es más frecuente en pacientes con más edad y más comorbilidad.

Mediante mapas SOM tampoco se ha encontrado una asociación entre la obesidad y la progresión de la enfermedad renal en nuestra muestra de pacientes.

En el análisis de la anemia también los mapas autoorganizados nos aportan información añadida. Así, mediante la estadística clásica no se ha podido demostrar

una asociación entre las cifras de Hb y la progresión de la enfermedad renal, ni tampoco con los depósitos de hierro ni con el tratamiento de la anemia. Sin embargo, mediante el SOM encontramos una asociación entre las cifras bajas de Hb y FG bajos durante ese seguimiento, así como con una mayor pendiente de descenso del FG, tal como confirman algunos trabajos publicados en la literatura (Johson 2008, Hsu 2009, Chase 2014).

Algo similar ocurre cuando se analiza la relación entre la hipertensión arterial y su tratamiento, con la progresión de la enfermedad renal. Mediante la estadística clásica no hemos podido demostrar una relación de dependencia, sin embargo, los mapas SOM nos permiten analizar la relación entre el FG y su pendiente de descenso, el tratamiento en global y con IECAs- ARA II y diuréticos, la enfermedad de base, etc. Además, nos permiten caracterizar a algún grupo concreto, como los pacientes con HTA refractaria, en los que podemos ver que no son los que tienen una mayor progresión de la enfermedad renal, pero sí que mantienen FG bajos, y son los que precisan tratamiento combinado con IECAs-ARA II, además de diuréticos y otros antihipertensivos, y además tienen un IMC mayor de 30.

Los mapas autoorganizados nos permiten también extraer nuevas conclusiones al analizar la acidosis metabólica. En este caso, la estadística clásica nos ha permitido establecer una relación de dependencia entre la progresión de la enfermedad renal y los niveles séricos de bicarbonato, aunque siendo estrictos, sin alcanzar la significación estadística ( $p=0,06$ ). Estos datos coinciden con algunos trabajos que defienden el papel de los niveles bajos de bicarbonato con el descenso más rápido del FG (Drawz 2013, Chase 2014). En cambio, los mapas SOM nos permiten corroborar de manera simultánea la relación existente entre la pendiente de descenso del FG con valores más bajos de FG, los niveles bajos de bicarbonato y la elevación del anión GAP.

También en el caso del metabolismo óseo-mineral, los mapas SOM nos han permitido ampliar la información obtenida del análisis de los datos mediante técnicas clásicas. Además de confirmar la relación existente entre los niveles bajos de calcio en sangre y elevados de fósforo con la progresión de la enfermedad renal, podemos además ver que esta situación se asocia a la presencia de niveles elevados de PTH y

### 3. RESULTADOS Y DESARROLLO ARGUMENTAL

---

bajos de vitamina D nativa, y que además, es este grupo de pacientes el que más tratamiento recibe con derivados activos de la vitamina D.

Por todo ello resulta evidente que el uso de los mapas SOM puede ser una herramienta muy útil en el campo de la progresión de la enfermedad renal, donde son tantos los factores que influyen de manera simultánea y muchas veces con una relación de dependencia entre ellos, que hace muy difícil extraer conclusiones claras. Este hecho justifica que el papel de los factores de progresión en esta enfermedad no se acaben de definir, pues los resultados publicados en los estudios no son concluyentes.





## 4. CONCLUSIONES

---



## 4. CONCLUSIONES

El estudio realizado “Análisis de la enfermedad renal crónica avanzada mediante técnicas de aprendizaje máquina” ha permitido extraer las siguientes conclusiones:

- Los pacientes analizados han presentado una pendiente media de descenso del filtrado glomerular de  $-1.4 \text{ ml/min/1,73 m}^2/\text{año}$ .

- Un 15% del total de pacientes ha presentado una progresión rápida de la enfermedad renal, de lo que deducimos que más de un 80 % de pacientes con enfermedad renal crónica en categorías 3b, 4 y 5 van a evolucionar de forma lenta, sin necesidad de terapia sustitutiva a corto plazo. Por este motivo es de vital importancia identificar a los pacientes que potencialmente van a presentar un curso más rápido y sus factores de progresión para frenar en la medida de lo posible esta tendencia y programar de forma adecuada el inicio de terapia sustitutiva de la función renal.

- En el estudio comparativo entre los pacientes progresores rápidos y lentos realizado mediante contraste de hipótesis se han identificado como factores de progresión: proteinuria durante el seguimiento, proteinuria basal, fosforemia, calcemia, creatinina sérica y pendiente de descenso del FG entre los dos primeros controles.

- El mejor predictor de progresión renal encontrado en el análisis de regresión logística univariante ha sido la proteinuria, tanto su nivel basal como la obtenida en los sucesivos controles de seguimiento. Este hallazgo está acorde con el cuerpo de conocimiento (o conocimientos actuales) acerca de la enfermedad renal crónica y apoya la validez del trabajo.

- De los modelos multivariantes aplicados, tanto clásicos (regresión logística multivariante), como de aprendizaje máquina (Naïve-Bayes y TAN), el que ha mostrado mayor capacidad predictiva en la progresión de enfermedad renal ha sido el

modelo TAN tras la selección de variables automática (Boruta), con un AUC de 0.93 y una probabilidad de acierto en los progresores rápidos de un 100 % en la matriz de confusión. A partir de este modelo es factible la creación de una aplicación informática en la que, a partir de los valores de proteinuria, filtrado glomerular estimado y niveles de PTH en un paciente concreto, nos permita conocer si va a presentar una progresión rápida o lenta.

- Se ha generado también un árbol de decisión aplicable a la clínica diaria a partir de las variables seleccionadas, con un AUC de 0.79. Aunque es algo inferior al modelo anterior, la ventaja de éste es que no precisa de la creación de una aplicación informática, pudiendo ser utilizado directamente.

- De la aplicación de mapas autoorganizados a la población analizada, concluimos que los paciente progresores rápidos tienen una edad de 50 a 60 años, un IMC de 30 kg/m<sup>2</sup>, niveles de hemoglobina más bajos, cifras de creatinina plasmática, PTH y fósforo en sangre más elevados, proteinuria más severa, mayor representación de nefropatías glomerulares, niveles más elevados de colesterol total y LDL y más bajos de HDL. En cambio, los progresores lentos son de mayor edad, con un índice de comorbilidad de Charlson más elevado, tienen niveles séricos más bajos de colesterol, creatinina y fósforo, niveles normales de bicarbonato en sangre, y niveles de hemoglobina más elevados.

La caracterización minuciosa de los pacientes con enfermedad renal crónica en categorías 3b, 4 y 5 mediante mapas autoorganizados, y la aplicación de modelos predictivos probabilísticos y árboles de decisión nos va a permitir ralentizar la evolución de la enfermedad renal incidiendo en los factores que han mostrado mayor relación con su progresión (proteinuria, metabolismo óseo-mineral, colesterol, y acidosis). Asimismo nos ayudará a optimizar el inicio del tratamiento sustitutivo de la función renal al poder predecir cuáles son los pacientes que van a tener una progresión más rápida de la enfermedad.

## 5. BIBLIOGRAFÍA

---



## 5. BIBLIOGRAFÍA

Abramowitz MK, Hostetter TH and Melamed ML. The serum anion gap is altered in early kidney disease and associates with mortality. *Kidney Int.* 2012; 82(6): 701–709.

Adler AI, Stevens RJ, Manley SE, Bilous RW, Cull CA and Holman RR, on behalf of the UKPDS Group. Development and progression of nephropathy in type 2 diabetes in the United Kingdom Prospective Diabetes Study (UKPDS 64). *Kidney Int.* 2003; 63: 225 - 232.

Ahmadi SF, Zahmatkesh G, Ahmadi E, Strega E, Rhee CM, Gillen DL, et al. Association of Body Mass Index with Clinical Outcomes in Non-Dialysis-Dependent Chronic Kidney Disease: A Systematic Review and Meta-Analysis. *Cardiorenal Med.* 2016; 6: 37–49.

Alpaydin E. Design and Analysis of Machine Learning Experiments. In: Dietterch T, ed. *Introduction to Machine Learning*. 2 ed. London: The MIT Press; 2010. p 475 - 516.

Astor BC, Matsushita K, Gansevoort RT, Van der Velde M, Woodward M, Levey AS, et al. Lower estimated glomerular filtration rate and higher albuminuria associated with mortality and end-stage renal disease. A collaborative meta-analysis of kidney disease population cohorts. *Kidney Int.* 2011;79(12):1331-40.

Baigent C, Landray MJ, Reith C, Emberson J, Wheeler DC, Tomson C, et al. The effects of lowering LDL cholesterol with simvastatin plus ezetimibe in patients with chronic kidney disease. *Study of Heart and Renal Protection: a randomised placebo-controlled trial.* *Lancet.* 2011; 377 (9784): 2181–2192

Bonnet F, Deprele C, Sassolas A, Moulin P, Alamartine E, Berthezène F, et al. Excessive body weight as a new independent risk factor for clinical and pathological progression in IgA nephritis. *Am J Kidney Dis.* 2001; 37(4) : 720–727.

Brancati F, Whelton P, Randall B, Neaton JD, Stamler J and Klag MJ. Risk of end-stage renal disease in diabetes mellitus: A prospective cohort study of men screened for MRFIT. Multiple Risk Factor Intervention Trial. *JAMA* 278 (23): 2069–2074.

Cerqueira DC, Soares CM, Silva VR, Magalhães JO, Barcelos IP, Duarte MG, et al. A Predictive Model of Progression of CKD to ESRD in a Predialysis Pediatric Interdisciplinary Program. *Clin J Am Soc Nephrol.* 2014; 9: 728–735.

Chase HS, Hirsch JS, Mohan S, Rao MK, Radhakrishnan J. Presence of early CKD-related metabolic complications predict progression of stage 3 CKD: a case-controlled study. *BMC Nephrology* [Internet]. 2014 [Citado 31 de Marzo de 2017]; 15:187. Disponible en: [\\_http://www.biomedcentral.com/1471-2369/15/187](http://www.biomedcentral.com/1471-2369/15/187)

Charlson ME, Charlson RE, Paterson JC, Marinopoulos SS, Briggs WM and Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol.* 2008; 61(12): 1234 -1240.

Collister D, Ferguson T, Komenda P and Tangri N. The patterns, risk factors, and prediction of progression in chronic kidney disease: A narrative review. *Seminars in Nephrology.* 2016; 36(4): 273 - 282.

Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, et al. Decline in estimated glomerular filtration rate and subsequent risk of end-stage renal disease and mortality. *JAMA.* 2014; 311 (24): 2518 - 31.

DCCT/EDIC Research Group, de Boer IH, Sun W, Cleary PA, Lachin JM, Molitch ME, et al. Intensive Diabetes Therapy and Glomerular Filtration Rate in Type 1 Diabetes. *N Engl J Med.* 2011; 365(25): 2366–2376.



De Brito-Ashurst I, Varaganam M, Raftery MJ and Yaqoob MM. Bicarbonate supplementation slows progression of CKD and improves nutritional status. *J Am Soc Nephrol.* 2009; 20: 2075-2084.

Dobre M, Yang W, Pan Q, Appel L, Bellovich K, Chen J et al. Persistent high serum bicarbonate and the risk of heart failure in patients with chronic kidney disease (CKD): a report from the Chronic Renal Insufficiency Cohort (CRIC) Study. *J Am Heart Assoc.* 2015; 4(4): 1-10.

Documento de consenso sobre enfermedad renal crónica. Sociedad Española de Nefrología. Noviembre 2012.

Drawz P and Rosenberg M. Slowing progression of chronic kidney disease. *Kidney International Supplements.* 2013; 3, 372-376.

Echouffo-Tcheugui JB and Kengne AP. Risk models to predict chronic kidney disease and its progression: a systematic review. *PLoS Med [Internet].* 2012; 9(11). Disponible en: <https://doi.org/10.1371/journal.pmed.1001344>

Eknoyan G. Obesity and chronic kidney disease. *Nefrología.* 2011; 31(4): 397-403

Elihimas UF, Elihimas HC, Lemos VM, Leão Mde A, França EE, Lemos A, et al. Smoking as risk factor for chronic kidney disease: systematic review. *J Bras Nefrol.* 2014; 36(4): 519-528.

Ferrario M, Moissl U, Garzotto F, Signorini MG, Cruz D, Tetta C, et al. Study of the autonomic response in hemodialysis patients with different fluid overload levels. 32nd Annual International IEEE Engineering in Medicine and Biology Society (EMBS) Conference; 2010 Aug 31- Sept 4. Buenos Aires, Argentina.

Flach P. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data.* Cambridge. University Press; 2012.

Goicoechea M, de Vinuesa SG, Verdalles U, Ruiz-Caro C, Ampuero J, Rincón A et al. Effect of allopurinol in chronic kidney disease progression and cardiovascular risk. *Clin J Am Soc Nephrol*. 2010; 5(8):1388-1393.

Goicoechea M, Garcia de Vinuesa S, Verdalles U, Verde E, Macias N, Santos A, et al. Allopurinol and progression of CKD and cardiovascular events: long-term follow-up of a randomized clinical trial. *Am J Kidney Dis*. 2015; 65(4): 543 - 9.

Gorostidi M, Santamaría R, Alcázar R, Fernández-Fresnedo G, Galcerán JM, Goicoechea M, et al. Documento de la Sociedad Española de Nefrología sobre las guías KDIGO para la evaluación y el tratamiento de la enfermedad renal crónica. *Nefrología* 2014; 34(3):302-16.

Gouva C, Nikolopoulos P, Ioannidis JP and Siamopoulos KC. Treating anemia early in renal failure patients shows the decline of renal function: a randomized controlled trial. *Kidney Int*. 2004; 66(2):753-760.

Grams ME, Li L, Greene TH, Tin A, Sang Y, Kao WH, Lipkowitz MS, et al. Estimating time to ESRD using kidney failure risk equations: Results from the African American Study of Kidney Disease and Hypertension (AASK). *Am J Kidney Dis*. 2015; 63(3): 394-402.

Halimi JM, Giraudeau B, Vol S, Caces E, Nivet H, Lebranchu Y, et al. Effects of current smoking and smoking discontinuation on renal function and proteinuria in the general population. *Kidney international*. 2000; 58: 1285– 92.

Hallan SI and Orth SR. Smoking is a risk factor in the progression to kidney failure. *Kidney Int*. 2011; 80: 516–523.

Hastie T, Tibshirani R and Friedman J. *The Elements of Statistical learning: Data Mining, Inference and Prediction*. Second edition. Springer; 2001.

Haynes R., Stalin N., Emerson J., Herrington W., Tomson C., Agodoa L., et al. Evaluating the Contribution of the Cause of Kidney Disease to Prognosis in CKD: Results From the Study of Heart and Renal Protection (SHARP). *Am J Kidney Dis.* 2014; 64(1): 40 - 48.

Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al. Global Prevalence of Chronic Kidney Disease – A Systematic Review and Meta-Analysis. *PLoS One* [Internet]. 2016;11(7):e0158765. Disponible en: <https://doi.org/10.1371/journal.pone.0158765>

Hunsicker LG, Adler S, Caggiula A, England BK, Greene T, Kusek JW, et al. Predictors of the progression of renal disease in the Modification on Diet in Renal Disease Study. *Kidney Int.* 1997 Jun; 51(6): 1908-19.

Hsu CY, McCulloch CE, Iribarren C, Darbinian J, Go AS. Body mass index and risk for end-stage renal disease. *Ann Intern Med.* 2006; 144(1): 21 - 8.

Hsu CY, Iribarren C, McCulloch CE, Darbinian J, Go AS. Risk factors for end-stage renal disease: 25-year follow-up. *Arch Intern Med.* 2009;169 (4): 342 - 350.

Inker LA, Levey AS, Pandya K, Woodward M, Stoycheff N, Okparavero A et al. Early Change in Proteinuria as a Surrogate Endpoint for Kidney Disease Progression: An Individual Patient Meta-analysis. *Am J Kidney Dis.* 2014; 64(1): 74–85.

Inrig JK, Barnhart HX, Reddan D, Patel UD, Sapp S, Califf RM, et al. Effect of hemoglobin target on progression of kidney disease: a secondary analysis of the CHOIR (Correction of Hemoglobin and Outcomes in Renal Insufficiency) trial. *Am J Kidney Dis.* 2012; 60(3): 390 - 401.

Iseki K, Ikemiya Y, Iseki C, and Takishita S. Proteinuria and the risk of developing end-stage renal disease. *Kidney Int.* 2003; 63:1468–1474.

Jalal DI, Chonchol M, Chen W and Targher G. Uric acid as a Target of Therapy in CKD. *Am J Kid Dis.* 2013; 61 (1): 134 - 46.

Jafar TH, Schmid CH, Stark PC, Toto R, Remuzzi G, Ruggenenti P et al. The rate of progression of renal disease may not be slower in women compared with men: a patient-level meta-analysis. *Nephrol Dial Transplant.* 2003; 18 (10): 2047-2053.

Jain G and Jaimes E. Nicotine signaling and progression of chronic kidney disease in smokers. *Biochem Pharmacol.* 2013; 86 (8): 1215 - 23.

Jeong J, Kwon SK and Kim HY. Effect of bicarbonate supplementation on renal function and nutritional indices in predialysis advanced chronic kidney disease. *Electrolyte Blood Press.* 2014; 12: 80-87.

Johnson ES, Thorp ML, Platt RW and Smith DH. Predicting de risk of dialysis and transplant among patients with CKD: a retrospective cohort study. *Am J Kidney Dis.* 2008; 52 (4): 653-60.

Kalaitzidis R and Elisaf M. The role of statins in Chronic kidney disease. *Am J Nephrol.* 2011; 34: 195-202.

Kanji T, Gandhi M, Clase CM and Yang R. Urate lowering therapy to improve renal outcomes in patients with chronic kidney disease: systematic review and meta-analysis. *BMC Nephrology* [Internet] 2015. [Citado 29 de Enero de 2017]. DOI 10.1186/s12882-015-0047-z

KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney Inter, Suppl* 2013; 3: 1-150.

Keane WF, Zhang Z, Lyle PA et al. Risk scores for predicting outcomes in patients with type 2 diabetes and nephropathy: The RENAAL Study. *Clin J Am Soc Nephrol*. 2006; 1 (4): 761-7.

klahr S and Morrissey J. Progresión of chronic renal disease. *Am J kidney Dis*. 2003; 41(3) Supl: S3-S7.

Klag MJ, Whelton PK, Randall BL, Neaton JD, Brancati FL, Ford CE et al. Blood pressure and end stage renal in men. *N Engl J Med*. 1996; 334(1):13-8.

Kohonen T. *Self-Organizing Maps*. 3ª edición. Springer; 2001.

Koller D and Friedman N. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press; 2009.

Kraut JA and Madias NE. Metabolic Acidosis of CKD: An Update. *Am J Kidney Dis*. 2016; 67(2): 307-317.

Krzanowski WJ and Hand D. *ROC Curves for Continuous Data*. CRC Press; 2009

Kursa MB and Rudnicki WR. Feature Selection with the Boruta Package. *Journal of Statistical Software*. 2010; (11): 1-13.

Lee J and Verleysen M. *Nonlinear Dimensionality Reduction*. Springer; 2007.

Levey A, Eckardt KU, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and clasification of crhonic kidney disease: a position statement for Kidney Disease: Improving Global Outcomes (KDIGO). *Kidney Int*. 2005; 67:2089-2100.

Levey A, Stevens L and Coresh J. Conceptual Model of CKD: Applications and Implications. *Am J Kidney Diseases*. 2009; 53 (Suppl 3): S4-S16

Levey A, Stevens L, Schmid C, Zhang Y, Castro A, Feldman H et al. A New Equation to Estimate Glomerular Filtration Rate. *Ann Intern Med.* 2009; 150(9): 604–612.

Li L, Chang A, Rostand SG, Hebert L, Appel LJ, Astor BC, et al. A Within-patient analysis for time-varying risk factors of CKD progression. *J Am Soc Nephrol.* 2014; 25: 606-613.

Li L, Astor BC, Lewis J, et al. Longitudinal progression trajectory of GFR among patients with CKD. *Am J Kidney Dis.* 2012; 59: 504-12.

Lorenzo V, Saracho R, Zamora J, Rufino M, Torres A. Similar renal decline in diabetic and non-diabetic patients with comparable levels of albuminuria. *Nephrol Dial Transplant.* 2010; 22: 2009-16

Lorenzo V. Enfermedad renal crónica. En: Lorenzo V, López J.M, editores. *Nefrología al día.* 1 ed. Barcelona: Edición del Grupo Editorial de la Sociedad Española de Nefrología; 2010.p.335-52.

Madero M, Sarnak MJ, Wang X, Greene T, Beck GJ, Kusek JW, et al. Uric acid and long term outcomes in CKD. *Am J Kidney Diseases.* 2009; 53(5): 796-83.

Mallamaci F and Tripepi G. Obesity and CKD progression: hard facts on fat CKD patients. *Nephrol Dial Transplant.* 2013; 28 (Suppl. 4):105–108.

Mancia G, Fagard R, Narkiewicz K, Redon J, Zanchetti A, Böhm M, et al. 2013 ESH/ESC Guidelines for the management of arterial hypertension. *Journal of hypertension.* 2013; 31(7): 1281-1357.

Martín J.D, Gómez F, Soria E, Schmidhuber J, Climente M, Jiménez V. A reinforcement learning approach for individualizing erythropoietin dosages in hemodialysis patients. *Expert Systems with Applications.* 2009; 36 (6): 9737- 42.

Martín JD, Marcelli D, Soria E, Mari F, Martínez JM, Soley I, et al. Self-Organizing Maps: A new way to screen the level of satisfaction of dialysis patients. *Expert Systems with Applications*. 2012 ;39(10): 8793–98.

Mcclellan W and Flanders D. Risk factors for progressive chronic kidney disease. *J Am Soc Nephrol*. 2003;14: S65-S70.

Menon V, Tighiouart H, Vaughn NS, Beck GJ, Kusek JW, Collins AJ, et al. Serum bicarbonate and long-term outcomes in CKD. *Am J Kidney Dis*. 2010; 56(5): 907-914.

Muntner P, Coresh J, Smith JC, Eckfeldt J and Klag MJ. Plasma lipids and risk of developing renal dysfunction: The Atherosclerosis Risk in Communities study. *Kidney Int* 2000; 58: 293–301.

Nacak H, van Diepen M, Qureschi A, Carrero JJ, Stijnen T, Dekker FW et al. Uric acid is not associated with decline in renal function or time to renal replacement therapy initiation in a referred cohort of patients with Stage III, IV and V chronic kidney disease. *Nephrol Dial Transplant*. 2015; 30: 2039–2045.

Nath KA, Hostetter MK and Hostetter TH. Pathophysiology of chronic tubulointerstitial disease in rats: Interactions of dietary acid load, ammonia, and complement component-C3. *J Clin Invest*. 1985; 76: 667-675.

National Kidney Foundation. KDOQI Clinical Practice Guidelines for chronic kidney disease: evaluation, classification and Stratification. *Am J Kidney Dis* 2002; 39: s1

Neugarten J, Acharya A and Silbiger SR. Effect of gender on the progression of non diabetic renal disease: a meta-analysis. *J Am Soc Nephrol*. 2000; 11 (2):319-329.

NICE Clinical Guidelines nº 73. National Collaborating Centre for Chronic Conditions. Chronic kidney disease: national clinical guideline for early identification and management in adults in primary and secondary care. London: Royal College of Physicians;2008.

Obermayr RP, Temml C, Gutjarh G, Knechtelsdorfer M, Oberbauer R and Klausner-Braun R. Elevated uric acid increases the risk for kidney disease. *J Am Soc Nephrol.* 2008; 19(12): 2407-2413.

O'Hare AM, Batten A, Burrows NR, Pavkov ME, Taylor L, Gupta I, et al. Trajectories of kidney function decline in the 2 years before initiation of long-term dialysis. *Am J Kidney Dis.* 2012; 59 (4): 513-22.

Orth SR, Stöckmann A, Conrads C, Ritz E, Ferro M, Kreuzer W, et al. Smoking as a risk factor for end-stage renal failure in men with primary renal disease. *Kidney Int.*1998; 54: 926- 31.

Otero A, De Francisco A, Gayoso P, García F, on behalf of the EPIRCE Study Group. Prevalence of chronic renal disease in Spain: Results of the EPIRCE study. *Nefrología* 2010; 30 (1): 78-86.

Peña D. *Fundamentos de Estadística.* Madrid: Alianza Editorial; 2008.

Pinto-Sietsma SJ, Mulder J, Janssen WM, Hillege HL, de Zeeuw D, and de Jong PE. Smoking is related to albuminuria and abnormal renal function in nondiabetic persons. *Ann Intern Med.* 2000; 133 (8): 585–91.

Portolés J, Gorriz JL, Rubio E, de Alvaro F, García F, Alvarez-Chivas V and NADIR-3 Study Group. The development of anemia is associated to poor prognosis in NKF/KDOQI stage 3 chronic kidney disease. *BMC Nephrol* [Internet] 2013. [Citado el 26 sept de 2016] 7;14:2. Disponible en <http://www.biomedcentral.com/1471-2369/14/2>.



Rahman M, Yang W, Akkina S, Alper A, Anderson A, Appel L, et al. Relation of Serum Lipids and Lipoproteins with Progression of CKD: The CRIC Study. *Clin J Am Soc Nephrol*. 2014; 9: 1190–1198.

Raphael KL, Zhang Y, Wei G, Greene T, Cheung AK and Beddhu S. Serum bicarbonate and mortality in adults in NHANES III. *Nephrol Dial Transplant*. 2013; 28(5): 1207-1213.

Remuzzi G, Chiurciu C and Ruggenenti P. Proteinuria predicting outcome in renal disease: Nondiabetic nephropaties (REIN). *Kidney int*. 2004; 66 (Suppl 92): S90-S96.

Remuzzi G, Benigni A and Remuzzi A: Mechanisms of progression and regression of renal lesions of chronic nephropathies and diabetes. *J Clin Invest*. 2006; 116: 288–296.

Remuzzi A, Sangalli F, Macconi D, Tomasoni S, Cattaneo, Rizzo P et al. Regression of Renal Disease by Angiotensin II Antagonism Is Caused by Regeneration of Kidney Vasculature. *J Am Soc Nephrol*. 2016; 27(3): 699-705..

Riffenburgh RH. *Statistics in Medicine* (2<sup>a</sup> ed.). London. Academic Press; 2006.

Rosert J, Levin A, Roger SD, Hörl WH, Fouqueray B, Gassmann-Mayer C, et al. Effect of early correction of anemia in the progression of CKD. *Am J Kidney Dis*. 2006; 47(5): 738 - 50.

Ruiz-Ortega M, Rupérez M, Esteban V, Rodríguez-Vita J, Sánchez-López E, Carvajal G, et al. Angiotensin II: a key factor in the inflammatory and fibrotic response in kidney diseases. *Nephrol Dial Transplant*. 2006; 21(1): 16 - 20.

Rucci P, Mandreoli M, Gibertoni D, Zuccala A, Fantini MP, Lenzi J, et al. A clinical stratification tool for chronic kidney disease progression rate based on classification tree analysis. *Nephrol Dial Transplant*. 2014; 29: 603 - 610.

---

Samuelsson O, Mulec H, Knight-Gibson C, Attman PO, Kron B, Larsson B, et al. Lipoprotein abnormalities are associated with increased rate of progression of human chronic renal insufficiency. *Nephrol Dial Transplant*. 1997; 12 (9): 1908 - 15.

Schaeffner ES, Kurth T., Curhan GC, Glynn RJ, Rexrode KM, Beignet C, et al. Cholesterol and the risk of renal dysfunction in apparently healthy men. *J Am Soc Nephrol*. 2003; 14(8): 2084 - 91.

Schwarz S, Trivedi BK, Kalantar-Zadeh K and Kovesdy CP. Association of disorders in mineral metabolism with progression of chronic kidney disease. *Clin J Am Soc Nephrol*. 2006; 1(4): 825-31

Shah SN, Abramowitz M, Hostetter TH and Melamed ML. Serum bicarbonate levels and the progression of kidney disease: a cohort study. *Am J Kidney Dis*. 2009; 54(2):270-277.

Stevens LA, Coresh J, Greene T and Levey AS. Assessing kidney function measured and estimated glomerular filtration rate. *N Eng L Med*. 2006 ; 354(23): 2473 - 83.

Siu YP, Leung KT, Tong MK and Kwan TH. Use of allopurinol in slowing the progression of renal disease through its ability to lower serum uric acid level. *Am J Kidney Dis*. 2006; 47 (1): 51-59.

Stanton RC. Clinical Challenges in Diagnosis and Management of Diabetic Kidney Disease. *Am J Kidney Dis*. 2014; 63(2)(suppl 2): S3 - S2.

Taal MW and Brenner BM. Predicting initiation and progression of chronic kidney disease: Developing renal risk scores. *Kidney Int*. 2006; 70(10):1694–1705.

Tangri N, Stevens LA, Griffith J, Tighiouart H, Djurdjev O, Naimark D, et al: A predictive model for progression of chronic kidney disease to kidney failure. *JAMA* 2011 Apr 20; 305(15): 1553-9.

Tangri N, Grams M, Levey A, Coresh J, Appel L, Astor B, et al. Multinational assessment of accuracy of equations for predicting risk of kidney failure: a meta-analysis. *JAMA* 2016; 315(2): 164–174.

Tangri N, Inker LA, Hiebert B, Wong J, Naimark D, Kent D et al. A Dynamic Predictive Model for Progression of CKD. *Am J Kidney Dis.* 2017; 69(4): 514 - 20.

Tarver-Karr ME, Brancati FL, Eberhart MS, et al. Proteinuria and the risk of chronic kidney disease (CKD) in the United States. *J Am Soc Nephrol.* 2000;11:168.

The Criteria Committee of the New York Heart Association. Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels, 9th ed, Little, Brown and Co, Boston, 1994, p.253.

Titapiccolo JI, Ferrario M, Cerutti S, Signorini M.G, Barbieri C, Mari F, et al. Mining Medical Data to Develop Clinical Decision Making Tools in Hemodialysis. *IEEE 12th International conference on Data Mining Workshops*; 2012 Dec 10-10. Brussels, Belgium.

Togawa A, Uyama S, Takanohashi S, Shimasaki M, Miyaji T, Endo H, et al. Adjusted Anion Gap is associated with glomerular filtration rate decline in chronic kidney disease. *Nephron Extra.* 2013;3 (1):113–117.

Tonelli M, Moyé L, Sacks FM, Cole T and Durham GC. Effect of pravastatin on loss of renal function in people with moderate chronic renal insufficiency and cardiovascular disease. *J Am Soc Nephrol.* 2003; 14 (6): 1605-1613.

Tsai WC, Wu HY, Peng YS, Ko MJ, Wu MS, Hung KY, et al. Risk factors for development and progression of chronic kidney disease. A systematic review and exploratory meta-analysis. *Medicine.* 2016; 95 (11): 1-9.

---

Upadhyay A, Earley A, Haynes SM and Uhlig K. Systematic review: blood pressure target in chronic kidney disease and proteinuria as an effect modifier. *Ann Intern Med.* 2011;154(8):541-8.

Van den Beukel O, Goeji M, Dekker F, Siegert C, Halbesma N, for de PREPARE Study Group. Differences in Progression to ESRD between Black and White Patients Receiving Predialysis Care in a Universal Health Care System. *Clin J Am Soc Nephrol.* 2013; 8 (9): 1540–1547.

Vargas F. Documento Marco sobre Enfermedad Renal Crónica (ERC) dentro de la Estrategia de Abordaje a la cronicidad en el SNS. 2015.

Vega A, Abad S, Verdalles U, Aragoncillo I, Velazquez K, Quiroga B, et al. Dose equivalence between continuous erythropoietin receptor activator (CERA), Darbepoetin and Epoetin in patients with advanced chronic kidney disease. *Hippokratia.* 2014; 18 (4): 315-318.

Wang Y, Chen X, Song Y, Caballero B and Checkin LJ. Association between obesity and kidney disease: a systematic review and meta-analysis. *Kidney Int.* 2008; 73(1):19-33.

Weiner DE, Tighiouart H, Elsayed EF, Griffith JL, Salem DN and Levey AS. Uric acid and incident kidney disease in the community. *J Am Soc Nephrol.* 2008; 19(6):1204-1211.

Xie X, Liu Y, Perkovic V, Li X, Ninomiva T, Hou W, et al. Renin-Angiotensin System Inhibitors and Kidney and Cardiovascular Outcomes in Patients With CKD: A Bayesian Network Meta-analysis of Randomized Clinical Trials. *Am J Kidney Dis.* 2016; 67(5): 728-741.

Yu HT. Progression of chronic renal failure. *Arch Intern Med.* 2003; 163 (12): 1417-29.

Yuste C, Barraca D, Aragoncillo-Sauco I, Vega-Martinez A, Abad S, Versalles-Guzman U, et al. Factores asociados con la progresión de la enfermedad renal crónica. *Nefrología*. 2013; 33(5): 685 - 91.

Zeeuw D, Remuzzi G, Parking H-H, Keane WF, Zhang Z, Shahinfar S, et al. Proteinuria, a target for renoprotection in patients with type 2 diabetic nephropathy: Lessons from RENAAL. *Kidney int* 2004; 65 (6): 2309 - 20.

Zheng H, Ji J, Zhao L, Chen M, Shi A, Pan L, et al. Prediction and diagnosis of renal cell carcinoma using nuclear magnetic resonance-based serum metabolomics and self-organizing maps. *Oncotarget* 2016; 7(37): 59189 - 59198.

Zoccali C, Ruggenenti P, Perna A, Leonardis D, Tripepi R, Tripepi G, et al. Phosphate may promote CKD progression and attenuate renoprotective effect of ACE inhibition. *J Am Soc Nephrol* 2011; 22 (10):1923-30.



## **6. ANEXOS**

---





**ANEXO I. COMPARACIÓN DE LAS DIFERENTES VARIABLES CONTINUAS ANALIZADAS EN AMBOS GRUPOS DE PACIENTES (TEST DE STUDENT PARA VARIABLES CON DISTRIBUCIÓN NORMAL Y TEST DE WILCOXON PARA DISTRIBUCIÓN NO NORMAL).**

	<b>NO PROGRESORES</b> mean+/- std, median +/- iqr *	<b>PROGRESORES</b> mean+/-std median +/-iqr *	p.value
EDAD	70.39 +/- 17.69 *	60.04 +/- 17.2	0.0854
DosisEPOmesKg	205.31 +/- 134.59 *	258.32 +/- 139.3	0.242
COL.TOT	179.16 +/- 32.88	196.66 +/- 42.23	0.11
COL.HDL	50.71 +/- 12.31	52.46 +/- 15.03	0.646
COL.LDL	100.57 +/- 26.17	117.38 +/- 35.11	0.067
TG	128.71 +/- 67.17 *	138.92 +/- 43.14	0.776
IMC	28.32 +/- 4.82	27.66 +/- 3.73	0.522
HB	12.3 +/- 1.01 *	12.06 +/- 1.07	0.324
IST	25.75 +/- 7.92 *	28.22 +/- 10.82	0.53
FERRITINA	163.56 +/- 141.32 *	135.75 +/- 222.54 *	0.655
UREA	117.69 +/- 26.43	124.18 +/- 25.48	0.334
<b>CREAT</b>	<b>3.27 +/- 0.82</b>	<b>3.97 +/- 0.87</b>	<b>0.00437</b>
<b>FOSFORO</b>	<b>4.01 +/- 0.53</b>	<b>4.44 +/- 0.53</b>	<b>0.00378</b>
<b>CALCIO</b>	<b>9.59 +/- 0.4 *</b>	<b>9.3 +/- 0.5 *</b>	<b>0.0437</b>
D3	22.66 +/- 10.09	23.98 +/- 10.24	0.661
PTH	181.06 +/- 88.85	240.96 +/- 127.34	0.0703
BICARBO	23.83 +/- 2.33	22.53 +/- 2.62	0.0616
ACIDO URICO	7.21 +/- 1.15	7.13 +/- 0.8	0.733
PROTEINURIA	0.73 +/- 1.16 *	3.24 +/- 1.92	1.21e-05
ALBUMINA	4.1 +/- 0.28	4.06 +/- 0.39 *	0.0626
HB.GLIC	6.45 +/- 1.39 *	6.54 +/- 0.77	0.982
GLUCOSA	100.6 +/- 21.8 *	103.8 +/- 23.71	0.479
PCR	6.1 +/- 9.86 *	9.8 +/- 32.97 *	0.554
CHARLSON	5.5 +/- 3 *	4.72 +/- 2.19	0.0989

	<b>NO PROGRESORES</b> mean+/- std, median +/- iqr *	<b>PROGRESORES</b> mean+/-std median +/-iqr *	p.value
CHARLSON	5.5 +/- 3 *	4.72 +/- 2.19	0.0989
FGeCKD	16.97 +/- 8.11 *	15.42 +/- 4.93	0.0926
<b>FGeMDRD_IDMS</b>	<b>18.25 +/- 5.38</b>	<b>15.43 +/- 4.71</b>	<b>0.051</b>
<b>FGeMDRD_4</b>	<b>19.39 +/- 5.72</b>	<b>16.4 +/- 5.01</b>	<b>0.051</b>
AGAPTrad	11.31 +/- 1.59	11.87 +/- 2.15	0.346
AGAP_AAlbumina	1.05 +/- 1.6	1.89 +/- 1.92	0.126
AGAP_ACalcio	8.37 +/- 1.49	8.92 +/- 1.78	0.272
PENDIENTE FG	-0.69 +/- 2.52 *	-7.04 +/- 1.66	1.87e-11
TIEMPO SEGUIM	2.98 +/- 3.78 *	1.33 +/- 0.49	1.15e-06
FG INICIAL	18.12 +/- 7.59 *	20.18 +/- 6.04	0.288
<b>PROTEINURIA IN</b>	<b>0.6 +/- 1.14 *</b>	<b>3.15 +/- 2.45</b>	<b>0.000157</b>
BICARBON INIC	24.45 +/- 2.95	23.04 +/- 3.69	0.165
HB INICIO	12.62 +/- 1.64	12.83 +/- 1.46	0.59
PTH INICIO	173 +/- 186.7 *	274.35 +/- 207.95	0.111
FOSFORO INI	3.93 +/- 0.62	4.04 +/- 0.63	0.524
URICO INI	7 +/- 2.31 *	7.44 +/- 1.14	0.563
<b>mFG INICIO</b>	<b>0.17 +/- 7.32 *</b>	<b>-6.59 +/- 12.12</b>	<b>0.002</b>
mPROTEIN INIC	0 +/- 0.81 *	0 +/- 2.5 *	0.345
mBICARBO INIC	0 +/- 5.42 *	-0.68 +/- 10.14 *	0.851
mHB INIC	-0.14 +/- 3.55 *	-0.21 +/- 4.74	0.426
mPTHIni	0 +/- 143.77 *	-193.06 +/- 641.14	0.178
mFOSFORO INIC	0 +/- 1.9 *	0.1 +/- 1.96 *	0.595
mURICO INIC	-0.13 +/- 2.8 *	-0.1 +/- 3.61	0.583

## **ANEXO II. LISTADO DE SÍMBOLOS, ABREVIATURAS Y SIGLAS**

AASK: *African American Study of Kidney Diseases and Hypertension.*

AGAP A albúmina: Anión GAP ajustado a albúmina

AGAP A calcio: Anión GAP ajustado a calcio

AGAP trad: Anión GAP tradicional

ALBUMI.0: Albúmina

ARA-II: Antagonistas del Receptor de Angiotensina II

ARIC: *Atherosclerosis Risk in Communities Study*

AUC: Area bajo la curva

Bicarbonato Ini: Niveles de bicarbonato en sangre en el primer control

CARE: *Cholesterol and Recurrent Events Trial*

CERA: Metoxi polietilen-glicol epoetina beta

CHOIR: *Correction of Hemoglobin and Outcomes in Renal Insufficiency*

CKD-EPI: *Chronic Kidney Diseases Epidemiology Collaboration*

ColCharlson: Índice de comorbilidad de Charlson

COL.HDL: Colesterol HDL

COL.LDL: Colesterol LDL

COL.TOT: Colesterol total

CREATE: *Cardiovascular Risk Reduction by Early Anemia Treatment with Epoetin Beta*

CRIC: *The Chronic Renal Insufficiency Cohort study*

D3: Niveles séricos de vitamina D3

DCCT/EDIC: *Diabetes Control and Complications Trial/ Epidemiology of Diabetes Interventions and Complications*

DM: Diabetes mellitus

DuraRegAnys: Tiempo de seguimiento

ECG: Electrocardiograma

EPIRCE: Epidemiología de la Insuficiencia Renal Crónica en España

ERC: Enfermedad renal crónica.

ERT: Enfermedad renal terminal

ET-1: Endotelina 1

FDA: *Food and Drug Administration*

FG: Filtrado glomerular

FGe: Filtrado glomerular estimado

FGeCKD: Filtrado glomerular estimado por la fórmula CKD-EPI

FGeIni: Filtrado glomerular estimado al inicio del seguimiento.

FGeMDRD4: Filtrado glomerular estimado por la fórmula MDRD-4

FGeMDRD-IDMS: Filtrado glomerular estimado por la fórmula MDRD con creatinina estandarizada para IDMS

Hb: Hemoglobina

HVI: Hipertrofia ventricular izquierda

IECAs: Inhibidores de la enzima convertora de la angiotensina

IncTiempo: Tiempo entre los dos primeros controles realizados

IMC: Índice de masa corporal

IST: Índice de saturación de transferrina

KDIGO: *Kidney Disease Improving Global Outcomes*

MDRD: *Modification of Diet in Renal Disease*

mFósforo Ini: Pendiente de modificación de los niveles de fósforo en sangre entre los dos controles iniciales realizados

mHb Ini: Pendiente de modificación de los niveles de hemoglobina en sangre entre los dos controles iniciales realizados

MRFIT: *Multiple Risk Factor Intervention Trial*

mFRPor Ini: Pendiente de modificación de los valores de filtrado glomerular estimados en los dos controles iniciales realizados

NADIR-3: *Study of non-anemic stage 3 CKD patients who develop renal anemia*

NFK: *National Kidney Foundation*

NFK-KDOQI: *National Kidney Foundation-Kidney Disease Outcomes Quality Initiative*

NYHA: *New York Heart Association*

ONTARGET: *Ongoing Telmisartan Alone and in Combination with Ramipril Global Endpoint Trial*

PCR: Proteína C Reactiva

PendienteAny: Pendiente anual de descenso del filtrado glomerular para cada paciente

PROTEINU.0: Proteinuria

PTH: Hormona paratiroidea

PTH Ini: Niveles de hormona paratiroidea al inicio del seguimiento.

REIN: *Ramipril Efficacy In Nephropathy*

RENAAL: *Reduction In End Points in Noninsulin-Dependent Diabetes Mellitus with the Angiotensin II Antagonist Losartan*

SHARP: *Study of Heart And Renal Protection*

SOM: *Self Organizing Maps*

SRAA: *Sistema Renina Angiotensina Aldosterona*

TAN: *Tree Augmented Network*

TGF- $\beta$ : *Tumoral growth factor  $\beta$*

TREAT: *Trial to Reduce Cardiovascular Events With Aranesp Therapy*

UKPDS: *United Kingdom Prospective Diabetes Study*

VIT.D: *Tratamiento con derivados activos de la vitamina D.*

**ANEXO III. RELACIÓN DE FIGURAS**

*Figura 1.1. Pronóstico de la enfermedad renal crónica en función de las categorías del FG y de la albuminuria.*

*Figura 1.2. Modelo conceptual para la enfermedad renal crónica (NFK)(Levey 2005).*

*Figura 2.1. Esquema de los métodos planteados en el análisis de los datos de la presente tesis doctoral.*

*Figura 2.2. Mapa de ganadoras del ejemplo propuesto.*

*Figura 2.3. Mapa de componentes del ejemplo propuesto.*

*Figura 2.4. Esquema del procedimiento Boruta.*

*Figura 2.5. Esquema del modelo Naïve Bayes.*

*Figura 2.6. Esquema del TAN (Tree Augmented Network).*

*Figura 2.7. Esquema de dos curvas ROC.*

*Figura 2.8. Esquema de un árbol de decisión (problema simulado).*

*Figura 3.1. Población estudiada y número de visitas realizadas por paciente.*

*Figura 3.2. Histograma que representa la edad y sexo de los pacientes estudiado.*

*Figura 3.3. Evolución de los pacientes al finalizar el estudio.*

*Figura 3.4. Nefropatía de base.*

*Figura 3.5. Representación de los diferentes diagramas de cajas (boxplot) de FG estimado por CKD-EPI para cada uno de los pacientes estudiados.*

*Figura 3.6. Histograma que representa la distribución de los valores del FG estimado por sexo.*

*Figura 3.7. Gráficos comparativos del filtrado glomerular estimado por las diferentes fórmulas empleadas.*

*Figura 3.8. Representación del índice de comorbilidad de Charlson y distribución por sexos.*

*Figura 3.9. Representación de la hipertensión arterial y su tratamiento.*

*Figura 3.10. Representación del índice de masa corporal.*

*Figura 3.11. Histogramas que muestran la distribución de las cifras de hemoglobina por sexos.*

*Figura 3.12. Tratamiento con suplementos de hierro.*

*Figura 3.13. Tratamiento con derivados eritropoyéticos.*

*Figura 3.14. Histograma que representa la dosis recibida de agentes eritropoyéticos.*

*Figura 3.15. Tratamiento con estatinas.*

*Figura 3.16. Tratamiento de diabetes mellitus.*

*Figura 3.17. Niveles de hemoglobina glicosilada.*

*Figura 3.18. Distribución de Anion GAP.*



*Figura 3.19. Niveles de calcio y fósforo.*

*Figura 3.20. Diagramas de cajas que representan el rango de fosfatemia para cada paciente.*

*Figura 3.21. Variabilidad individual en los niveles de PTH (diagrama de cajas).*

*Figura 3.22. Tratamiento con fármacos quelantes del fósforo.*

*Figura 3.23. Tratamiento con derivados activos de vitamina D.*

*Figura 3.24. Tratamiento de hiperuricemia con alopurinol.*

*Figura 3.25. Distribución de proteinuria.*

*Figura 3.26. Distribución de los niveles de albúmina por sexos.*

*Figura 3.27. Tratamiento diurético.*

*Figura 3.28. Tratamiento con inhibidores del sistema renina-angiotensina-aldosterona.*

*Figura 3.29. Hipertrofia ventricular izquierda.*

*Figura 3.30. Insuficiencia cardíaca.*

*Figura 3.31. Evolución del FGe en un paciente.*

*Figura 3.32. Representación de la trayectoria de descenso del filtrado glomerular mediante regresión lineal.*

*Figura 3.33. Esquema del procedimiento experimental planteado.*

*Figura 3.34. Curva sensibilidad/especificidad para el modelo lineal usando la proteinuria.*

*Figura 3.35. Curva sensibilidad/especificidad para el modelo lineal usando la proteinuria basal.*

*Figura 3.36. Curva sensibilidad/especificidad para el modelo lineal usando el fósforo.*

*Figura 3.37. Curva sensibilidad/especificidad para el modelo lineal usando la creatinina.*

*Figura 3.38. Curva sensibilidad/especificidad para el modelo lineal usando el calcio.*

*Figura 3.39. Curva sensibilidad/especificidad para el modelo lineal usando las variables estadísticamente significativas del análisis univariante.*

*Figura 3.40. Curva sensibilidad/especificidad para el Naïve Bayes usando las variables estadísticamente significativas del análisis univariante.*

*Figura 3.41. Curva sensibilidad/especificidad para el TAN usando las variables estadísticamente significativas.*

*Figura 3.42. Esquema de las relaciones encontradas en el modelo TAN.*

*Figura 3.43. Comparación entre los diferentes modelos variando los conjuntos de entrenamiento/validación.*

*Figura 3.44. Curva sensibilidad/especificidad para el modelo lineal usando las variables seleccionadas por el Boruta.*

*Figura 3.45. Curva sensibilidad/especificidad para el Naïve Bayes usando las variables seleccionadas por el Boruta.*

*Figura 3.46. Curva sensibilidad/especificidad para el TAN usando las variables seleccionadas por el Boruta.*

*Figura 3.47. Esquema de las relaciones encontradas en el modelo TAN.*

*Figura 3.48. Comparación entre los diferentes modelos variando los conjuntos de entrenamiento/validación.*

*Figura 3.49. Curva sensibilidad/especificidad para el modelo lineal usando las variables seleccionadas por el Boruta a partir de las estadísticamente significativas.*

*Figura 3.50. Curva sensibilidad/especificidad para Naïve Bayes usando las variables seleccionadas por el Boruta a partir de las estadísticamente significativas.*

*Figura 3.51. Curva sensibilidad/especificidad para TAN usando las variables seleccionadas por el Boruta a partir de las estadísticamente significativas.*

*Figura 3.52. Esquema de las relaciones encontradas en el modelo TAN tras aplicar Boruta a las variables significativas.*

*Figura 3.53. Árbol de decisión construido con las variables más importantes obtenidas en los modelos anteriores.*

*Figura 3.54. Curva sensibilidad/especificidad para el árbol de decisión desarrollado.*

*Figura 3.55. Representación mediante SOM.*

*Figura 3.56. Representación mediante SOM de la pendiente de descenso del FG.*

*Figura 3.57. Representación mediante SOM de la relación entre el sexo de los pacientes y la progresión de la enfermedad renal.*

*Figura 3.58. Representación de la relación entre el descenso del FG y la nefropatía de base.*

*Figura 3.59. Progresión de enfermedad renal e HTA.*

*Figura 3.60. Tabaco y progresión.*

*Figura 3.61. Relación entre DM y progresión.*

*Figura 3.62. Relación entre alteraciones en perfil lipídico y progresión.*

*Figura 3.63. Relación entre anemia, su tratamiento y progresión.*

*Figura 3.64. Relación entre HTA, su tratamiento y progresión.*

*Figura 3.65. Relación entre equilibrio ácido-base y progresión de la enfermedad renal.*

*Figura 3.66. Relación entre alteraciones del metabolismo óseo-mineral y progresión de la enfermedad renal.*

**ANEXO IV. RELACIÓN DE TABLAS**

*Tabla 3.1. Nefropatía de base.*

*Tabla 3.2. Correlación lineal entre variables continuas y el FGe por CKD\_EPI.*

*Tabla 3.3. Correlación lineal entre variables continuas y pendiente de modificación del FG.*

*Tabla 3.4. Características de los pacientes de ambos grupos.*

*Tabla 3.5. Variables continuas en las que se han encontrado diferencias significativas entre los pacientes con progresión rápida y lenta de la enfermedad renal.*

*Tabla 3.6. Coeficiente de regresión para proteinuria.*

*Tabla 3.7. Matriz de confusión para proteinuria.*

*Tabla 3.8. Coeficiente de regresión para proteinuria basal.*

*Tabla 3.9. Matriz de confusión para proteinuria inicial.*

*Tabla 3.10. Coeficiente de regresión para fósforo.*

*Tabla 3.11. Matriz de confusión para fósforo.*

*Tabla 3.12. Coeficiente de regresión para creatinina.*

*Tabla 3.13. Matriz de confusión para creatinina.*

*Tabla 3.14. Coeficiente de regresión para calcio.*

*Tabla 3.15. Matriz de confusión para calcio.*

*Tabla 3.16. Coeficientes de regresión de las diferentes variables en el modelo multivariante.*

*Tabla 3.17. Matriz de confusión para modelo de regresión logística.*

*Tabla 3.18. Matriz de confusión para Naïve-Bayes.*

*Tabla 3.19. Matriz de confusión para modelo TAN.*

*Tablas 3.20 y 3.21. Comparación de modelos multivariantes.*

*Tabla 3.22. Coeficiente de regresión para regresión logística.*

*Tabla 3.23. Matriz de confusión para regresión logística tras selección de variables por Boruta.*

*Tabla 3.24. Matriz de confusión para el conjunto de test.*

*Tabla 3.25. Matriz de confusión para el modelo TAN tras selección de variables por Boruta.*

*Tablas 3.26 y 3.27. Comparación de modelos.*

*Tabla 3.28. Coeficientes de regresión logística.*

*Tabla 3.29. Matriz de confusión para regresión logística.*

*Tabla 3.30. Matriz de confusión para regresión logística tras selección mixta de variables.*

*Tabla 3.31. Matriz de confusión para el modelo TAN.*

*Tabla 3.32. Matriz de confusión del árbol de decisión.*

