



Universidad
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES
DE POSGRADO Y DOCTORADO

ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL

Año académico 2016/17

DOCTORANDO: GOYA GATO, FRANCISCO

PROGRAMA DE DOCTORADO: D423-FARMACIA

DEPARTAMENTO DE: CIENCIAS BIOMÉDICAS

TITULACIÓN DE DOCTOR EN: DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ

En el día de hoy 05/07/17, reunido el tribunal de evaluación nombrado por la Comisión de Estudios Oficiales de Posgrado y Doctorado de la Universidad y constituido por los miembros que suscriben la presente Acta, el aspirante defendió su Tesis Doctoral, elaborada bajo la dirección de JOSEFA B. ESCALERA IZQUIERDO.

Sobre el siguiente tema: ESTUDIO DE DETERMINANTES DEL DAÑO RENO-CARDIOVASCULAR EN UNA MUESTRA DE POBLACIÓN DEL CORREDOR DEL HENARES

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL¹ de (no apto, aprobado, notable y sobresaliente): SOBRESALIENTE

Alcalá de Henares, 5 de Julio de 2017

EL PRESIDENTE


Fdo.: Francisco Pedragosa
García

EL SECRETARIO


Fdo.: Manuel Córdoba Díaz

EL VOCAL


Fdo.: EDUARDO PARRA

Con fecha 24 de Julio de 2017 la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado, a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- Conceder la Mención de "Cum Laude"
 No conceder la Mención de "Cum Laude"

La Secretaria de la Comisión Delegada



FIRMA DEL ALUMNO,


Fdo.: FRANCISCO GOYA

¹ La calificación podrá ser "no apto" "aprobado" "notable" y "sobresaliente". El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad.

INCIDENCIAS / OBSERVACIONES:

El presente informe tiene como finalidad informar a la Junta de Gobierno del desarrollo de las actividades programadas para el primer trimestre de 2024, así como de los resultados obtenidos y las incidencias observadas durante el periodo.

El informe se divide en tres partes: un resumen ejecutivo, un análisis detallado de las actividades y un apartado de conclusiones y recomendaciones.

El resumen ejecutivo muestra que se han cumplido con el 85% de las actividades programadas, con un presupuesto ejecutado del 70%. Se han observado incidencias en la gestión de recursos humanos y en la ejecución de algunos proyectos, que se detallan en el apartado de incidencias.

Las conclusiones y recomendaciones indican que se debe mejorar la gestión de recursos humanos y la ejecución de los proyectos para garantizar el cumplimiento de las actividades programadas en el resto del trimestre.

En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 24 de julio, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por *GOYA GATO, FRANCISCO*, el día 5 de julio de 2017, titulada *ESTUDIO DE DETERMINANTES DEL DAÑO RENO-CARDIOVASCULAR EN UNA MUESTRA DE POBLACIÓN DEL CORREDOR DEL HENARES*, para determinar, si a la misma, se le concede la mención "cum laude", arrojando como resultado el voto favorable de todos los miembros del tribunal.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado resuelve otorgar a dicha tesis la

MENCIÓN "CUM LAUDE"

Alcalá de Henares, 27 julio de 2017
EL PRESIDENTE DE LA COMISIÓN DE ESTUDIOS
OFICIALES DE POSGRADO Y DOCTORADO



Firmado digitalmente por VELASCO
PEREZ JUAN RAMON - DNI
03087239H
Fecha: 2017.07.30 18:15:09 +02'00'

Juan Ramón Velasco Pérez

Copia por e-mail a:
Doctorando: GOYA GATO, FRANCISCO
Secretario del Tribunal: MANUEL CÓRDOBA DÍAZ.
Directora de Tesis: JOSEFA B. ESCALERA IZQUIERDO



Universidad
de Alcalá

ESCUELA DE DOCTORADO
Servicio de Estudios Oficiales de
Posgrado

DILIGENCIA DE DEPÓSITO DE TESIS.

Comprobado que el expediente académico de D./D^a _____
reúne los requisitos exigidos para la presentación de la Tesis, de acuerdo a la normativa vigente, y habiendo
presentado la misma en formato: soporte electrónico impreso en papel, para el depósito de la
misma, en el Servicio de Estudios Oficiales de Posgrado, con el nº de páginas: _____ se procede, con
fecha de hoy a registrar el depósito de la tesis.

Alcalá de Henares a _____ de _____ de 20 _____



Fdo. El Funcionario



PROGRAMA DE DOCTORADO EN BIOMEDICINA

**ESTUDIO DE DETERMINANTES DEL DAÑO RENO-CARDIOVASCULAR EN
UNA MUESTRA DE POBLACIÓN DEL CORREDOR DEL HENARES.**

Tesis Doctoral presentada por

FRANCISCO GOYA GATO

Directora:

DRA. BEGOÑA ESCALERA IZQUIERDO

Alcalá de Henares, 2017



Universidad
de Alcalá

DPTO. CIENCIAS BIOMÉDICAS
UNIDAD DOCENTE FARMACIA Y TECNOLOGÍA
FARMACÉUTICA
Facultad de Farmacia
Ctra. Barcelona, Km. 33,600
28871 Alcalá de Henares (Madrid)
Telf.: 91 885 46 57
e-mail: dpto.biomedicas@uah.es

**FRANCISCO ZARAGOZÁ GARCÍA, CATEDRÁTICO DE
FARMACOLOGÍA Y DIRECTOR DEL DEPARTAMENTO DE
CIENCIAS BIOMÉDICAS DE LA UNIVERSIDAD DE ALCALÁ**

CERTIFICA:

Que la memoria para optar al Grado de Doctor, elaborada por Francisco Goya Gato, ha sido realizada en este Departamento bajo la dirección de la Dra. Begoña Escalera Izquierdo, Profesora Titular del Departamento de Ciencias Biomédicas de esta Universidad, y se encuentra ya concluida, por lo que autorizo su presentación a fin de que pueda ser juzgada por el tribunal correspondiente.

Y para que así conste, firmo el presente certificado en Alcalá de Henares a diecinueve de abril de dos mil diecisiete.





Universidad
de Alcalá

DPTO. CIENCIAS BIOMÉDICAS
UNIDAD DOCENTE FARMACIA Y TECNOLOGIA
FARMACÉUTICA
Facultad de Farmacia
Ctra. Barcelona, Km. 33,600
28871 Alcalá de Henares (Madrid)
Telf.: 91 885 46 57
e-mail: dpto.biomedicas@uah.es

DÑA. BEGOÑA ESCALERA IZQUIERDO, PROFESORA TITULAR DEL
DEPARTAMENTO DE CIENCIAS BIOMÉDICAS DE LA
UNIVERSIDAD DE ALCALÁ

CERTIFICA:

Que la memoria para optar al Grado de Doctor, elaborada por Francisco Goya Gato, cuyo título es

“Estudio de Determinantes del Daño Reno-Cardiovascular en una Muestra de Población del Corredor del Henares”

ha sido realizada bajo mi dirección y reúne los requisitos necesarios para su juicio y calificación.

Y para que así conste, firmo el presente certificado en Alcalá de Henares a diecinueve de abril de dos mil diecisiete.

3 Escalera

AGRADECIMIENTOS

A la Dra. Begoña Escalera Izquierdo, mi directora, por confiar en mí, y guiarme en este largo camino.

A los Dres. Francisco J. de Lucio Cazaña y Manuel R. Ramírez Chamond, auténticos artífices de esta Tesis. Sin vosotros no habría sido posible.

A los Dres. Pedro Miranda Menéndez y Nirian Martín Apaloaza, por hacerme más sencillo el mundo de la estadística.

A los Dres. Fidel Ortega Ortiz de Apodaca y Rosario Aberturas Ramos, mis directores del DEA. Con vosotros empezó el todo.

A mi hermano Ramón, por ser una guía constante y un espejo en el que mirarme cada día.

A mis padres, Ramón y Charo, porque gracias a su esfuerzo y sobre todo su ejemplo, han hecho de mí, la persona que soy.

A mi mujer, Elisa, por su amor, su compañía, su comprensión, su paciencia, su apoyo constante, y por darme lo mejor que me ha pasado en mi vida.

A Elisa

A mis padres

-INDICE-

GLOSARIO DE ABREVIATURAS

<u>1.- INTRODUCCIÓN</u>	1
1.1.- SÍNDROME CARDIO-RENAL	3
1.2.- ENFERMEDAD RENAL CRÓNICA	5
1.3.- ENFERMEDAD CARDIOVASCULAR	7
1.4.- FILTRADO GLOMERULAR COMO ÍNDICE DE LA FUNCIÓN RENAL	8
1.4.1.- Aclaramiento de Inulina.....	10
1.4.2.- Concentración sérica de creatinina.....	10
1.4.3.- Aclaramiento de creatinina.....	10
1.4.4.- Ecuaciones de estimación del filtrado glomerular.....	11
<u>2.- RELACIÓN HISTÓRICA ENTRE ENFERMEDAD CARDIOVASCULAR Y ENFERMEDAD RENAL CRÓNICA: BASES PARA UNA HIPÓTESIS</u>	13
<u>3.- OBJETIVOS Y PLAN DE TRABAJO</u>	19
3.1.- OBJETIVOS	21
3.2.- PLAN DE TRABAJO	21
<u>4.- MATERIAL Y MÉTODOS</u>	23
4.1.- MÉTODO DE EXTRACCIÓN SANGUÍNEA	25
4.2.- DETERMINACIÓN DE CREATININA	25
4.3.- DETERMINACIÓN DE HEMOGLOBINA GLICOSILADA	26
4.4.- DETERMINACIÓN DE LEUCOCITOS	27
4.5.- DETERMINACIÓN DE VELOCIDAD DE SEDIMENTACIÓN GLOMERULAR	29
4.6.- DETERMINACIÓN DE COLESTEROL	30

4.7.- DETERMINACIÓN DE TRANSAMINASAS	30
4.7.1.- Determinación de Glutamato-oxalacetato transaminasa (GOT).....	30
4.7.2.- Determinación de Glutamato-piruvato transaminasa (GPT).....	31
4.8.- DETERMINACIÓN DE GAMMA GLUTAMIL TRANSPEPTIDASA (GGT)	32
4.9.- ELECCIÓN DEL TIPO DE ANÁLISIS ESTADÍSTICO	33
4.10.- AGRUPACIÓN DE LAS MUESTRAS PARA EL ESTUDIO ESTADÍSTICO	34
<u>5.- RESULTADOS Y DISCUSIÓN</u>	37
5.1.- VALORACIÓN PROMEDIO DE LAS VARIABLES EXPLICATIVAS.....	39
5.2.- ESTUDIO ESTADÍSTICO.....	49
<u>6.- CONCLUSIONES</u>	119
<u>7.- BIBLIOGRAFÍA</u>	123

GLOSARIO DE ABREVIATURAS

ALT: Alanina aminotransferasa

AST: Aspartato aminotransferasa

ECV: Enfermedad cardiovascular

EDTA: Ácido etilendiaminotetraacético

EPIRCE: Estudio sobre la Epidemiología de la Insuficiencia Renal Crónica en España

ERC: Enfermedad renal crónica

FGR: Filtrado glomerular

GGT: Gamma glutamil transpeptidasa

GOT: Glutamato oxalacetato transaminasa

GPT: Glutamato piruvato transaminasa

HbA1C: Hemoglobina glicosilada

K/DOQI: Kidney Disease Outcomes Quality Initiative

MDRD: Modification of Diet in Renal Disease

NKDPEP: National Kidney Disease Education Program

SCR: Síndrome cardio-renal

VSG: Velocidad de sedimentación glomerular

1.- INTRODUCCIÓN

1.1 SÍNDROME CARDIO-RENAL

Las alteraciones combinadas del corazón y riñón se clasifican hoy en día como Síndrome cardio-renal (SCR). La definición más reciente incluye alteraciones agudas o crónicas donde el órgano afectado puede ser primariamente el corazón y secundariamente el riñón o viceversa, así se han clasificado 5 tipos de SCR, como se muestra en la figura 1 (Ronco C et al, 2010), dependiendo del órgano inicialmente lesionado y la naturaleza aguda o crónica del daño (Santamaría R et al, 2012):

Síndrome cardio-renal tipo 1
Deterioro agudo de la función cardíaca que produce un daño y/o disfunción renal aguda
Síndrome cardio-renal tipo 2
Anomalías crónicas de la función cardíaca que conducen a disfunción renal crónica
Síndrome cardio-renal tipo 3
Deterioro agudo de la función renal que conduce a un deterioro agudo de la función cardíaca
Síndrome cardio-renal tipo 4
Anomalías crónicas de la función renal que conducen a enfermedad cardíaca
Síndrome cardio-renal tipo 5
Condiciones sistémicas que causan de manera simultanea una disfunción cardíaca y renal

Figura 1: Clasificación del síndrome cardio-renal

La fisiopatología del síndrome cardio-renal no se entiende completamente, pero puede ser causada por una combinación compleja de activación del sistema neurohormonal, disfunción endotelial, proteinuria, estrés oxidativo, toxinas urémicas y otros factores. El manejo del síndrome cardio-renal sigue siendo un desafío terapéutico importante en la práctica clínica porque muchos de los fármacos usados para controlar la insuficiencia cardíaca pueden empeorar la función renal y viceversa (Fu Q et al, 2014; Núñez J et al, 2015).

Los factores predisponentes del SCR son fundamentalmente (Lekuona I, 2012):

Obesidad: La glomerulopatía asociada a la obesidad ha sido descrita como una situación de hiperfiltración en sujetos obesos sin DM que al final lleva a ERC y predispone al SCR tipo 1. De forma añadida el Síndrome metabólico en ausencia de DM franca se asocia a un incremento de 3-7 veces del SCR tipo 1 en diferentes situaciones clínicas.

Caquexia: problemas combinados del corazón y del riñón se desarrollan en presencia de caquexia y sarcopenia y que se asocian a una interrelación de los 2 órganos a través de factores como el de necrosis tumoral alfa y otras citoquinas que van a aumentar su concentración de 6 a 8 veces.

Proteinuria: el daño endotelial, mesangial o de los podocitos en presencia de HTA y DM va a dar lugar a un exceso de albúmina en la cápsula de Bowman, las células tubulares proximales deben reabsorber este exceso lo que las lleva a apoptosis a pérdida de nefronas y a progresión de la enfermedad renal. La albuminuria predice el desarrollo de IC, y en aquellos con IC establecida está presente en un 30% de los casos y se asocia con aumento de la hospitalización y de la mortalidad. La micro-albuminuria es un marcador de enfermedad renal y probablemente un factor patogénico en la progresión de la ERC.

Retención de solutos urémicos: La uremia causa disfunción de los miocitos alterando el movimiento del calcio en el citosol afectando la contracción, contribuye a acelerar la fibrosis y el remodelado cardiaco adverso después de un infarto agudo de miocardio.

Anemia. Es común en la insuficiencia cardiaca y se asocia a incrementos de la mortalidad y morbilidad, así como empeoramiento de la función.

Episodios repetidos de daño renal subclínico. Es altamente probable que algunos individuos hayan sufrido episodios repetidos de daño renal subclínico a lo largo de su vida.

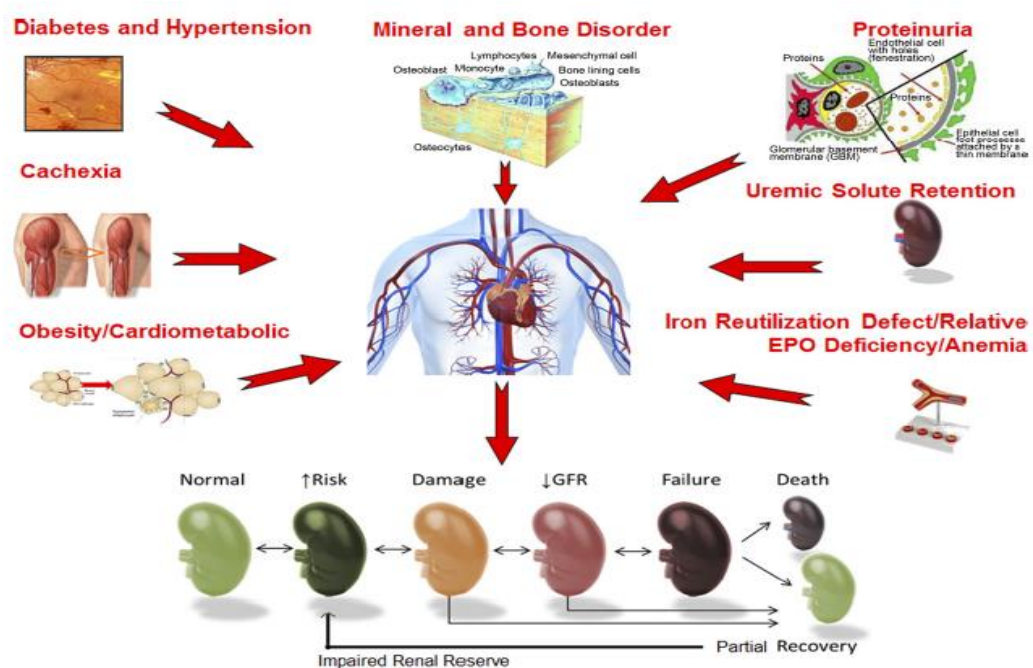


Figura 2: Factores predisponentes SRC. Ronco C. 2012

Respecto al tratamiento, como no hay uno definitivo disponible, la prevención debe ser una estrategia clave. Cualquier paciente con con disfunción del corazón o del riñón corre un alto riesgo de desarrollar un SCR. Dada que la congestión juega un papel clave en la patogénesis de casi todos los tipos de SRC, prevenir la sobrecarga de volúmen es una parte clave de la gestión de los pacientes con este síndrome, además son importantes las estrategias de prevención primarias centradas en la modificación de factores de riesgo como la obesidad, la hipertensión o la Diabetes Mellitus (*Gnanaraj J. et al, 2016*).

1.2 ENFERMEDAD RENAL CRONICA

La Enfermedad Renal Crónica (ERC), se ha convertido en un problema de salud mundial, encontrándose España, según los datos de la Sociedad Española de Nefrología (SEN), entre los países europeos, con alta prevalencia de ERC terminal, y al estar el problema estrechamente relacionado con el envejecimiento de la población y la elevada prevalencia de diabetes e hipertensión arterial, se prevé que podría aumentar en los próximos años. El estudio EPIRCE, sitúa en un 11% la población adulta en España que sufre ERC, en alguno de sus estadios. La presencia de ERC se ha relacionado con un riesgo elevado de Insuficiencia Renal Crónica terminal (IRct), enfermedad cardiovascular y muerte (*Otero A et al, 2005*).

Ante el previsible aumento de la población con ERC, los esfuerzos se están dirigiendo hacia la prevención y detección precoz de la enfermedad, para evitar o enlentecer su progresión, por ello en 2002, la National Kidney Foundation estadounidense, a través del proyecto K/DOQI, publicó una serie de guías de práctica clínica sobre la evaluación, clasificación y estratificación de la ERC, que han sido aceptadas por la mayoría de las Sociedades Científicas (*National Kidney Foundation, 2002*). Los objetivos de estas guías fueron los siguientes:

- La definición de Enfermedad Renal Crónica y su clasificación en estadios a partir del valor del filtrado glomerular y la presencia de lesión renal, independientemente de la etiología de la misma.
- Determinar las pruebas de laboratorio más adecuadas para la evaluación de la función renal.
- Asociar los estadios de la función renal con las complicaciones de la ERC.

- Estratificar el riesgo de progresión de la ERC y de desarrollar complicaciones cardiovasculares.

Según estas guías, se entiende por Enfermedad Renal Crónica (ERC):

- La presencia de un filtrado glomerular menor de $60\text{ml}/\text{min}/1,73\text{m}^2$, durante un tiempo igual o superior a 3 meses.
- La presencia de lesión renal, con o sin descenso del filtrado glomerular, durante un tiempo igual o superior a 3 meses. Por lesión renal, se entiende la presencia de alteraciones estructurales o funcionales del riñón, puestas de manifiesto:
 - Directamente por alteraciones histológicas en una biopsia renal.
 - Indirectamente, por la presencia de albuminuria, proteinuria, alteraciones en el sedimento urinario, o mediante técnicas de imagen.

La combinación de ambos criterios, es la base para la clasificación de la Enfermedad Renal Crónica en 5 estadios:

Tabla 1: Clasificación en estadios de la enfermedad renal crónica según las guías de la K/DOQI 2002 de la National Kidney Foundation (Soriano, S 2004)

Estadio	Descripción	Filtrado glomerular (mL/min/1,73 m ²)
1	Lesión renal con filtrado glomerular normal o aumentado	≥ 90
2	Lesión renal con disminución leve del filtrado glomerular	60 – 89
3	Disminución moderada del filtrado glomerular	30 – 59
4	Disminución severa del filtrado glomerular	15 – 29
5	Fallo renal o diálisis	< 15

Se ha seleccionado el valor de corte del filtrado glomerular en $60\text{ ml}/\text{min}/1.73\text{m}^2$, como definición de ERC, porque este valor representa una reducción de más del 50%, respecto del valor normal del mismo.

En los estadios 1 y 2, el valor de filtrado glomerular no es diagnóstico por sí mismo y precisa de la presencia de algún marcador asociado a lesión renal. Es más, el estadio 2, en ausencia de marcador de lesión renal, se cataloga como “descenso del filtrado glomerular” y no como ERC (Soriano, S 2004).

1.3 ENFERMEDAD CARDIOVASCULAR

Se trata de un concepto amplio que engloba al corazón y los vasos sanguíneos. Con frecuencia va ligada a la arterioesclerosis, que consiste en una acumulación de grasas y colesterol en las arterias en forma de placa, produciendo un engrosamiento de la pared arterial, y que con el paso de los años, esta placa disminuye la luz de las arterias. Se produce asimismo una pérdida de elasticidad arterial; en el corazón se bloquean de manera progresiva las arterias coronarias y sus ramas principales. Se trata de un proceso crónico de lenta evolución, y su prevalencia es mayor en hombres que en mujeres. (Newby DE et al, 2012; Goldman L et al, 2016; Toth PP et al, 2016)

Respecto a los factores de riesgo, no fue hasta finales de los años 40, cuando se inició un estudio epidemiológico en una población del estado de Massachusetts, el que ha generado la mayor parte de la evidencia sobre el origen de la enfermedad isquémica y con posterioridad del resto de enfermedades cardiovasculares con sustrato ateromatoso. Desde 1947 la ciudad de Framingham, próxima a Boston fue seleccionada para realizar un estudio comunitario que proporcionara información sobre la historia natural de la enfermedad isquémica cardíaca, datos de incidencias reales y su pronóstico. Fue el NHBLI (Instituto Nacional de Corazón, Pulmón y Sangre) el encargado de dirigirlo, en la figura de Thomas R. Dawber. Inicialmente el estudio contó con 5.209 hombres y mujeres de entre 30 y 62 años exentos de enfermedad cardíaca y cerebro-vascular. Entre sus avances cabe destacar el haber utilizado un estudio prospectivo, que proporcionó una evidencia de asociación causal más potente que los estudios clínicos transversales previos, y además dio origen al concepto *factor de riesgo*, como condición que precede al comienzo de la enfermedad y aumenta el riesgo de padecerla (Dawber T et al, 1962). Con los años el seguimiento de esta población ha llevado a la identificación de los principales factores de riesgo asociados a la enfermedad, como pueden ser (Eckel RH et al, 2014):

- **Aumento del colesterol:** Desde 1961 había reunida suficiente evidencia para asegurar que los niveles séricos elevados de colesterol predecían la aparición de enfermedad isquémica cardíaca

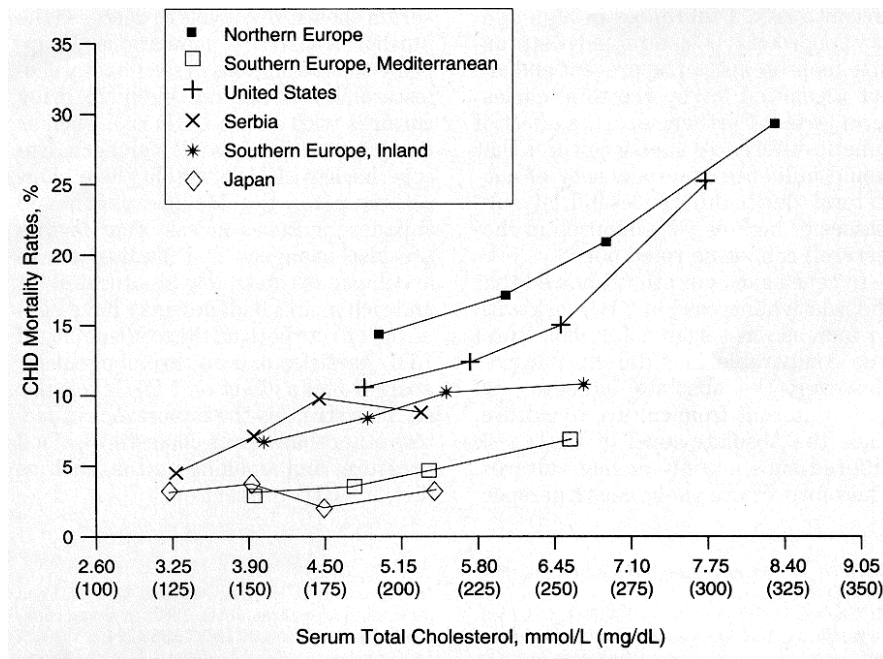


Figura 3: Riesgo absoluto entre mortalidad por cardiopatía isquémica y valores de colesterol total entre distintos países y regiones (Verschuren M et al, 1995).

- **Hipertensión arterial:** Mucho antes de los estudios de Framingham se conocía la relación entre cifras elevadas de presión arterial y mortalidad cardiaca. Se considera un factor de riesgo mayor para todas las enfermedades cardiovasculares.
- **Tabaquismo:** Se evidenció una relación causal con el ictus isquémico y la arteriopatía periférica.
- **Diabetes Mellitus:** Fue la más tardía en identificarse dentro de los factores principales para la cardiopatía isquémica, fundamentalmente por el escaso número de pacientes diabéticos que participaron inicialmente en el estudio. El seguimiento posterior de los pacientes ha podido asociarla con enfermedades cardiovasculares como la arteriopatía diabética, la insuficiencia cardiaca y el ictus.
- **Obesidad:** Hoy en día se considera que la acumulación de grasa perivisceral abdominal es la que induce mayor riesgo a través de la producción de diversas hormonas u mediadores inflamatorios que inducen la resistencia a la insulina, hipertensión arterial, cambios en el perfil lipémico y actividad procoagulante, lo que a su vez incrementa el riesgo de padecer diabetes y enfermedad cardiovascular.

- **Edad:** Las personas mayores tienen un mayor riesgo de sufrir enfermedades del corazón. Aproximadamente 4 de cada 5 muertes debidas a una enfermedad cardíaca se producen en personas mayores de 65 años de edad. Con la edad, la actividad del corazón tiende a deteriorarse. Puede aumentar el grosor de las paredes del corazón, las arterias pueden endurecerse y perder su flexibilidad y, cuando esto sucede, el corazón no puede bombear la sangre tan eficientemente como antes a los músculos del cuerpo.
- **Sexo:** En general, los hombres tienen un riesgo mayor que las mujeres de sufrir un ataque al corazón. Gracias a sus hormonas sexuales, las mujeres generalmente están protegidas de las enfermedades del corazón hasta la menopausia, que es cuando su riesgo comienza a aumentar.

1.4 FILTRADO GLOMERULAR COMO INDICE DE LA FUNCION RENAL

La valoración del filtrado glomerular es el mejor indicador para medir la función renal, la monitorización del mismo permite calcular la progresión de la enfermedad renal y también ajustar las dosis de los fármacos que se eliminan por vía renal, y evitar así posibles problemas de toxicidad.

El valor del filtrado glomerular, varía en función de la edad, sexo y la masa corporal, siendo 140 ml/min/1,73 m² el valor considerado normal en individuos adultos, jóvenes sanos (*Gracia S et al, 2006*).

La forma ideal de medirlo es a través del aclaramiento de una sustancia, que es el volumen del plasma, del que ésta es totalmente eliminada por unidad de tiempo (*Swan SK, 1997*). Lo ideal es usar una molécula cuya concentración plasmática sea estable, que sea fisiológicamente inerte, que se filtre libremente por el glomérulo renal y que no sea sintetizada, reabsorbida, metabolizada ni secretada por el riñón, de forma que toda la cantidad que se filtra en el glomérulo renal aparezca en la orina.

1.4.1 ACLARAMIENTO DE INULINA

La inulina es una molécula que cumple con gran parte de las características antes mencionadas, y por ello se la considera el “gold-standard” para la medida del filtrado glomerular. El problema es que es una molécula exógena que necesita ser perfundida, por lo que su medida es muy compleja a nivel del laboratorio clínico (*Gracia S et al, 2006; Gómez A et al, 2009*).

1.4.2 CONCENTRACIÓN SÉRICA DE CREATININA

La concentración sérica de creatinina, es la medida que se utiliza habitualmente para evaluar la función renal, pues tiene una serie de características que se aproximan a las de la molécula ideal para la medición del filtrado glomerular; sin embargo está afectada por distintas fuentes de variabilidad biológica, como son la edad, el sexo, la raza, la masa muscular o el tipo de dieta, múltiples interferencias analíticas e importantes problemas de estandarización. Por todos estos inconvenientes, no se debe utilizar únicamente el valor de creatinina en sangre para evaluar la función renal (*Gowans EM et al, 1988; Swedko PJ et al, 2003*).

1.4.3 ACLARAMIENTO DE CREATININA

El aclaramiento de creatinina, calculado a partir del valor de la concentración sérica de creatinina, y de su excreción en orina de 24 horas, es el método más utilizado como medida del filtrado glomerular, sin embargo también tiene una serie de limitaciones (*Payne RB, 2000; Gracia S et al, 2006*)

- Sobreestima el filtrado glomerular, porque la creatinina, además de ser filtrada por el glomérulo renal, es secretada por los túbulos renales.
- Inconveniente de tener que recoger una orina de 24 horas, con los errores que ello conlleva, sobre todo en niños y ancianos.

1.4.4 ECUACIONES PARA LA ESTIMACIÓN DEL FILTRADO GLOMERULAR

Las ecuaciones desarrolladas para la estimación del filtrado glomerular, han intentado eliminar los problemas anteriormente mencionados para la creatinina y el aclaramiento de creatinina. Tratan de obtener una estimación del mismo a partir de la concentración sérica de creatinina y de algunas variables demográficas y antropométricas (edad, sexo, peso, talla, etnia). La valoración del filtrado glomerular a través de estas ecuaciones es más exacta y precisa (Landry DW et al, 2016).

La primera ecuación de este tipo que se desarrolló, fue la de Cockcroft-Gault, que lo que estimaba era el aclaramiento de creatinina (Cockcroft DW, Gault MH, 1976):

$$\text{Aclaramiento de creatinina estimado} = \frac{(140 - \text{edad}) \times \text{peso}}{72 \times \text{creatinina}} \times (0,85 \text{ si mujer}) \quad (\text{ec. 1})$$

Hoy en día, la ecuación más utilizada es la del estudio MDRD (*Modification of Diet in Renal Disease, 1992*), que mejora la ecuación de Cockcroft-Gault, puesto que esta es una estimación del filtrado glomerular (FGR) y no del aclaramiento de creatinina (Stevens LA et al, 2007).

En concreto, el FGR viene dado por la siguiente expresión:

$$\text{FG estimado} = 186 \times (\text{creatinina (mg/dl)})^{-1,154} \times (\text{edad})^{-0,203} \times (0,742 \text{ si mujer}) \times (1,210 \text{ si raza negra}) \quad (\text{ec. 2})$$

A pesar de que hoy en día están en estudio otras ecuaciones, como puede ser la de CKD-EPI, la que todas las sociedades científicas recomiendan utilizar es la MDRD-4 (Levey AS et al, 1999; Gracia S et al, 2006; Alcázar R, et al, 2008). Informando, siempre que se pida una determinación de creatinina, en un análisis de sangre, de este valor, que se informará con el valor obtenido siempre que sea inferior a 60 ml/min/1,73 m² y se informará como mayor de 60 ml/min/1,73m², siempre que el resultado obtenido sea superior a ese valor.

Esto se hace así para homogeneizar los datos, en el sentido que los valores en las dos poblaciones sean comparables, es decir, estén medidas en la misma escala; de esta forma, se pondrá el mismo valor en los dos casos como frontera entre los individuos con un valor FGR significativo.

**2.- RELACIÓN HISTÓRICA ENTRE ENFERMEDAD
CARDIOVASCULAR Y ENFERMEDAD RENAL CRÓNICA: BASES
PARA UNA HIPÓTESIS.**

Las muertes cardiovasculares atribuidas a la disfunción renal superan en número a las muertes por la propia insuficiencia renal en todo el mundo. En 2013 la insuficiencia renal se asoció con el 4% de las muertes en todo el mundo, 2,2 millones de fallecimientos. Más de la mitad de estas, 1,2 millones fueron cardiovasculares, mientras que solo 0,96 millones fueron por enfermedad renal (Thomas B et al, 2017). La figura 4 muestra las muertes atribuidas a la reducción de la tasa de filtrado glomerular por 100.000 habitantes en 2013 en todo el mundo.

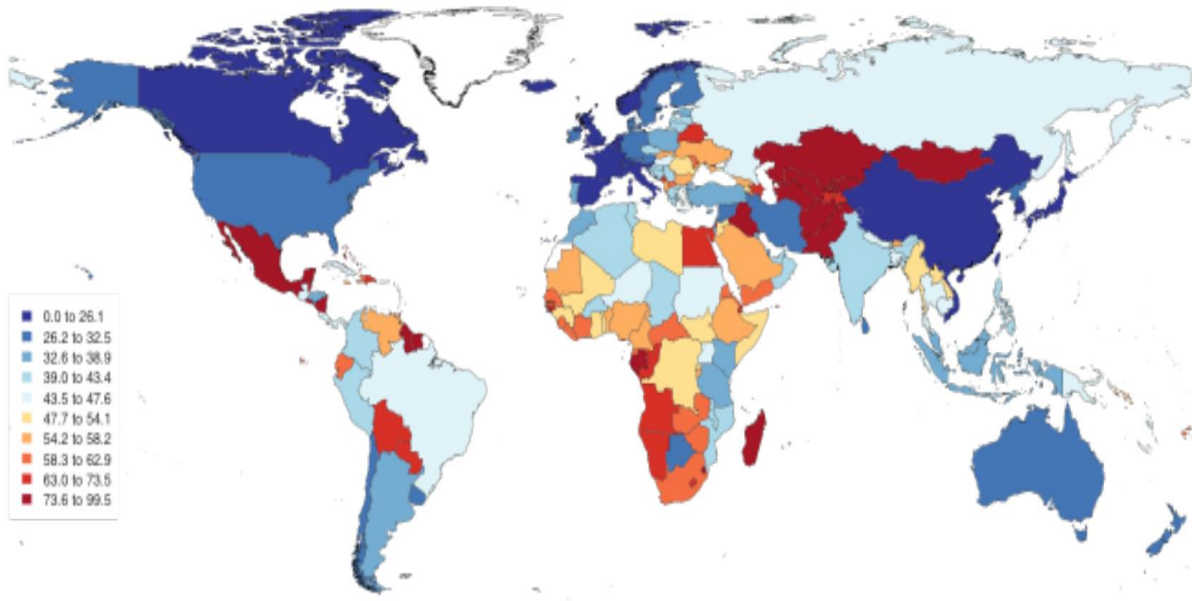


Figura 4: Muertes atribuidas a la reducción de la tasa de filtrado glomerular por 100.000 habitantes. (Thomas B et al 2017.)

La relación entre el riñón y la patología vascular es muy estrecha. El riñón está implicado de una forma u otra en el origen de la hipertensión y es fundamental en la perpetuación de ésta. Además, el riñón es afectado por la hipertensión, convirtiéndose en un órgano de los llamados diana, y su daño, es usado en la valoración del riesgo de estos pacientes (Siu AL et al, 2015; Victor RG 2016). Se ha demostrado en diversos estudios epidemiológicos que la insuficiencia renal desde estadios muy iniciales es un factor de riesgo cardiovascular de primer orden. La gran mayoría de estos pacientes nunca llegarán a diálisis pues fallecen antes, de la pérdida de la función renal por causas cardiovasculares. (De Francisco A, Gonzalez E, 2004; Inker LA et al, 2015)

Los pacientes con insuficiencia renal crónica (IRC) presentan un aumento muy importante de la morbimortalidad cardiovascular en relación a la población general. Entre el 40% y el 75% de los pacientes que comienzan los programas de diálisis tiene enfermedad cardiovascular (ECV). La ECV es responsable del 44% de las muertes de pacientes en esta

situación y constituye, tras los ajustes para edad y sexo, la causa más importante de morbilidad cardiovascular y de mortalidad total (Go AS et al, 2004; Tonelli M et al, 2006).

La elevada prevalencia de ECV en la población de diálisis sugiere que el daño vascular puede comenzar en los estadios iniciales de la enfermedad renal crónica (ERC). Aunque los estudios epidemiológicos de la ECV en dichos estadios son escasos, existen evidencias crecientes de su interrelación (Levin A et al, 1996). En los informes iniciales de los grandes estudios epidemiológicos (Framingham, MRFIT, *Multiple Risk Factors Intervention Trial*) la concentración de creatinina sérica no fue incluida como parámetro de análisis entre los posibles factores de riesgo cardiovascular. Los pacientes con ERC, pese a ser un grupo con alto riesgo cardiovascular, fueron sistemáticamente excluidos de la mayoría de los ensayos terapéuticos clásicos. En los últimos años se han publicado varios estudios que han demostrado que los niveles de creatinina sérica ligeramente elevados constituyen un factor de riesgo cardiovascular independiente. Se trata de análisis *post hoc* sobre grupos de pacientes con patología cardiovascular asociada. En el estudio HOPE (*Heart Outcomes and Prevention Evaluation*),(Mann JFE et al, 2001) los 980 enfermos que tenían una creatinina basal mayor o igual a 1,4 mg/dl tuvieron una prevalencia de infarto de miocardio, muerte cardiovascular y mortalidad total más alta que los 8.307 pacientes con niveles de creatinina menor de 1,4 mg/dl. Resultados similares se comunicaron en el estudio HOT (*Hypertension Optimal Treatment*) (Ruilope KM et al, 2001). También se han publicado los datos conjuntos de cuatro estudios poblacionales básicos (*Atherosclerosis Risk in Communities, Cardiovascular Health Study, Framingham Heart Study y Framingham Offspring Study*) que incluyeron un total de 22.634 individuos. Tanto basalmente como tras el seguimiento medio de 86 meses, pudo comprobarse que los sujetos con ERC, definida en este análisis por un filtrado glomerular (FG) estimado < 60 ml/min, tuvieron una mayor prevalencia de morbilidad vascular y mortalidad total que aquellos con un filtrado glomerular más alto. La ERC fue por sí misma un factor de riesgo CV independiente y tan poderoso como otros factores bien reconocidos como la diabetes, la presión arterial (PA) sistólica o la hipertrofia ventricular izquierda (HVI). (Weiner DE et al, 2004).

El informe americano JNC-7 sobre HTA ha considerado, por primera vez, a la microalbuminuria elevada y al descenso del filtrado glomerular estimado por debajo de 60 ml/min como factores de riesgo cardiovascular mayores. (Chovanian AV et al, 2003).

En la nefropatía diabética y en la nefropatía vascular los mecanismos etiopatogénicos de la enfermedad renal y de la ECV son, probablemente, muy similares y el progreso de ambos procesos ocurre en muchos casos de un modo simultáneo (Marín R, 2004), como se muestra en la figura 5.

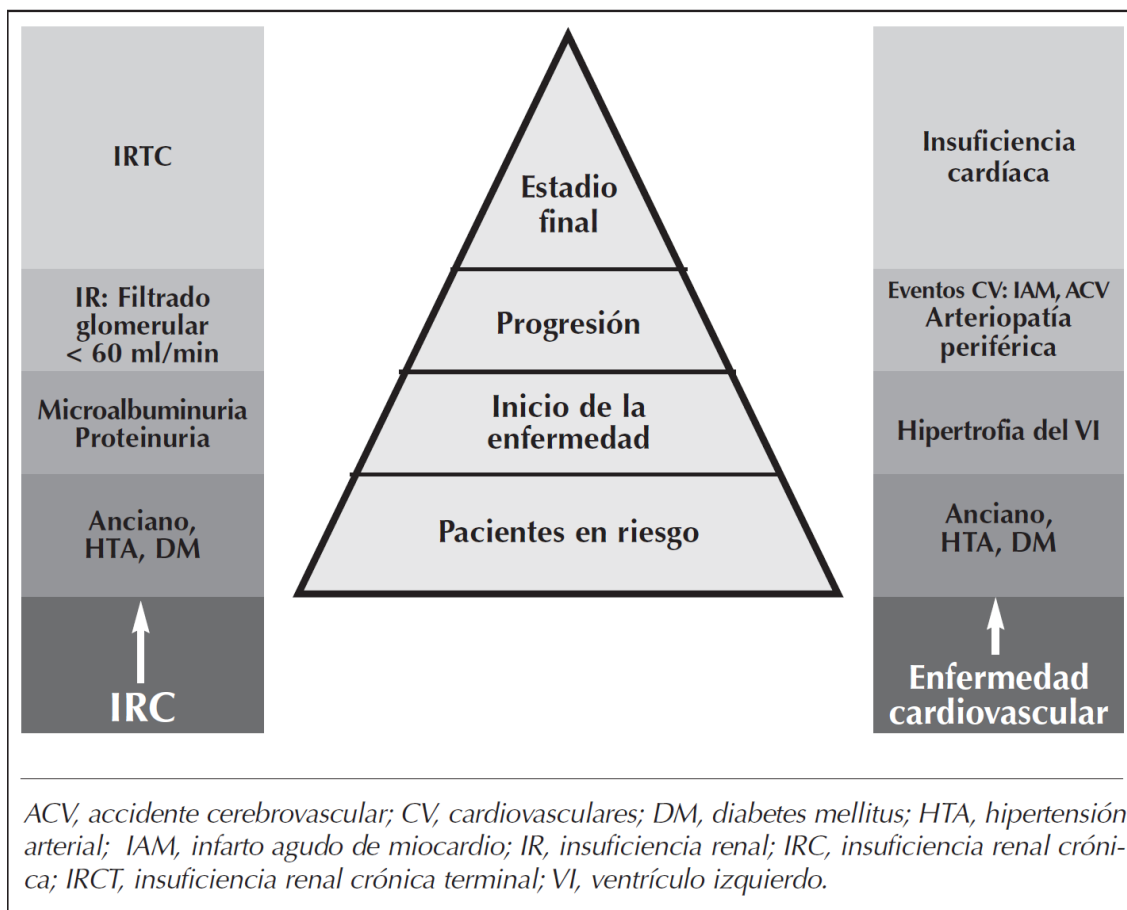


Figura 5: Estadios evolutivos paralelos de la enfermedad renal crónica y de la enfermedad cardiovascular (Sarnak MJ y Levey AS, 2000).

En el presente estudio se parte de la hipótesis de que es posible encontrar un modelo que explique la posibilidad de llegar a tener una disminución del filtrado glomerular (FGR) a partir de una serie de marcadores. Para ello, se analizaron una serie de variables que se considera que pueden influir en el valor de FGR, lo que se conoce como variables explicativas, son las siguientes:

Leucocitos, Velocidad de sedimentación glomerular en una hora, Colesterol total, Transaminasa GOT, Transaminasa GPT, Gamma-glutamyl transpeptidasa, Hemoglobina glicosilada.

El uso de estas variables se debe a que tanto los leucocitos como la velocidad de sedimentación glomerular son marcadores agudos de inflamación y por tanto estarán elevados en cuanto haya una lesión renal. Las transaminasas, son indicadores de una posible alteración hepática. Y por último el colesterol elevado como marcador de una posible dislipemia y la hemoglobina glicosilada como indicador de Diabetes Mellitus. Ésta última es la principal causa de fallo renal crónico por la alteración renal que produce (Fox CS et al, 2012).

3.- OBJETIVO Y PLAN DE TRABAJO

3.1 OBJETIVO

El objetivo del presente proyecto es caracterizar en una cohorte de individuos de la población del Corredor del Henares, sin patología renal aparente, si el filtrado glomerular estimado (FGR) muestra algún tipo de dependencia con los valores de analitos como:

- Colesterol.
- Recuento de leucocitos,
- Velocidad de sedimentación globular
- Aquellos asociados a deterioro hepático como la Glutamato-oxalacetato transaminasa (GOT), la Glutamato-piruvato transaminasa (GPT) o la Gamma glutamil transpeptidasa (GGT).
- Aquellos asociados al estado de glucemia del individuo como la hemoglobina glicosilada.

De esta manera podríamos ver la posible relación existente en personas que tienen valores de analitos asociados a una posible enfermedad cardiovascular elevados, si a su vez tienen un valor de FGR que pueda ser sugerente de llegar a desarrollar una enfermedad renal crónica.

Para la consecución del objetivo propuesto se describe a continuación el plan de trabajo seguido.

3.2 PLAN DE TRABAJO:

- Solicitud al Comité de Ética de la investigación y de experimentación animal de la Universidad de Alcalá, de la autorización para la realización del estudio.
- Selección de pacientes: Para realizar este estudio se dispone de muestras de 9.750 individuos, obtenidas desde enero de 2013 hasta diciembre de 2015, que parece un tamaño más que suficiente para que sea representativo de toda la población de referencia. Se ha tomado la información de todos los pacientes, por lo que no se ha utilizado ninguna aleatorización en la selección de individuos. Además, en este estudio, cada dato corresponde a un individuo en un momento concreto y no se

hace un seguimiento del mismo, por lo que el estudio no puede ser calificado ni prospectivo ni retrospectivo. Aunque, desde luego, se trata de un estudio observacional en función del objetivo propuesto, del tipo de muestras y del hecho que no se ha intervenido en la selección de individuos. Este hecho condiciona que no sea necesario un tamaño de muestra mínimo.

No se aplicarán en este estudio criterios de inclusión o exclusión. No se excluirá ni se incluirá a ningún individuo por ninguna característica. Los datos que se recogen corresponden a los resultados de análisis de individuos en un laboratorio, sin ningún tipo de restricción ni condición.

- Estudio del promedio de los valores de todas las variables explicativas, frente a los valores de referencia normales de dichas variables así como el cálculo del error estimado de la media.
- Estudio estadístico cruzando los valores de las variables explicativas frente a la estimación del filtrado glomerular de cada paciente.

4. MATERIAL Y MÉTODOS

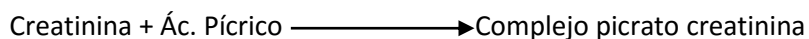
En este apartado, se describen las técnicas realizadas en todas las fases del proceso, desde la obtención de la muestra sanguínea, pasando por la determinación de las distintas pruebas, el uso de la ecuación de estimación del Filtrado Glomerular, MDRD-4, así como la elección del método estadístico utilizado para analizar el problema.

4.1.- MÉTODO DE EXTRACCIÓN SANGUINEA

La muestra de sangre se obtiene por venopunción en la zona de flexión del antebrazo, normalmente en la vena mediana cubital o en la mediana cefálica (*Gonzalez de Buitrago JM, 2.004; Harbert KR, 2013*). Una vez realizada la extracción, se llena el tubo correspondiente, con ácido etilendiaminotetraacético (EDTA) para la determinación de hemoglobina glicosilada (HbA1c), un tubo con citrato sódico para la determinación de la velocidad de sedimentación glomerular y un tubo de separación de suero para la determinación del resto de variables explicativas.

4.2.- DETERMINACIÓN DE CREATININA

La determinación de creatinina, se hace en el auto analizador OLYMPUS AU 2700 (imagen 1), mediante una prueba colorimétrica cinética (método de Jaffé) (*Jaffé M, 1886*). Se basa en que en un medio alcalino, la creatinina forma con el ácido pícrico un compuesto de coloración amarillo-anaranjada. La velocidad del cambio de la absorbancia a 520/800 nm es proporcional a la concentración de creatinina en la muestra. Se puede esquematizar así:



Se utiliza el calibrador OSR 6578, cuyo valor de creatinina es atribuible al material de referencia normalizado (SRM) 909b.



Imagen 1: Auto analizador Olympus AU 2700

4.3.- DETERMINACION DE HEMOGLOBINA GLICOSILADA (HbA_{1c})

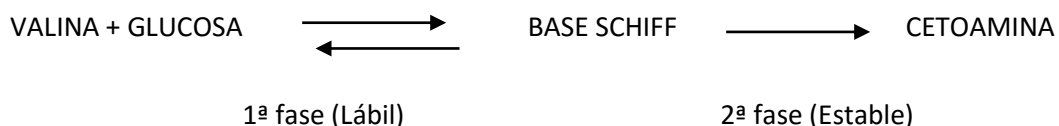
La determinación se hace en el auto analizador ADAMS A1c-HA-8160 (imagen 2), mediante la técnica de HPLC (Sacks DB et al, 2011; Lorenzo-Medina M et al, 2013) .

El fenómeno de la glicosilación no enzimática se produce al entrar en contacto una proteína con una pentosa o una hexosa, creando primero un enlace lábil y luego uno estable. El producto resultante experimenta mutaciones en las características físico-químicas y se dice entonces que la proteína está glicada. La intensidad de la glicosilación de las hemoglobinas depende únicamente del tiempo de exposición de la molécula de la hemoglobina a la glucosa.

La brevedad de la vida eritrocitaria (120 días máximo) permite solamente realizar las dos primeras etapas de la glicosilación:

- La primera, rápida y aún reversible, produce una base de schiff con formación del enlace aldímínico y que está influenciada por la concentración de glucosa en el medio.
- La segunda, más lenta, consistente en la formación de una fructo-amina, mediante un enlace irreversible ceto-amínico, y que refleja la medida de la glucosa durante la vida del hematíe.

ESQUEMA DE GLICOSILACIÓN



Cuando el hemolizado de hematíes se somete a una cromatografía de intercambio catiónico en fase reversa, se evidencian cuatro fracciones de hemoglobina, separándose según las distintas cargas eléctricas: HbF, HbA₁, HbA_{1c} y HbA₂.



Imagen 2: Auto analizador Adams A1c-HA-8160

4.4.- DETERMINACION DE LEUCOCITOS

La determinación de leucocitos se hace en el auto analizador BC-5380 (imagen 3), mediante la técnica de citometría de flujo. Se basa en medir las características de dispersión de luz y fluorescencia que poseen las células conforme se las hace pasar a través de un rayo de luz.

Para su análisis por citometría de flujo, las células deben encontrarse de forma individual en suspensión en un fluido. El recuento y clasificación de las células sanguíneas se basa en sus características morfológicas (Morris M et al, 2016). Las células en suspensión atraviesan un finísimo tubo transparente sobre el que incide un fino rayo de luz láser, la luz transmitida y dispersada por el paso de las células a través del tubo se recoge por medio de unos dispositivos de detección, permitiendo hacer inferencias en cuanto a tamaño y complejidad de las células.



Imagen 3: Auto analizador BC-5380

Los reactivos que se utilizan para la determinación de leucocitos son:

Envolvente:

Proporciona un entorno estable para el recuento y la medición de células sanguíneas.

Lisantes

- El lisante M-53LEO (I): Sirve para dividir las paredes de los glóbulos rojos y coopera con el lisante M-53LH (II) para distinguir entre cuatro tipos de células blancas sanguíneas y tiñe los eosinófilos.
- El lisante M-53LH.

Limpiadores

- El limpiador M-53: Se trata de una disolución de limpieza creada para cebar y limpiar las líneas fluídicas y los tubos y para eliminar la albúmina en sangre y los sedimentos. Se utiliza cada vez que se apaga el equipo para dejar rellenos todos los canales. Esto es importante ya que si no se hiciese esto, habría que lavar todas las cámaras, lo que implicaría una gran pérdida de tiempo.
- El limpiador de sondas M-53P: Se utiliza para limpiar el analizador con regularidad, normalmente una vez por semana.

Se utilizan también controles, que son productos de sangre completa preparados comercialmente, que se utilizan para comprobar que el analizador funciona correctamente. Están disponibles en los niveles bajo, normal y alto, pero en el laboratorio solo se utiliza el nivel normal. El uso diario del control comprueba el buen funcionamiento del analizador y asegura la obtención de resultados fiables.

4.5.- DETERMINACION DE VELOCIDAD DE SEDIMENTACIÓN GLOMERULAR (VSG)

La determinación de Velocidad de Sedimentación Glomerular (VSG), se hace en el analizador Therma NE (imagen 4), consiste en determinar la velocidad con que se sedimentan los hematíes al colocar la sangre en una columna durante una unidad de tiempo (1 hora).

Los hematíes poseen una carga electrostática negativa en su membrana que provoca fenómenos de repulsión y los mantiene en suspensión sin aglomerarse. En condiciones normales, si se coloca la sangre en un tubo en posición vertical, se produce la sedimentación de las células en el fondo debido a la acción de la gravedad. Si la carga electrostática disminuye por alguna causa, los hematíes tienden a aglomerarse aumentando la V.S.G. Ésta se expresa como los milímetros que han descendido la células durante un periodo de tiempo (1 hora).

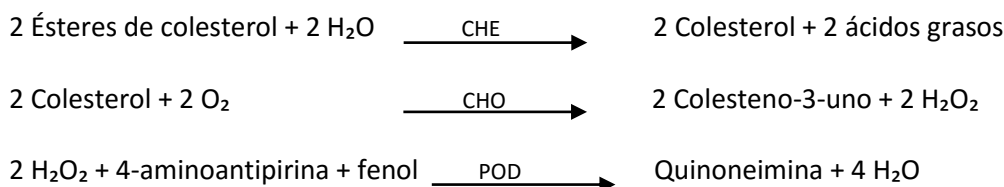
El control se realiza por intercomparación de los resultados obtenidos en el analizador Therma del laboratorio y un analizador Therma del mismo modelo y fabricante, en un laboratorio externo, y mediante un método manual de referencia, que es el método de Westergreen. Este control se realiza con una periodicidad de dos muestras semanales, una de rango alto y otra de rango bajo.



Imagen 4: Auto analizador Therma NE

4.6.- DETERMINACION DE COLESTEROL

La determinación de colesterol, se hace en el auto analizador OLYMPUS AU 2700 (imagen 5), mediante una prueba colorimétrica cinética para cuantificar el colesterol en el plasma y suero humanos. En dicho procedimiento, los ésteres de colesterol de una muestra se hidrolizan con colesterol esterasa (CHE). El colesterol libre resultante se oxida con oxidasa de colesterol (CHO) y se transforma en colesteno-3-uno, produciéndose simultáneamente peróxido de hidrógeno (H₂O₂) que se combina oxidativamente con 4-aminoantipirina y fenol en presencia de peroxidasa (POD), para dar lugar a un cromóforo. El colorante de quinoneimina roja que se forma puede cuantificarse espectrofotométricamente a 540/600 nm como un incremento de la absorbancia. El proceso se puede esquematizar así (*Bachorick PS et al, 2016*):



Como calibrador se utiliza el System Calibrator, N° de Cat. 66300. El valor de colesterol del calibrador es conforme al material de referencia SRM 909b nivel 1 del National Institute of Standards and Technology (NIST) (espectrometría de masa con dilución isotópica).

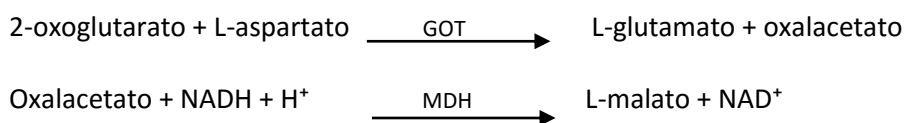
4.7.- DETERMINACION DE TRANSAMINASAS

4.7.1.- DETERMINACION DE GLUTAMATO-OXALACETATO TRANSAMINASA (GOT)

La determinación de GOT, también llamada aspartato aminotransferasa (AST) se hace en el auto analizador OLYMPUS AU 2700 (imagen 5), mediante una prueba UV cinética, basado en las recomendaciones de la Federación Internacional de Química Clínica (IFCC) (*Schumann G et al, 2002*) para cuantificar la aspartato-aminotransferasa del suero y plasma humanos. La GOT se encuentra en una amplia diversidad de tejidos (hígado, miocardio, músculo esquelético, cerebro, riñones, pulmones, páncreas, eritrocitos y leucocitos, etc.), siendo máxima su actividad en el hígado y en el músculo esquelético. La cuantificación de la AST permite diagnosticar, diferenciar y controlar la enfermedad hepatoiliar, el infarto de miocardio y el deterioro del músculo esquelético.

El método de detección se basa en que la GOT cataliza la transaminación del aspartato y del 2-oxoglutarato, formando L-glutamato y oxalacetato.

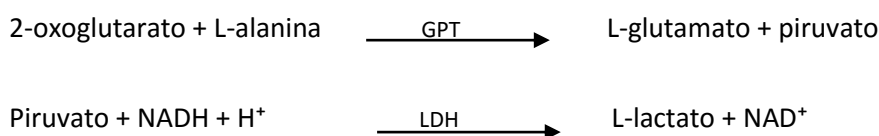
La adición de fosfato de piridoxal a la mezcla reactiva garantiza la máxima actividad catalítica de la AST. La malato-deshidrogenasa (MDH) reduce el oxalacetato a L-malato, en tanto que el NADH se transforma en NAD⁺. La pérdida de absorbencia atribuible al consumo del NADH, cuantificada en 340 nm, es proporcional a la actividad de la AST de la muestra. El piruvato endógeno se elimina, mediante la reacción de la lactato deshidrogenasa (LDH), durante el período de incubación. El proceso se puede esquematizar como:



Como calibrador se utiliza el System Calibrator, N° de Cat. 66300, en el modo de calibración AB.

4.7.2.- DETERMINACION DE GLUTAMATO-PIRUVATO TRANSAMINASA (GPT)

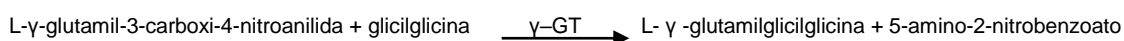
La determinación de GPT, también llamada alanina aminotransferasa (ALT), se hace en el auto analizador OLYMPUS AU 2700 (imagen 5), mediante una prueba UV cinética, basado en las recomendaciones de la Federación Internacional de Química Clínica (IFCC) (*Schumann G et al, 2002*) para cuantificar la alanina-aminotransferasa del suero y plasma humanos. La GPT es una aminotransferasa, que corresponde a un grupo de enzimas que cataliza la transformación reversible de los cetoácidos en aminoácidos mediante transferencia de grupos amino. La GPT transfiere el grupo amino desde la alanina al 2-oxoglutarato, para formar piruvato y glutamato. La adición de fosfato de piridoxal a la mezcla reactiva garantiza la máxima actividad catalítica de la GPT. La reacción del piruvato con el NADH, catalizada por la lactato deshidrogenasa (LDH), produce lactato y NAD⁺. La pérdida de absorbencia atribuible al consumo del NADH, cuantificada en 340 nm, es proporcional a la actividad de la GPT de la muestra. El piruvato endógeno se elimina durante el período de incubación. El proceso se puede esquematizar como:



4.8.- DETERMINACION DE GAMMA GLUTAMIL TRANSPEPTIDASA (GGT)

La determinación de GGT, se hace en el auto analizador OLYMPUS AU 2700 (imagen 5), mediante una prueba colorimétrica cinética para cuantificar la γ -glutamyl-transferasa en el plasma y suero humanos, basada en las recomendaciones de la Federación Internacional de Química Clínica (IFCC) (Schumann G et al, 2002). La γ -glutamyl-transpeptidasa (GGT) pertenece a un grupo de peptidasas que catalizan la transferencia de aminoácidos entre dos péptidos y, en consecuencia, actúan a modo de aminoácido transferasas. Esta enzima sólo reacciona ante péptidos o compuestos similares que contengan un residuo de glutamato terminal unido al resto del compuesto a través del carboxilo terminal.

La GGT cataliza la transferencia del grupo γ -glutámico, entre el sustrato (γ -glutamyl-3-carboxilo-4-nitroanilida) y la glicilglicina, produciendo 5-amino-2-nitrobenzoato. El cambio de absorbencia a 410/480 nm, debido a la formación de 5-amino-2-nitrobenzoato, es directamente proporcional a la actividad de la GGT de la muestra. Esquema del proceso:



Como calibrador se utiliza el System Calibrator, Nº de Cat. 66300, en el modo de calibración AB.



Imagen 5: Auto analizador Olympus AU 2700

4.8.- ELECCIÓN DEL TIPO DE ANÁLISIS ESTADÍSTICO.

Los test que se utilizan para el estudio de nuestro problema son los test de Regresión logística (*Bagley SC et al, 2001, Milton JS, 2001*), Probit (*Bliss C, 1934, McCullag P et al, 1989*) y LOG-LOG (*Paton D et al, 2002; Pourhoseingholi A et al, 2010*), porque se ajustan más a nuestro problema ya que los modelos de regresión logística tratan de explicar el resultado de una variable binaria Y en función de un grupo de variables explicativas. Estos modelos se usan muy frecuentemente en estudios médicos en los que se busca explicar la probabilidad de enfermar a partir de unas variables explicativas. La diferencia con un modelo de regresión múltiple, es que en éste la variable que se quiere explicar se supone que sigue una distribución normal, mientras que en el modelo de regresión logística, la probabilidad de que ocurra un suceso sigue una distribución binomial, donde la probabilidad de éxito depende de los valores de las variables explicativas.

Una vez hallado el modelo a utilizar, es preciso determinar hasta qué punto las variables incluidas son necesarias, para ello se usan varios contrastes: AIC, SC y el test de la razón de verosimilitudes. Lo siguiente es ver si el modelo halla una constante o estima parámetros, para ello se utilizan tres contrastes: Test de Wald, Scores de Rao y test de cociente de verosimilitudes. En el caso de la regresión logística, se estudian las “odds”, para cada una de las variables explicativas, es decir, ver si para cada una de ellas el valor de probabilidad de éxito aumenta, al aumentar una unidad el valor de la correspondiente variable explicativa. Si el valor de la “odd” de una variable explicativa es mayor de 1, significa que al aumentar el valor de dicha variable en una unidad, aumenta la probabilidad de que el experimento lleve el suceso al éxito. Por último, hay que ver en estos modelos si el ajuste es suficientemente bueno, para ello existen varios estadísticos, pero en nuestro caso nos interesa el de Hosmer y Lemeshow, ya que tiene observaciones tipo Bernoulli en vez de binomiales.

Entre los test potencialmente aplicables a nuestro problema que fueron inicialmente descartados, cabe citar:

Análisis por Componentes Principales: ninguna variable demostró especial relevancia, para explicar la variación, en comparación con el resto de variables. Los coeficientes de todas las variables son similares.

Análisis Multivariante (Clúster): con este análisis se buscaba obtener grupos de individuos muy homogéneos, para estudiar las relaciones entre las distintas variables. Pero se

vio que en unos casos salían grupos con muy pocos individuos, e incluso alguno con uno solo, y en otros casos grupos muy numerosos.

Análisis de Varianza: el primer impedimento fue tener un número de niveles reducido para cada variable, por lo que el problema era muy complicado a nivel computacional. No se utilizó porque no se cumplían las condiciones del análisis de varianza, como la normalidad de todas las subpoblaciones y la homocedasticidad (igualdad de varianzas).

4.9.- AGRUPACION DE LAS MUESTRAS PARA EL ESTUDIO ESTADISTICO

Se realizó el estudio agrupando todas las muestras de dos maneras, primeramente se hace el estudio de toda la población en conjunto, y posteriormente separándolo por sexos, y además debido a que la NKDEP no recomienda utilizar el filtrado glomerular (FGR) en pacientes menores de 18 años o mayores de 70 años, se realiza el mismo estudio usando solamente las muestras que cumplen ese requisito, primero de toda la población en ese rango de edad y posteriormente separándolo por sexos.

Los valores de las variables explicativas se obtienen de la BBDD del laboratorio, de los análisis realizados en multitud de pacientes a lo largo de varios años. En cuanto al valor de FGR se obtiene como ya se ha explicado anteriormente, mediante la ecuación MDRD-4, pues la valoración del Filtrado Glomerular a través de estas ecuaciones es más exacta y precisa. Como esta ecuación ya homogeniza el valor de FGR para hombres y mujeres, en el estudio se ha considerado de acuerdo con las guías publicadas por la NKDEP, que había indicios de enfermedad renal si el valor de FGR era inferior a 60, para ambas subpoblaciones.

Se utiliza para la resolución del problema la versión 9.4 del programa estadístico SAS.

Para la selección del modelo hay que tener en cuenta que en ocasiones se tiene un número muy elevado de variables explicativas, y es posible que se pueda conseguir un mejor ajuste con un número menor de variables. En SAS hay cuatro procedimientos para determinar el modelo más adecuado; en nuestro caso, se ha decidido utilizar el procedimiento de selección backward (eliminación hacia atrás); este procedimiento consiste en considerar inicialmente todas las variables e ir eliminando en cada paso la variable que aporte menos al modelo; para seleccionar esta variable se realiza un contraste sobre cada uno de los parámetros y se elige la variable a abandonar el modelo entre las variables cuyo parámetro tenga un p-valor menor que un valor prefijado (normalmente 0.05). Una vez que se ha

determinado la variable que debe salir, se vuelve a ajustar el modelo con las variables que quedan y así sucesivamente hasta que ninguna de las variables del modelo tenga un coeficiente significativo. Es una forma de actuar que sigue la misma filosofía que el modelo backwards de regresión múltiple.

Se decidió no recodificar los valores de las variables explicativas; en consecuencia, cada individuo tiene una combinación de valores de las variables explicativas que no se repetían en otros individuos. Por ello, se tiene una sola observación por cada combinación de valores de las variables explicativas y nuestra distribución en cada combinación pasa de ser una distribución binomial a una distribución de tipo Bernoulli. Esto influye en el test de bondad de ajuste que se selecciona, se elige el de Hosmer y Lemeshow, tal y como se explicó anteriormente.

5.- RESULTADOS Y DISCUSIÓN

5.1.- VALORACIÓN PROMEDIO DE LAS VARIABLES EXPLICATIVAS.

Antes de realizar el estudio estadístico propiamente dicho, en el que valoramos si las variables explicativas tenían correlación con la estimación del daño renal a través del FGR, hemos querido valorar si el promedio de dichas variables se ajusta a los valores de referencia utilizados por el laboratorio. Por un lado, lo realizamos de toda la población que intervino en el estudio y posteriormente separada por sexo, debido a que ciertos valores de referencia varían en función del mismo. A continuación realizamos el mismo estudio teniendo en cuenta solamente la población con edades comprendidas entre 18 y 70 años, pues como veremos posteriormente en el estudio estadístico, la NKDEP, advierte que el FGR no debe ser utilizado como criterio de daño renal en pacientes menores de 18 años o mayores de 70. Del mismo modo lo hicimos primero con toda la población comprendida en ese rango de edad, y posteriormente separada por sexo.

El uso de valores de referencia es hoy en día algo imperativo en los laboratorios clínicos, pues es la manera de comparar si los resultados obtenidos en los pacientes se encuentran dentro de la normalidad, o por el contrario están alterados. El concepto de valor de referencia fue diseñado en la década de 1970 por un grupo escandinavo, para a continuación ser desarrollado y completado por varias sociedades científicas a nivel mundial, especialmente al Federación Internacional de Química Clínica y Medicina del Laboratorio (IFCC-LM), en los años 1980. Más recientemente un grupo de trabajo conjunto del Clinical and Laboratory Standards Institute (CLSI) y la IFCC-LM, han revisado todos los documentos publicados con anterioridad, y publicado un documento llamado "Definir, establecer y verificar intervalos de referencia en el laboratorio clínico: Approved Guideline-Third Edition - C28-A3 11-2010" (*Henny J, 2011*).

El intervalo de referencia se corresponde con el intervalo especificado de la distribución de los valores obtenidos a partir de poblaciones de sujetos sanos. Por lo general se establece en un intervalo correspondiente al 95% de la población, centrado en la mediana. Puede variar dependiendo de la muestra estándar primaria seleccionada y el método utilizado. En algunos casos, se puede seleccionar un solo límite de referencia, por lo general el límite superior. La determinación de la referencia de intervalo se basa en cálculos estadísticos. Es puramente descriptiva de una población dada (*Henny J, 2011*).

Tabla 2: Promedio de Leucocitos frente a valores de referencia y error estándar.

LEUCOCITOS	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	6.49		1.93	0.02
HOMBRES	6.71	4-10.6	1.86	0.03
MUJERES	6.27	3.5-11	1.97	0.03
TOTAL 18-70 años	6.48		1.89	0.03
HOMBRES 18-70 años	6.74	4-10.6	1.92	0.03
MUJERES 18-70 años	6.23	3.5-11	1.82	0.03

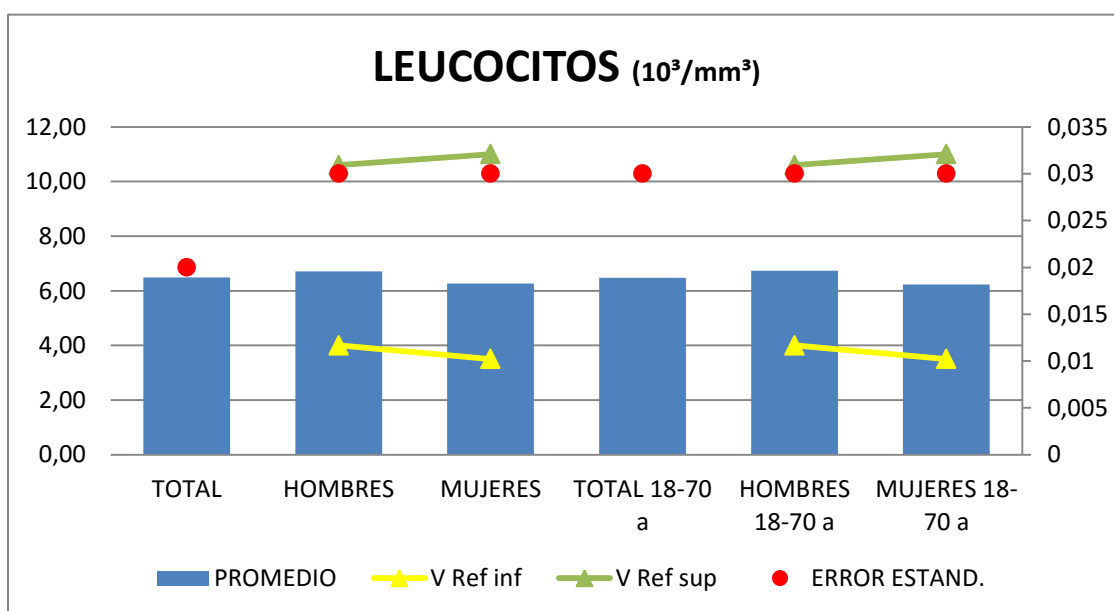


Figura 6: Promedio de Leucocitos frente a valores de referencia y error estándar.

Tabla 3: Promedio de VSG frente a valores de referencia y error estándar.

VSG	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	11		9	0.10
HOMBRES	8	<20	9	0.13
MUJERES	13	<25	9	0.13
TOTAL 18-70 años	10		8	0.13
HOMBRES 18-70 años	7	<20	8	0.13
MUJERES 18-70 años	12	<25	8	0.13

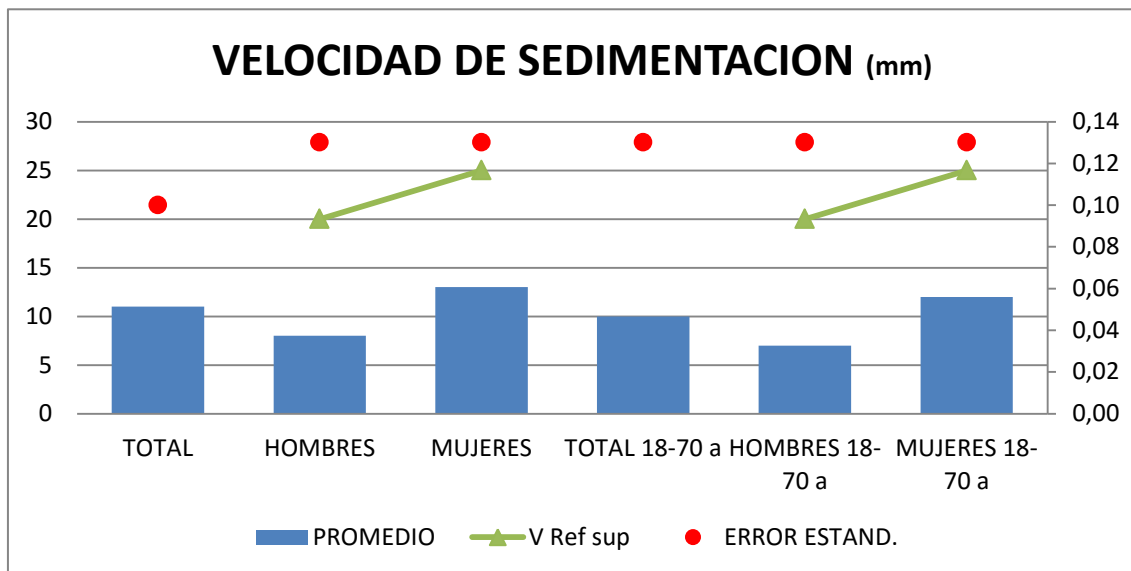


Figura 7: Promedio de VSG frente a valores de referencia y error estándar.

Tabla 4: Promedio de Colesterol frente a valores de referencia y error estándar.

COLESTEROL	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	193	<200	38	0.38
HOMBRES	187	<200	38	0.55
MUJERES	200	<200	36	0.52
TOTAL 18-70 años	196	<200	38	0.61
HOMBRES 18-70 años	190	<200	39	0.64
MUJERES 18-70 años	202	<200	36	0.59

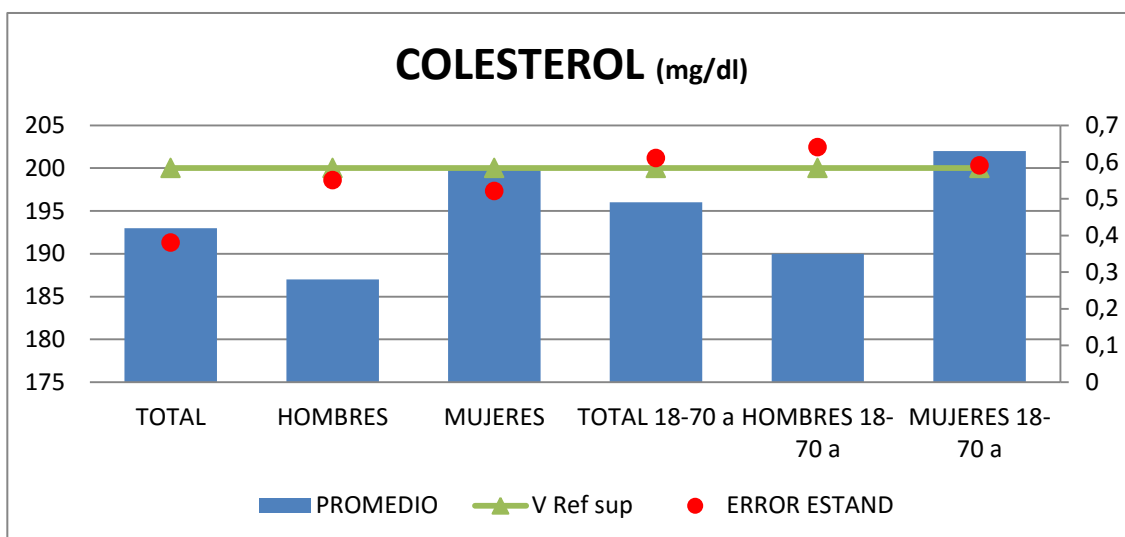


Figura 8: Promedio de Colesterol frente a valores de referencia y error estándar.

Tabla 5: Promedio de Creatinina frente a valores de referencia y error estándar.

CREATININA	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	0.87	0.5-1.2	0.25	0.003
HOMBRES	0.99	0.5-1.2	0.26	0.004
MUJERES	0.74	0.5-1.2	0.17	0.002
TOTAL 18-70 años	0.85	0.5-1.2	0.22	0.003
HOMBRES 18-70 años	0.98	0.5-1.2	0.21	0.003
MUJERES 18-70 años	0.72	0.5-1.2	0.14	0.002

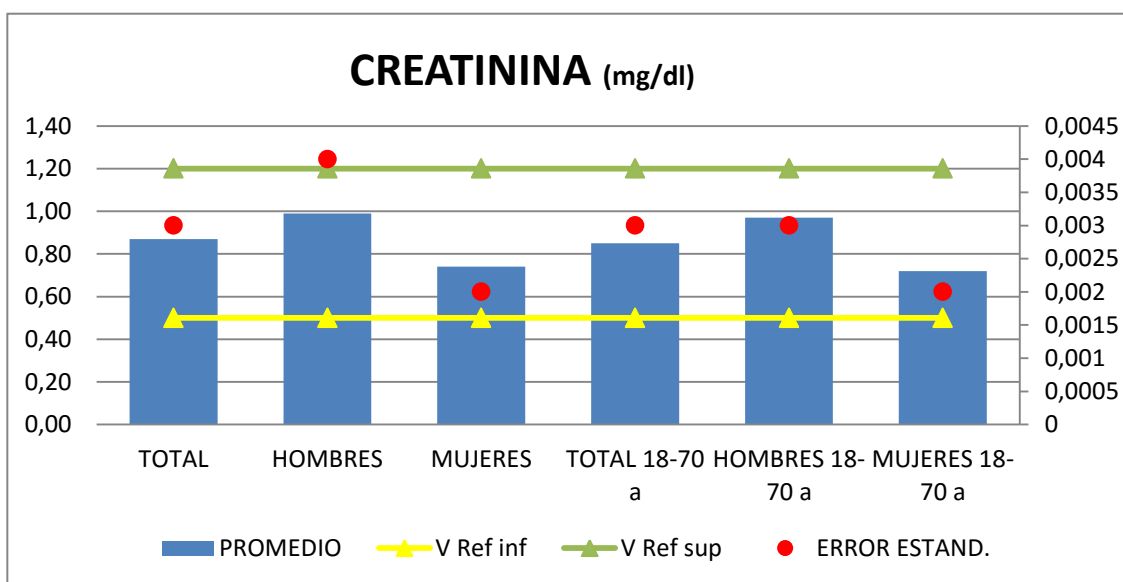


Figura 9: Promedio de Creatinina frente a valores de referencia y error estándar.

Tabla 6: Promedio de GOT frente a valores de referencia y error estándar.

GOT	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	25	<40	13	0.13
HOMBRES	26	<40	13	0.19
MUJERES	23	<40	12	0.17
TOTAL 18-70 años	25	<40	13	0.21
HOMBRES 18-70 años	26	<40	13	0.21
MUJERES 18-70 años	23	<40	13	0.21

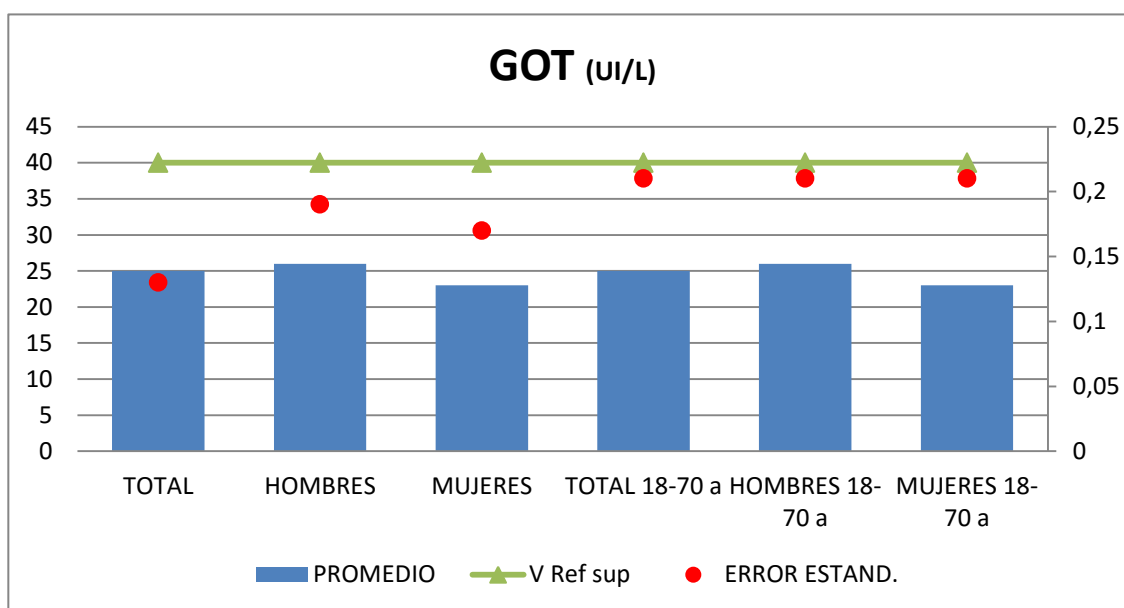


Figura 10: Promedio de GOT frente a valores de referencia y error estándar.

Tabla 7: Promedio de GPT frente a valores de referencia y error estándar.

GPT	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	25	<35	19	0.19
HOMBRES	28	<35	20	0.29
MUJERES	22	<35	17	0.25
TOTAL 18-70 años	26	<35	19	0.31
HOMBRES 18-70 años	30	<35	20	0.33
MUJERES 18-70 años	23	<35	18	0.29

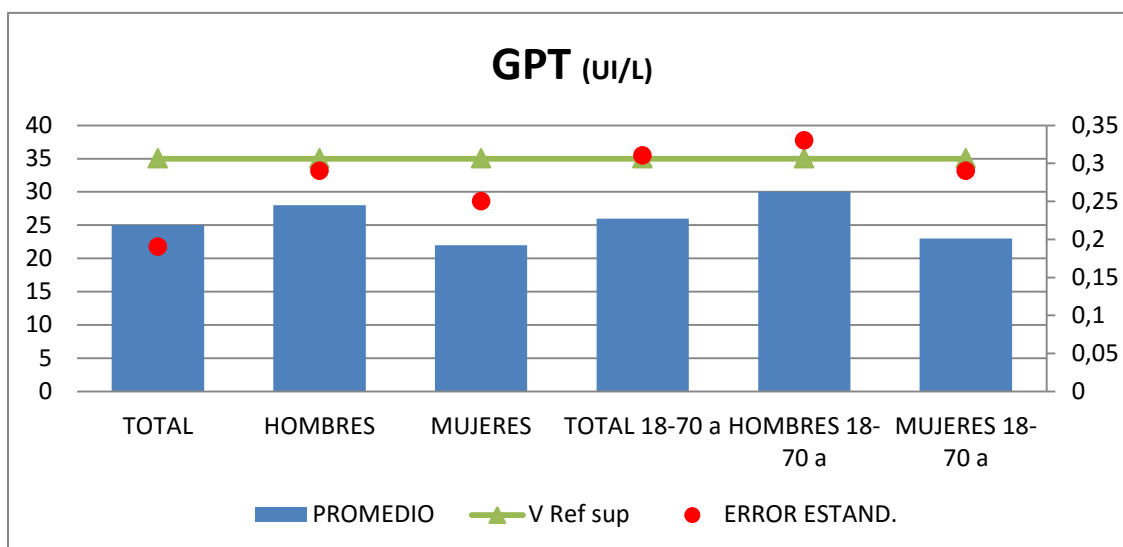


Figura 11: Promedio de GPT frente a valores de referencia y error estándar.

Tabla 8: Promedio de GGT frente a valores de referencia y error estándar.

GGT	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	35		49	0.49
HOMBRES	43	<50	57	0.82
MUJERES	28	<40	38	0.54
TOTAL 18-70 años	35		45	0.72
HOMBRES 18-70 años	43	<50	52	0.86
MUJERES 18-70 años	28	<40	35	0.55

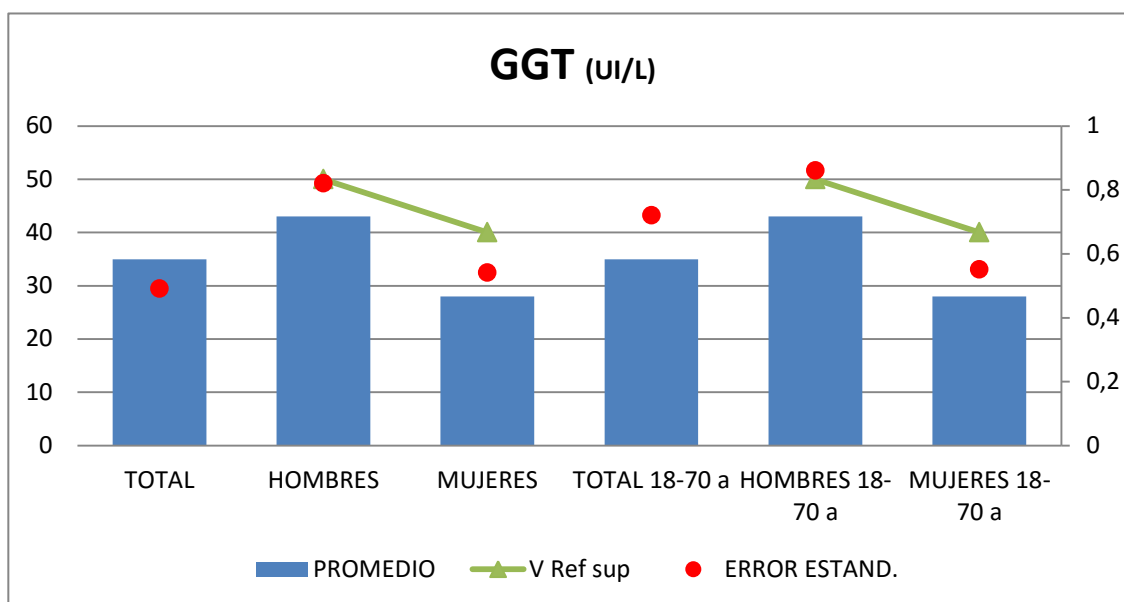


Figura 12: Promedio de GGT frente a valores de referencia y error estándar.

Tabla 9: Promedio de HBA1C frente a valores de referencia y error estándar.

HBA1C	PROMEDIO	VALORES DE REFERENCIA	DESVIACION ESTANDAR	ERROR ESTANDAR
TOTAL	5.4	4-6	1.2	0.01
HOMBRES	5.6	4-6	1.4	0.02
MUJERES	5.2	4-6	1.0	0.01
TOTAL 18-70 años	5.3	4-6	1.3	0.02
HOMBRES 18-70 años	5.5	4-6	1.5	0.02
MUJERES 18-70 años	5.1	4-6	0.9	0.01

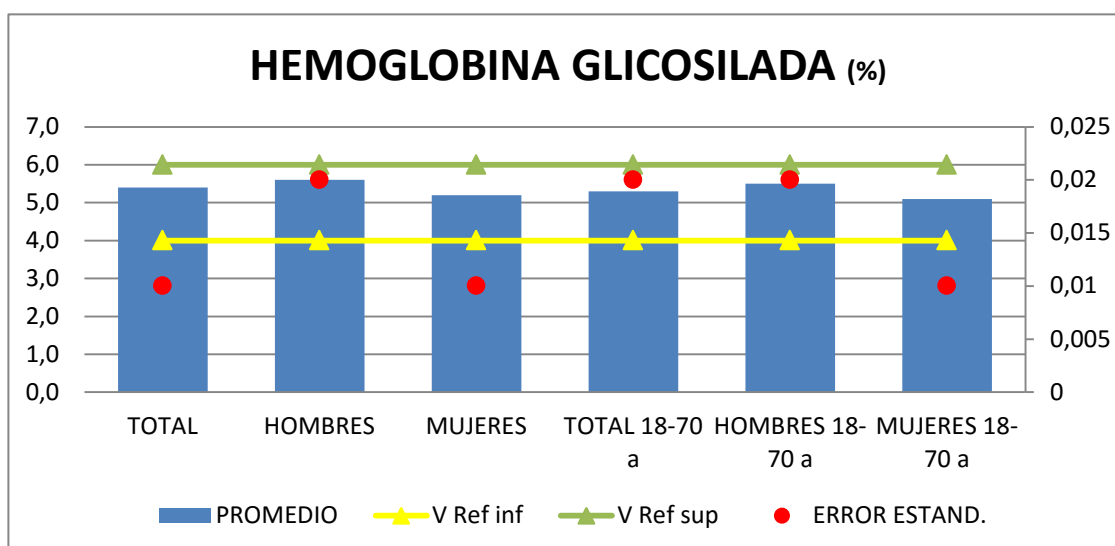


Figura 13: Promedio de HBA1c frente a valores de referencia y error estándar.

Las tablas 2 a la 9 y las figuras 6 a la 13 muestran los resultados obtenidos, y se puede observar que todos los promedios estudiados, excepto el colesterol total en mujeres y la gamma glutamil transpeptidasa en hombres, que exceden ligeramente el límite de los valores de referencia, se encuentran perfectamente delimitados por dichos valores, siendo además el error estándar muy poco significativo. Los promedios están comparados frente a los valores de referencia del laboratorio, y nos muestra que los individuos sobre los que se ha realizado el estudio estadístico no tienen ningún valor especialmente significativo frente a la población normal.

5.2.- ESTUDIO ESTADISTICO

El estudio se realizó con una muestra de población de 9.750 personas, inicialmente incluimos todas las muestras sin ningún tipo de discriminación en función de la edad, y analizamos las muestras sin ningún tipo de desagregación por sexo, utilizando tres modelos, logit, probit y log-log, a continuación desagregamos por sexo y realizamos el estudio de los tres mismos modelos, primero solo a las muestras de hombres y luego solo las muestras de mujeres. Posteriormente seleccionamos de inicio solamente aquellas muestras que cumplan el criterio de edad de estar entre los 18 y los 70 años, debido a las recomendaciones del National Kidney Disease Education Program (NKDEP) de Estados Unidos advierte que el valor de FGR no debe ser usada en edades menores a 18 ni mayores a 70 años. (Myers GL et al, 2006), y volvimos a realizar el estudio de los tres modelos en la población conjunta y a continuación con las muestras desagregadas, por un lado los hombres y por otro las mujeres. A continuación se muestran los resultados obtenidos:

POBLACION CONJUNTA CON EDADES DE 2-99 AÑOS

Modelo logit

En el caso de la población conjunta, sin tener en cuenta las diferencias entre hombres y mujeres, tenemos 9.750 datos. El número de individuos con valor FGR inferior a 60 (valor 1) y con valor FGR superior a 60 (valor 0) vienen dados por:

Tabla 10: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	9750
Number of Observations Used	9750

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	695
2	sano(>=60)	9055

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo. La tabla 11 muestra las estimaciones de los parámetros del modelo.

Tabla 11: Análisis de ajuste de los datos al modelo

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.5402	0.3104	130.1152	<.0001
Leucocitos	1	0.0677	0.0174	15.1213	0.0001
V1HORA	1	0.0447	0.00319	196.1938	<.0001
COL	1	-0.00481	0.00109	19.4276	<.0001
GOT	1	0.0321	0.00535	36.0247	<.0001
GGT	1	0.00187	0.000695	7.2566	0.0071
GPT	1	-0.0291	0.00453	41.3306	<.0001
HBA1C	1	0.1275	0.0306	17.3814	<.0001

En la segunda columna de resultados de esta tabla se tienen los valores de los estimadores de los parámetros beta del modelo. Entonces, la probabilidad de que la variable FGR tome un valor significativo (por debajo de 60) viene dado por:

$$\pi(x) = \frac{\exp(-3.5402 + 0.0677Leucocitos + \dots + 0.1275HBA1C)}{1 + \exp(-3.5402 + 0.0677Leucocitos + \dots + 0.1275HBA1C)} \quad (ec. 3)$$

La tercera columna de resultados da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. Se puede observar que todas las estimaciones son distintas de cero.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 12: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	5012.382	4650.176
SC	5019.567	4707.656
-2 Log L	5010.382	4634.176

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 13: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	376.2059	7	<.0001
Score	524.9740	7	<.0001
Wald	354.6427	7	<.0001

Note: No (additional) effects met the 0.05 significance level for removal from the model.

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

La Tabla 14 recoge las “odds” de las distintas variables explicativas.

Tabla 14: Odds de las variables explicativas

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Leucocitos	1.070	1.034	1.107
V1HORA	1.046	1.039	1.052
COL	0.995	0.993	0.997
GOT	1.033	1.022	1.044
GGT	1.002	1.001	1.003
GPT	0.971	0.963	0.980
HBA1C	1.136	1.070	1.206

En este caso se observa que las variables Leucocitos, V1Hora, GOT y HBA1C hacen que aumente la probabilidad de llegar a desarrollar una enfermedad renal por disminución del FGR por debajo de 60, mientras que las otras variables hacen que la probabilidad disminuya.

Finalmente, falta por comprobar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow (Paul P et al, 2013), cuyos resultados se recogen en la tabla 15:

Tabla 15: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	975	23	25.88	952	949.12
2	975	32	36.04	943	938.96
3	975	28	41.60	947	933.40
4	976	35	46.94	941	929.06
5	975	46	52.10	929	922.90
6	975	48	57.99	927	917.01
7	975	69	65.28	906	909.72
8	975	74	75.59	901	899.41
9	975	110	94.75	865	880.25
10	974	230	198.83	744	775.17

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
20.3453	8	0.0091

A partir de esta tabla, se ve que el modelo no es adecuado, por lo que se descarta.

Modelo probit

Vamos a considerar ahora el modelo probit. Como en el caso anterior tenemos 9750 datos y el número de éxitos y fracasos es el mismo que en el caso anterior.

Tabla 16: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	9750
Number of Observations Used	9750

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	695
2	sano(>=60)	9055

Probability modeled is FGR_='enfermo(<60)'.

A continuación se procede a estimar los parámetros del modelo. Estas estimaciones se recogen en la siguiente tabla 17:

Tabla 17: Análisis de ajuste de los datos al modelo

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.0264	0.1491	184.6935	<.0001
Leucocitos	1	0.0372	0.00906	16.8681	<.0001
V1HORA	1	0.0243	0.00169	207.3134	<.0001
COL	1	-0.00247	0.000534	21.3241	<.0001
GOT	1	0.0170	0.00260	42.8296	<.0001
GGT	1	0.00111	0.000373	8.8843	0.0029
GPT	1	-0.0147	0.00210	48.9586	<.0001
HBA1C	1	0.0674	0.0136	24.4606	<.0001

La segunda columna de resultados de esta tabla muestra los valores de los estimadores de los parámetros beta del modelo. Nótese que no se han obtenido los mismos valores que en el modelo logit, lo que era de esperar ya que se trata de modelos diferentes.

Entonces, la probabilidad de que la variable FGR tome un valor significativo (por debajo de 60) viene dado por la siguiente expresión:

$$\pi(x) = \Phi(-2.0264 + 0.0372Leucocitos + \dots + 0.0674HBA1C) \quad (ec. 4)$$

La tercera columna de resultados muestra las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. Con estos resultados se ve que ninguna de las estimaciones puede considerarse distinta de cero.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 18: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	5012.382	4631.327
SC	5019.567	4688.808
-2 Log L	5010.382	4615.327

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla 19 muestra los valores obtenidos:

Tabla 19: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	395.0544	7	<.0001
Score	406.7813	7	<.0001
Wald	386.1132	7	<.0001

Note: No (additional) effects met the 0.05 significance level for removal from the model.

Como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, falta por comprobar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 20: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	975	21	22.57	954	952.43
2	975	32	33.06	943	941.94
3	975	26	39.11	949	935.89
4	975	35	44.97	940	930.03
5	976	50	50.86	926	925.14
6	976	48	57.38	928	918.62
7	975	67	65.39	908	909.61
8	975	77	76.80	898	898.20
9	975	108	98.09	867	876.91
10	973	231	201.35	742	771.65

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
15.3484	8	0.0527

A partir de los datos obtenidos y recogidos en estas tablas, se ve que el modelo sí podría ajustar bien los datos, aunque estamos en un p-valor muy cercano al valor límite 0.05. Veremos más adelante que es posible conseguir un p-valor mucho más alto en otras condiciones; quede claro sin embargo que sí se podría admitir que el modelo probit es capaz de explicar los datos de forma bastante satisfactoria.

Modelo log-log

Pasamos finalmente a estudiar el modelo log-log. La tabla 21, que nos da información sobre los datos, es común con la de los otros modelos.

Tabla 21: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	9750
Number of Observations Used	9750

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	695
2	sano(>=60)	9055

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo; se obtiene la siguiente tabla de estimaciones de los parámetros del modelo.

Tabla 22: Análisis de ajuste de los datos al modelo

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.0393	0.2440	155.1994	<.0001
Leucocitos	1	0.0520	0.0119	19.0388	<.0001
V1HORA	1	0.0341	0.00204	278.8349	<.0001
COL	1	-0.00442	0.00102	18.6481	<.0001
GOT	1	0.0225	0.00322	49.0359	<.0001
GGT	1	0.00170	0.000535	10.0502	0.0015
GPT	1	-0.0229	0.00378	36.7781	<.0001
HBA1C	1	0.0761	0.0136	31.4655	<.0001

En la segunda columna de resultados de esta tabla se reflejan los valores de los estimadores de los parámetros beta del modelo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la siguiente expresión:

$$\log(-\log(\pi(x))) = -3.0393 + 0.0520Leucocitos + \dots + 0.0761HBA1C \quad (ec. 5)$$

La tercera columna de resultados corresponde a las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. A la vista de estos resultados se concluye que todas las estimaciones son distintas de cero.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 23: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	5012.382	4684.707
SC	5019.567	4742.187
-2 Log L	5010.382	4668.707

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla 24 muestra los siguientes valores:

Tabla 24: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	341.6746	7	<.0001
Score	572.1301	7	<.0001
Wald	513.6991	7	<.0001

Note: No (additional) effects met the 0.05 significance level for removal from the model.

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, queda por comprobar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la tabla 25:

Tabla 25: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	975	24	30.68	951	944.32
2	976	36	41.03	940	934.97
3	975	23	46.27	952	928.73
4	975	41	50.91	934	924.09
5	975	43	55.54	932	919.46
6	975	50	60.49	925	914.51
7	975	62	66.48	913	908.52
8	975	73	74.56	902	900.44
9	975	109	88.79	866	886.21
10	974	234	170.66	740	803.34

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
55.3298	8	<.0001

A partir de estas tablas, se ve que el modelo no ajusta bien con los datos, por lo que se descarta también este modelo.

Una vez analizados los datos conjuntos con los tres modelos podemos decir que la Bondad de Ajuste de Hosmer y Lemeshow para los datos agregados (la población total sin separación por sexos) es difícilmente aceptable, aunque quizá el modelo probit estaría al límite de los aceptable con un nivel de significación de 0.05

- $\text{logit}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{COL}, \text{GOT}, \text{GGT}, \text{GPT}, \text{HBA1C}$; $p\text{-valor} = 0.0091$.
- $\text{probit}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{COL}, \text{GOT}, \text{GGT}, \text{GPT}, \text{HBA1C}$; $p\text{-valor} = \mathbf{0.0527}$.
- $\text{cloglog}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{COL}, \text{GOT}, \text{GGT}, \text{GPT}, \text{HBA1C}$; $p\text{-valor} = < 0.0001$.

En el segundo paso, vamos a realizar el mismo estudio por separado para las poblaciones de hombres y mujeres con el objetivo de estudiar diferencias de comportamiento en función del sexo. Nótese que FGR toma valores homogéneos, ya que para ello la fórmula con la que se calcula (MDRD-4), tiene en cuenta esta diferencia, por ello posibles diferencias entre estas dos subpoblaciones serían debidas a diferencias en el modelo.

POBLACION EXCLUSIVAMENTE DE HOMBRES CON EDADES DE 2-99 AÑOS

Modelo logit

En el caso de la población conjunta, la población masculina corresponde a 4.820 datos. El número de individuos con valor FGR inferior a 60 (valor 1) y con valor FGR superior a 60 (valor 0) vienen dados por:

Tabla 26: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	4820
Number of Observations Used	4820

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	398
2	sano(>=60)	4422

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo; se obtiene la tabla 27 que muestra las estimaciones de los parámetros del modelo.

Tabla 27: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	HBA1C	1	6	1.6918	0.1934	HBA1C
2	GGT	1	5	2.0022	0.1571	GGT
3	Leucocitos	1	4	2.7325	0.0983	Leucocitos

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.9313	0.3021	40.8747	<.0001
V1HORA	1	0.0554	0.00474	136.5818	<.0001
COL	1	-0.00509	0.00149	11.6280	0.0006
GOT	1	0.0220	0.00592	13.7432	0.0002
GPT	1	-0.0266	0.00551	23.3961	<.0001

Se observa ya una primera diferencia con el caso de la población conjunta: en este caso se han eliminado varias variables del modelo. En concreto, se han eliminado del modelo las variables HBA1C, GGT y Leucocitos. Por lo tanto, tenemos ahora un modelo con sólo cuatro variables explicativas. En la segunda columna de resultados de esta tabla se recogen los valores de los estimadores de los parámetros beta del modelo, que son los que aparecen en la fórmula que explica la probabilidad de éxito. Entonces, la probabilidad de que se la variable FGR tome un valor significativo viene dado por la siguiente ecuación:

$$\pi(x) = \frac{\exp(-1.9313 + 0.0554V1HORA - 0.00509COL + 0.022GOT - 0.0266GPT)}{1 + \exp(-1.9313 + 0.0554V1HORA - 0.00509COL + 0.022GOT - 0.0266GPT)} \quad (ec. 6)$$

La segunda columna muestra las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald; en la tabla 27 se dan primeramente los valores para las variables que deben ser eliminadas del modelo, y posteriormente se dan los correspondientes valores para las

variables que han sido incluidas en la última iteración backwards. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna; en la primeras variables los p-valores son grandes, por lo que estas no se incluyen en el modelo, mientras que en la siguientes variables los p-valores son pequeños y entonces esas se incluyen en el modelo.

Si ahora se observa la bondad de ajuste del modelo seleccionado se obtiene la siguiente tabla:

Tabla 28: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2749.478	2533.805
SC	2755.958	2566.207
-2 Log L	2747.478	2523.805

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 29: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	223.6731	4	<.0001
Score	317.0849	4	<.0001
Wald	197.9354	4	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Las “odds” de las distintas variables explicativas vienen dadas en la siguiente tabla.

Tabla 30: Odds de las variables explicativas

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
V1HORA	1.057	1.047	1.067
COL	0.995	0.992	0.998
GOT	1.022	1.010	1.034
GPT	0.974	0.963	0.984

En este caso se ve que las variables V1Hora y GOT hacen que aumente la probabilidad de tener la enfermedad, mientras que COL y GPT hacen que la probabilidad disminuya.

Finalmente, nos queda por observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 31: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	482	13	13.53	469	468.47
2	482	19	20.48	463	461.52
3	482	23	24.24	459	457.76
4	482	17	27.37	465	454.63
5	482	33	30.25	449	451.75
6	483	27	33.59	456	449.41
7	482	35	37.62	447	444.38
8	483	39	43.17	444	439.83
9	482	62	52.59	420	429.41
10	480	130	115.16	350	364.84

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
11.0684	8	0.1979

A partir de estos resultados, se ve que el modelo ajusta bien con los datos, bastante mejor que el modelo que se tenía para la población conjunta con el modelo probit.

Modelo probit

Como en el caso anterior, tenemos entonces 4. 820 datos, y la tabla 32 es común con el caso anterior.

Tabla 32: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	4820
Number of Observations Used	4820

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	398
2	sano(>=60)	4422

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo; se obtiene la siguiente tabla de estimaciones de los parámetros del modelo.

Tabla 33: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	Leucocitos	1	6	2.0398	0.1532	Leucocitos
2	GGT	1	5	2.1540	0.1422	GGT
3	HBA1C	1	4	2.9319	0.0868	HBA1C

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1848	0.1504	62.0724	<.0001
V1HORA	1	0.0299	0.00244	149.7888	<.0001
COL	1	-0.00254	0.000741	11.7861	0.0006
GOT	1	0.0121	0.00319	14.3480	0.0002
GPT	1	-0.0135	0.00269	25.1889	<.0001

Nótese las similitudes con el modelo anterior en el sentido de que han sido descartadas las mismas variables. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo óptimo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\Phi(\pi(x)) = -1.1848 + 0.0299V1HORA - 0.00254COL + 0.0121GOT - 0.0135GPT \quad (ec. 7)$$

La segunda columna da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. En las primeras variables aparecen p-valores grandes porque se excluyen del modelo. En las siguientes variables los p-valores son pequeños porque se incluyen en el modelo. Entonces se concluye que todas las estimaciones de la segunda tabla son distintas de cero aunque los valores de los parámetros son cercanos a cero; esto es debido a que los valores que toman las variables explicativas son grandes y entonces el producto toma un valor que es claramente distinto de cero.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 34: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2749.478	2528.144
SC	2755.958	2560.547
-2 Log L	2747.478	2518.144

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global para los parámetros del modelo considerado (es decir, el reducido)

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 35: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	229.3338	4	<.0001
Score	227.7856	4	<.0001
Wald	214.5111	4	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 36: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	482	14	12.20	468	469.80
2	482	21	19.48	461	462.52
3	482	20	23.50	462	458.50
4	482	19	26.85	463	455.15
5	482	32	29.96	450	452.04
6	482	25	33.42	457	448.58
7	482	34	37.80	448	444.20
8	482	41	43.63	441	438.37
9	482	62	53.63	420	428.37
10	482	130	114.12	352	367.88

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
10.7557	8	0.2159

A partir de estas tablas, se ve que el modelo ajusta bastante bien los datos, mejor incluso de lo que pasaba con el modelo logit.

Modelo log-log

Comenzamos como siempre con los datos de la muestra, con lo que se tiene una tabla común con la de los modelos anteriores.

Tabla 37: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	4820
Number of Observations Used	4820

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	398
2	sano(>=60)	4422

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo; se obtienen la siguiente tablas de estimaciones de los parámetros del modelo.

Tabla 38: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	Leucocitos	1	6	1.3167	0.2512	Leucocitos
2	GGT	1	5	2.1728	0.1405	GGT

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.0826	0.3109	44.8605	<.0001
V1HORA	1	0.0356	0.00255	195.2109	<.0001
COL	1	-0.00473	0.00142	11.1569	0.0008
GOT	1	0.0183	0.00437	17.6127	<.0001
GPT	1	-0.0250	0.00506	24.3787	<.0001
HBA1C	1	0.0514	0.0211	5.9214	0.0150

Nótese que en este modelo sí se incluye la variable HBA1C, que había sido descartada en los dos modelos anteriores. En la segunda columna de la tabla se tienen los valores de los estimadores de los parámetros beta del modelo elegido por el procedimiento backwards. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\log(-\log(\pi(x))) = -2.0826 + 0.0356V1HORA + \dots + 0.054HBA1C \quad (ec. 8)$$

La segunda columna da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. En la tabla 38 aparecen primero los p-valores de las variables descartadas y a continuación los p-valores de las variables incluidas en el modelo.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 39: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2749.478	2558.514
SC	2755.958	2597.398
-2 Log L	2747.478	2546.514

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 40: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	200.9633	5	<.0001
Score	364.0372	5	<.0001
Wald	298.4263	5	<.0001

Y como en los casos anteriores, se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda por observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 41: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	482	15	15.00	467	467.00
2	482	14	22.54	468	459.46
3	482	25	26.38	457	455.62
4	482	21	29.60	461	452.40
5	482	27	32.63	455	449.37
6	482	35	35.67	447	446.33
7	482	26	39.27	456	442.73
8	482	49	43.89	433	438.11
9	482	51	51.59	431	430.41
10	482	135	95.93	347	386.07

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
32.5988	8	<.0001

A partir de estas tablas, al contrario de los otros modelos, se ve que el modelo no ajusta bien con los datos, con lo que vamos a descartar este modelo como adecuado.

En este caso analizando los datos únicamente de los hombres, lo que observamos es que la Bondad de Ajuste de Hosmer y Lemeshow es bastante aceptable tomando el modelo probit y también el modelo logit.

- $\text{logit}(\text{FGR}_i) = \text{V1HORA}, \text{COL}, \text{GOT}, \text{GPT}; p\text{-valor}=\mathbf{0.1979}$.
- $\text{probit}(\text{FGR}_i) = \text{V1HORA}, \text{COL}, \text{GOT}, \text{GPT}; p\text{-valor}=\mathbf{0.2159}$.
- $\text{cloglog}(\text{FGR}_i) = \text{V1HORA}, \text{COL}, \text{GOT}, \text{GPT}, \text{HBA1C}; p\text{-valor}=\mathbf{<0.0001}$

POBLACION EXCLUSIVAMENTE DE MUJERES CON EDADES DE 2-99 AÑOS

Modelo logit

Pasamos ahora a estudiar por separado la población de mujeres. En este caso tenemos 4.930 datos. El número de individuos con valor FGR inferior a 60 (valor 1) y con valor FGR superior a 60 (valor 0) vienen dados por

Tabla 42: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	4930
Number of Observations Used	4930

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	297
2	sano(>=60)	4633

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo. La tabla 43 muestra las estimaciones de los parámetros del modelo.

Tabla 43: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.0837	0.7723	COL

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.9432	0.3283	327.6252	<.0001
Leucocitos	1	0.0776	0.0240	10.4505	0.0012
V1HORA	1	0.0388	0.00475	66.6553	<.0001
GOT	1	0.0644	0.0106	36.6955	<.0001
GGT	1	0.00420	0.00148	8.0892	0.0045
GPT	1	-0.0522	0.00905	33.2434	<.0001
HBA1C	1	0.2945	0.0496	35.2563	<.0001

Se observa que en este caso se ha descartado la variable COL en el modelo, pero esto no es lo que había ocurrido en el caso de la población masculina ni en el caso de la población conjunta. Como en los otros casos, en la segunda columna de esta tabla se tienen los valores de los estimadores de los parámetros beta del modelo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\pi(x) = \frac{\exp(-5.9432 + 0.0776Leucocitos + \dots + 0.2945HBA1C)}{1 + \exp(-5.9432 + 0.0776Leucocitos + \dots + 0.2945HBA1C)} \quad (ec. 9)$$

La segunda columna da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. Primero aparece el p-valor para la variable descartada COL (que es muy grande) y a continuación los p-valores para las variables explicativas que se han incluido en el modelo (que son muy pequeños).

Observando la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 44: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2246.498	2054.040
SC	2253.001	2099.561
-2 Log L	2244.498	2040.040

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 45: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	204.4588	6	<.0001
Score	296.3321	6	<.0001
Wald	193.3110	6	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Las “odds” de las distintas variables explicativas vienen dadas en la tabla 46.

Tabla 46: Odds de las variables explicativas

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Leucocitos	1.081	1.031	1.133
V1HORA	1.040	1.030	1.049
GOT	1.066	1.045	1.089
GGT	1.004	1.001	1.007
GPT	0.949	0.932	0.966
HBA1C	1.342	1.218	1.480

En este caso se ve que todas las variables incluidas en el modelo excepto GPT hacen que aumente la probabilidad de tener la enfermedad, mientras que GPT hace que la probabilidad disminuya si aumenta su valor.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 47: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	493	13	9.89	480	483.11
2	493	8	14.00	485	479.00
3	493	5	16.21	488	476.79
4	493	5	18.38	488	474.62
5	493	15	20.73	478	472.27
6	494	20	23.53	474	470.47
7	493	26	26.82	467	466.18
8	493	48	31.79	445	461.21
9	493	57	40.97	436	452.03
10	492	100	94.68	392	397.32

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
40.0586	8	<.0001

A partir de estas tablas, se ve que el modelo no ajusta bien con los datos, con lo que vamos a descartar este modelo como adecuado. Nótese la diferencia con la población de hombres, en el que el modelo sí ajustaba bastante bien.

Modelo probit

La primera tabla es común con el modelo anterior.

Tabla 48: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	4930
Number of Observations Used	4930

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	297
2	sano(>=60)	4633

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo, la siguiente tabla muestra las estimaciones de los parámetros del modelo.

Tabla 49: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.3809	0.5371	COL

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.2507	0.1666	380.5629	<.0001
Leucocitos	1	0.0420	0.0121	12.1130	0.0005
V1HORA	1	0.0210	0.00255	67.7983	<.0001
GOT	1	0.0297	0.00472	39.5315	<.0001
GGT	1	0.00212	0.000661	10.3025	0.0013
GPT	1	-0.0226	0.00385	34.5147	<.0001
HBA1C	1	0.1563	0.0262	35.4684	<.0001

Al igual que en modelo logit, se ha descartado la variable COL del modelo. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\Phi(\pi(x)) = -3.2507 + 0.0420Leucocitos + \dots + 0.1563HBA1C \quad (ec. 10)$$

La segunda columna da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. Al igual que en el caso anterior, se puede ver primero que el p-valor para COL es muy grande, mientras que para las otras variables explicativas, el p-valor que aparece a continuación es muy pequeño.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 50: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2246.498	2043.409
SC	2253.001	2088.931
-2 Log L	2244.498	2029.409

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 51: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	215.0894	6	<.0001
Score	226.3791	6	<.0001
Wald	212.4843	6	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda estudiar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 52: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	493	12	8.77	481	484.23
2	493	6	12.61	487	480.39
3	493	5	15.03	488	477.97
4	493	8	17.41	485	475.59
5	493	14	19.93	479	473.07
6	493	22	22.89	471	470.11
7	493	30	26.72	463	466.28
8	493	46	32.20	447	460.80
9	493	51	42.20	442	450.80
10	493	103	95.52	390	397.48

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
28.2942	8	0.0004

A partir de estas tablas, se ve que el modelo no ajusta bien con los datos, con lo que vamos a descartar este modelo como adecuado. Véase que en el caso de la población de hombres el modelo sí ajustaba bastante bien, y en la población conjunta se podía llegar a aceptar.

Modelo log-log

La primera tabla es común en los tres modelos.

Tabla 53: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	4930
Number of Observations Used	4930

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	297
2	sano(>=60)	4633

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo; se obtiene la siguiente tabla de estimaciones de los parámetros del modelo.

Tabla 54: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.2088	0.6477	COL

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.0223	0.2374	447.6045	<.0001
Leucocitos	1	0.0592	0.0135	19.0944	<.0001
V1HORA	1	0.0293	0.00368	63.1568	<.0001
GOT	1	0.0403	0.00580	48.2572	<.0001
GGT	1	0.00239	0.000717	11.1239	0.0009
GPT	1	-0.0336	0.00636	27.9532	<.0001
HBA1C	1	0.2082	0.0392	28.2595	<.0001

Nuevamente se ha descartado la variable COL. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo para las variables incluidas. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\log(-\log(\pi(x))) = -5.0223 + 0.0592Leucocitos + \dots + 0.2082HBA1C \quad (ec. 11)$$

La segunda columna de la segunda tabla da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y en la última el p-valor, muy alto para COL (la variable descartada) y muy pequeño para las otras variables.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 55: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2246.498	2081.198
SC	2253.001	2126.719
-2 Log L	2244.498	2067.198

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 56: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	177.3007	6	<.0001
Score	321.0077	6	<.0001
Wald	302.1395	6	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda estudiar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 57: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	493	13	13.55	480	479.45
2	493	5	17.34	488	475.66
3	493	8	19.17	485	473.83
4	493	6	20.99	487	472.01
5	493	15	22.86	478	470.14
6	493	19	24.97	474	468.03
7	493	35	27.50	458	465.50
8	493	42	31.12	451	461.88
9	493	48	37.60	445	455.40
10	493	106	75.44	387	417.56

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
55.3741	8	<.0001

A partir de estas tablas, se ve que el modelo no ajusta bien con los datos, con lo que vamos a descartar este modelo como adecuado.

En este otro caso teniendo solamente en cuenta la población de mujeres la Bondad de Ajuste de Hosmer y Lemeshow para los datos no es aceptable para ninguno de los tres modelos

- $\text{logit}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{GOT}, \text{GGT}, \text{GPT}, \text{HBA1C}; p\text{-valor} < 0.0001.$
- $\text{probit}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{GOT}, \text{GGT}, \text{GPT}, \text{HBA1C}; p\text{-valor} = 0.0004.$
- $\text{cloglog}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{GOT}, \text{GGT}, \text{GPT}, \text{HBA1C}; p\text{-valor} < 0.0001.$

En definitiva, los tres modelos ajustan mal en el caso de la población de mujeres y ajustan bastante mejor en la población de hombres.

Se ha visto que los modelos anteriores no son completamente satisfactorios ya que el test de Hosmer y Lemeshow no da un p-valor muy grande. Ello parece sugerir que las variables explicativas no son capaces de explicar los valores de FGR a partir de estos modelos. Por otra parte, el National Kidney Disease Education Program (NKDEP) de Estados Unidos advierte que el valor de FGR no debe ser usada en edades menores a 18 ni mayores a 70 años. (*Myers GL et al, 2006*). En consecuencia, se ha decidido repetir el estudio, pero restringiéndonos a la población con edades comprendidas entre 18 y 70 años.

POBLACION CONJUNTA CON EDADES DE 18-70 AÑOS

Modelo logit

En el caso de la población conjunta, sin tener en cuenta las diferencias entre hombres y mujeres; tenemos 7.543 datos. El número de individuos con valor FGR inferior a 60 (valor 1) y con valor FGR superior a 60 (valor 0) vienen dados por

Tabla 58: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	7543
Number of Observations Used	7543

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	275
2	sano(>=60)	7268

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo, la tabla 59 muestra las estimaciones de los parámetros del modelo.

Tabla 59: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	GGT	1	6	0.0024	0.9607	GGT
2	COL	1	5	1.0187	0.3128	COL

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.9865	0.2632	358.9812	<.0001
Leucocitos	1	0.0795	0.0249	10.2268	0.0014
V1HORA	1	0.0402	0.00478	70.9172	<.0001
GOT	1	0.0210	0.00619	11.5189	0.0007
GPT	1	-0.0134	0.00527	6.5221	0.0107
HBA1C	1	0.0993	0.0339	8.5662	0.0034

Podemos observar que las variables GGT y COL se han descartado. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo para las variables incluidas en el mismo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\pi(x) = \frac{\exp(-4.9865 + 0.0795Leucocitos + \dots + 0.0993HBA1C)}{1 + \exp(-4.9865 + 0.0795Leucocitos + \dots + 0.0993HBA1C)} \quad (ec. 12)$$

La segunda columna de la segunda tabla da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en la siguiente columna y el p-valor correspondiente del contraste viene dado en la última columna. Entonces, puede verse que el p-valor es grande para las variables descartadas y pequeño para las variables que se incluyen en el modelo.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 60: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2363.232	2266.694
SC	2370.161	2308.265
-2 Log L	2361.232	2254.694

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros. Nótese sin embargo que estas diferencias no son tan acusadas como en el caso de usar todos los datos.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 61: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	106.5379	5	<.0001
Score	172.1955	5	<.0001
Wald	117.2921	5	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

La tabla 62 recoge las “odds” de las distintas variables explicativas.

Tabla 62: Odds de las variables explicativas

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Leucocitos	1.083	1.031	1.137
V1HORA	1.041	1.031	1.051
GOT	1.021	1.009	1.034
GPT	0.987	0.977	0.997
HBA1C	1.104	1.033	1.180

En este caso se ve que todas las variables incluidas en el modelo salvo GPT hacen que aumente la probabilidad de tener la enfermedad si aumentan su valor, mientras que GPT hace que disminuya.

Finalmente, nos queda por observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla

Tabla 63: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	755	11	14.99	744	740.01
2	754	19	17.37	735	736.63
3	754	20	18.90	734	735.10
4	754	17	20.37	737	733.63
5	754	16	21.93	738	732.07
6	754	20	23.74	734	730.26
7	754	26	25.87	728	728.13
8	754	30	28.77	724	725.23
9	754	47	34.42	707	719.58
10	756	69	68.65	687	687.35

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
9.0141	8	0.3411

A partir de estas tablas, se ve que el modelo ajusta bien los datos, con lo que vamos a aceptar el modelo. Nótese que el p-valor es mucho más significativo de lo que se había obtenido en los modelos sin restricción de edad.

Modelo probit

Empezamos con la información de la muestra, que es común a los tres modelos.

Tabla 64: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	7543
Number of Observations Used	7543

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	275
2	sano(>=60)	7268

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo. La siguiente tabla muestra las estimaciones de los parámetros del modelo.

Tabla 65: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	GGT	1	6	0.0516	0.8202	GGT
2	COL	1	5	1.0175	0.3131	COL

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.6815	0.1288	433.6326	<.0001
Leucocitos	1	0.0377	0.0125	9.0403	0.0026
V1HORA	1	0.0202	0.00244	68.5776	<.0001
GOT	1	0.0107	0.00323	10.9694	0.0009
GPT	1	-0.00685	0.00255	7.2034	0.0073
HBA1C	1	0.0586	0.0164	12.7792	0.0004

Observamos que en este caso se han descartado las variables GGT y COL. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo para las variables incluidas en el mismo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\Phi(\pi(x)) = -2.6815 + 0.0377Leucocitos + \dots + 0.05986HBA1C \quad (ec. 13)$$

La segunda columna de la segunda tabla da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. Y como en casos anteriores, se tiene un p-valor grande para las variables descartadas y pequeño para las variables incluidas en el modelo. .

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 66: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2363.232	2261.829
SC	2370.161	2303.399
-2 Log L	2361.232	2249.829

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 67: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	111.4037	5	<.0001
Score	122.1999	5	<.0001
Wald	116.0557	5	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 68: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	754	11	13.31	743	740.69
2	754	18	16.05	736	737.95
3	756	19	17.89	737	738.11
4	755	18	19.56	737	735.44
5	754	17	21.37	737	732.63
6	754	21	23.47	733	730.53
7	754	24	25.99	730	728.01
8	754	30	29.45	724	724.55
9	754	45	35.99	709	718.01
10	754	72	71.37	682	682.63

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
4.5795	8	0.8014

A partir de estas tablas, se ve que el modelo ajusta extraordinariamente bien los datos, con lo que vamos a aceptar este modelo como el modelo adecuado para explicar el comportamiento de la variable FGR.

Modelo log-log

La primera tabla (tabla 69) es común con los otros dos modelos.

Tabla 69: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	7543
Number of Observations Used	7543

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	275
2	sano(>=60)	7268

Probability modeled is FGR_='enfermo(<60)'.

Comenzamos por las estimaciones de los parámetros.

Tabla 70: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	GGT	1	6	0.0046	0.9459	GGT
2	COL	1	5	0.8659	0.3521	COL

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.7638	0.2084	522.3874	<.0001
Leucocitos	1	0.0765	0.0219	12.2044	0.0005
V1HORA	1	0.0339	0.00348	95.0146	<.0001
GOT	1	0.0188	0.00495	14.5027	0.0001
GPT	1	-0.0117	0.00455	6.5925	0.0102
HBA1C	1	0.0732	0.0204	12.8406	0.0003

Nuevamente se han descartado las variables GGT y COL del modelo. Véase que el p-valor del test de Wald es muy grande, lo que lleva a su exclusión del modelo. Las otras variables explicativas sí están incluidas en el modelo ya que el p-valor para el contraste de Wald es muy pequeño. En la segunda columna se tienen los valores de los estimadores de los parámetros beta del modelo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\log(-\log(\pi(x))) = -4.7638 + 0.0765Leucocitos + \dots + 0.0732HBA1C \quad (ec. 14)$$

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 71: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2363.232	2271.181
SC	2370.161	2312.751
-2 Log L	2361.232	2259.181

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 72: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	102.0514	5	<.0001
Score	183.1392	5	<.0001
Wald	166.9159	5	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 73: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	755	12	16.19	743	738.81
2	756	18	18.46	738	737.54
3	754	19	19.81	735	734.19
4	754	16	21.17	738	732.83
5	754	21	22.55	733	731.45
6	755	18	24.20	737	730.80
7	754	27	26.04	727	727.96
8	754	31	28.59	723	725.41
9	754	46	33.47	708	720.53
10	753	67	62.88	686	690.12

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
9.6551	8	0.2901

A partir de estas tablas, se ve que el modelo ajusta bastante bien los datos, aunque no es tan bueno como el modelo probit.

En este caso vemos que la Bondad de Ajuste de Hosmer y Lemeshow para los datos agregados es muy buena especialmente para el modelo probit

- $\text{logit}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{GOT}, \text{GPT}, \text{HBA1C}; p\text{-valor}=0.3411.$
- $\text{probit}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{GOT}, \text{GPT}, \text{HBA1C}; p\text{-valor}=\mathbf{0.8014}.$
- $\text{cloglog}(\text{FGR}_i) = \text{Leucocitos}, \text{V1HORA}, \text{GOT}, \text{GPT}, \text{HBA1C}; p\text{-valor}=0.2901.$

Vamos ahora a considerar un estudio por separado para la población de hombres y de mujeres.

POBLACION EXCLUSIVAMENTE DE HOMBRES CON EDADES DE 18-70 AÑOS

Modelo logit

En este caso tenemos 3.657 datos. El número de individuos con valor FGR inferior a 60 (valor 1) y con valor FGR superior a 60 (valor 0) vienen dados por

Tabla 74: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	3657
Number of Observations Used	3657

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	159
2	sano(>=60)	3498

Probability modeled is FGR_='enfermo(<60)'.

Las estimaciones de los parámetros del modelo vienen dadas por:

Tabla 75: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.0409	0.8396	COL
2	Leucocitos	1	5	0.2276	0.6333	Leucocitos
3	GGT	1	4	1.9345	0.1643	GGT
4	HBA1C	1	3	3.0809	0.0792	HBA1C

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.5783	0.1879	362.8314	<.0001
V1HORA	1	0.0553	0.00687	64.6477	<.0001
GOT	1	0.0164	0.00752	4.7396	0.0295
GPT	1	-0.0161	0.00677	5.6368	0.0176

Llama la atención que se han descartado cuatro variables explicativas, con lo que ha quedado un modelo muy simple, con sólo tres variables explicativas significativas. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo para las variables incluidas en el mismo. Entonces, la probabilidad de que la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\pi(x) = \frac{\exp(-3.5783 + 0.0553V1HORA + 0.0164GOT - 0.0161GPT)}{1 + \exp(-3.5783 + 0.0553V1HORA + 0.0164GOT - 0.0161GPT)} \quad (ec. 15)$$

La segunda columna da las desviaciones estándar de las estimaciones. El test de Wald se utiliza para ver qué variables hay que incluir en el modelo. Primero se dan los p-valores (grandes) para las variables excluidas y a continuación los p-valores (pequeños) para las variables incluidas en el modelo.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 76: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1310.072	1239.014
SC	1316.276	1263.832
-2 Log L	1308.072	1231.014

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 77: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	77.0572	3	<.0001
Score	145.6042	3	<.0001
Wald	72.4878	3	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Las “odds” de las distintas variables explicativas vienen dadas en la siguiente tabla.

Tabla 78: Odds de las variables explicativas

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
V1HORA	1.057	1.043	1.071
GOT	1.017	1.002	1.032
GPT	0.984	0.971	0.997

En este caso se ve que las variables V1HORA y GOT y HBA1C hacen que aumente la probabilidad de tener la enfermedad cuando aumentan de valor, mientras que GPT hace que la probabilidad disminuya.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 79: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	366	4	8.18	362	357.82
2	367	14	10.31	353	356.69
3	368	11	11.24	357	356.76
4	367	10	11.96	357	355.04
5	367	12	12.64	355	354.36
6	367	14	13.52	353	353.48
7	366	13	14.56	353	351.44
8	366	10	16.09	356	349.91
9	366	24	18.52	342	347.48
10	357	47	41.97	310	315.03

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
8.9113	8	0.3498

A partir de estas tablas, se ve que el modelo ajusta bastante bien los datos.

Modelo probit

La primera tabla es común a los tres modelos.

Tabla 80: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	3657
Number of Observations Used	3657

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	159
2	sano(>=60)	3498

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo; se obtiene la siguiente tabla de estimaciones de los parámetros del modelo.

Tabla 81: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.1113	0.7387	COL
2	Leucocitos	1	5	0.3525	0.5527	Leucocitos
3	GGT	1	4	1.5079	0.2195	GGT
4	HBA1C	1	3	3.1895	0.0741	HBA1C

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.9529	0.0849	528.6463	<.0001
V1HORA	1	0.0286	0.00343	69.5594	<.0001
GOT	1	0.00817	0.00398	4.2150	0.0401
GPT	1	-0.00814	0.00326	6.2551	0.0124

Es significativo que se han descartado las mismas variables que con el modelo logit. En la segunda columna de la segunda tabla se tienen los valores de los estimadores de los parámetros beta del modelo para las variables incluidas en el mismo. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\Phi(\pi(x)) = -1.9529 + 0.0286V1HORA + 0.00817GOT - 0.00814GPT \quad (ec. 16)$$

La segunda columna da las desviaciones estándar de las estimaciones. Para comprobar si estas estimaciones pueden ser consideradas distintas de cero, se realiza el test de Wald. El valor del estadístico del contraste viene dado en las siguientes columnas y los p-valores correspondientes del contraste vienen dado en la última columna.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 82: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1310.072	1235.868
SC	1316.276	1260.686
-2 Log L	1308.072	1227.868

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 83: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	80.2037	3	<.0001
Score	83.2634	3	<.0001
Wald	77.6750	3	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 84: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	366	4	7.29	362	358.71
2	368	14	9.70	354	358.30
3	366	10	10.68	356	355.32
4	368	11	11.60	357	356.40
5	366	12	12.33	354	353.67
6	367	15	13.38	352	353.62
7	366	12	14.59	354	351.41
8	366	10	16.36	356	349.64
9	366	24	19.15	342	346.85
10	358	47	43.20	311	314.80

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
8.5046	8	0.3858

A partir de estas tablas, se ve que el modelo ajusta bien los datos, aunque no tan bien como ocurría en el caso de la población conjunta.

Modelo log-log

Información sobre los datos, común con los otros modelos.

Tabla 85: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	3657
Number of Observations Used	3657

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	159
2	sano(>=60)	3498

Probability modeled is FGR_='enfermo(<60)'.
 (Note: The original image contains a typo '60' in the text, which has been corrected to '60' based on the table content.)

La tabla 86 nos muestra las estimaciones de los parámetros del modelo.

Tabla 86: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	Leucocitos	1	6	0.0026	0.9593	Leucocitos
2	COL	1	5	0.0357	0.8502	COL
3	GGT	1	4	1.7004	0.1922	GGT

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.8092	0.2410	249.8935	<.0001
V1HORA	1	0.0406	0.00397	104.6543	<.0001
GOT	1	0.0159	0.00625	6.4915	0.0108
GPT	1	-0.0151	0.00636	5.6143	0.0178
HBA1C	1	0.0574	0.0273	4.4076	0.0358

Este modelo descarta menos variables que los otros dos modelos. Esto dará lugar a un modelo más complejo. En este caso, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\log(-\log(\pi(x))) = -3.8092 + 0.0406V1HORA + \dots + 0.0574HBA1C \quad (ec. 17)$$

La segunda columna da la segunda tabla da las desviaciones estándar de las estimaciones. Y se ha usado el test de Wald para determinar las variables explicativas que finalmente se incluyen en el modelo.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 87: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1310.072	1244.986
SC	1316.276	1276.008
-2 Log L	1308.072	1234.986

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 88: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	73.0856	4	<.0001
Score	167.6289	4	<.0001
Wald	120.7096	4	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 89: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	366	4	8.88	362	357.12
2	366	17	10.93	349	355.07
3	366	6	11.78	360	354.22
4	366	8	12.48	358	353.52
5	366	17	13.15	349	352.85
6	366	6	13.91	360	352.09
7	366	15	14.84	351	351.16
8	366	15	16.19	351	349.81
9	367	19	18.29	348	348.71
10	362	52	36.71	310	325.29

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
23.8707	8	0.0024

A partir de estas tablas, se ve que el modelo no ajusta bien con los datos, con lo que vamos a descartar este modelo como adecuado. Véase la diferencia entre este modelo y los otros dos: no solamente considera unas variables explicativas diferentes, sino que incluso teniendo en cuenta más variables, no es capaz de explicar satisfactoriamente el comportamiento de la variable FGR.

Con estos resultados podemos decir que la Bondad de Ajuste de Hosmer y Lemeshow para los datos correspondientes a los hombres es bastante aceptable tomando el modelo probit y también el modelo logit.

- $\text{logit}(\text{FGR}_) = \text{V1HORA, GOT, GPT}$; p-valor=**0.3498**.
- $\text{probit}(\text{FGR}_) = \text{V1HORA, GOT, GPT}$; p-valor=**0.3858**.
- $\text{cloglog}(\text{FGR}_) = \text{V1HORA, GOT, GPT, HBA1C}$; p-valor=0.0024.

POBLACION EXCLUSIVAMENTE DE MUJERES CON EDADES DE 18-70 AÑOS

Modelo logit

En este caso, la información muestral es la siguiente:

Tabla 90: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	3886
Number of Observations Used	3886

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	116
2	sano(>=60)	3770

Probability modeled is FGR_='enfermo(<60)'.

Se estudia cómo se ajustan los datos al modelo y se obtiene la tabla 91 de estimaciones de los parámetros del modelo.

Tabla 91: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.0476	0.8272	COL
2	GGT	1	5	1.2377	0.2659	GGT
3	GPT	1	4	3.0954	0.0785	GPT
4	GOT	1	3	2.9057	0.0883	GOT

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.5813	0.4456	156.8981	<.0001
Leucocitos	1	0.0994	0.0356	7.8000	0.0052
V1HORA	1	0.0340	0.00766	19.7317	<.0001
HBA1C	1	0.1892	0.0737	6.5826	0.0103

Se observa que se han descartado cuatro variables, lo mismo que pasaba en la población de hombres, pero las variables descartadas no coinciden en los dos casos. Esto explica que en el modelo con la población global no se hayan descartado tantas variables. En la segunda columna se tienen los valores de los estimadores de los parámetros beta del modelo para las variables incluidas. Entonces, la probabilidad de que se la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\pi(x) = \frac{\exp(-5.5813 + 0.0994Leucocitos + 0.034V1HORA + 0.1892HBA1C)}{1 + \exp(-5.5813 + 0.0994Leucocitos + 0.034V1HORA + 0.1892HBA1C)} \quad (ec. 18)$$

La segunda columna da las desviaciones estándar de las estimaciones. El test de Wald determina las variables que se deben incluir en el modelo.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 92: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1045.181	1011.932
SC	1051.446	1036.993
-2 Log L	1043.181	1003.932

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 93: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	39.2484	3	<.0001
Score	58.1573	3	<.0001
Wald	45.6885	3	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Las tabla 94 nos muestra las “odds” de las distintas variables explicativas.

Tabla 94: Odds de las variables explicativas

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Leucocitos	1.105	1.030	1.184
V1HORA	1.035	1.019	1.050
HBA1C	1.208	1.046	1.396

En este caso se ve que todas las variables incluidas en el modelo hacen que aumente la probabilidad de tener la enfermedad al aumentar su valor.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 95: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	389	3	6.33	386	382.67
2	390	5	7.29	385	382.71
3	389	8	7.95	381	381.05
4	389	6	8.61	383	380.39
5	389	7	9.29	382	379.71
6	389	9	10.03	380	378.97
7	389	14	11.06	375	377.94
8	389	19	12.59	370	376.41
9	390	23	15.21	367	374.79
10	383	22	27.63	361	355.37

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
13.5822	8	0.0933

A partir de estas tablas, se ve que el modelo ajusta bien los datos, pero no tan bien como ajustaba con la población global y con la población de hombres.

Modelo probit

La información muestral es la misma que en el caso anterior.

Tabla 96: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	3886
Number of Observations Used	3886

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	116
2	sano(>=60)	3770

Probability modeled is FGR_='enfermo(<60)'.

En primer lugar, se estudia cómo se ajustan los datos al modelo y se obtiene la siguiente tabla de estimaciones de los parámetros del modelo.

Tabla 97: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.1028	0.7484	COL
2	GGT	1	5	1.7148	0.1904	GGT
3	GPT	1	4	3.4962	0.0615	GPT
4	GOT	1	3	2.8646	0.0905	GOT

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	
Intercept	1	-2.9522	0.2166	185.7725	<.0001	
Leucocitos	1	0.0434	0.0180	5.8372	0.0157	
V1HORA	1	0.0169	0.00381	19.6798	<.0001	
HBA1C	1	0.1071	0.0366	8.5870	0.0034	

Nótese que da una información idéntica a la del modelo logit en cuanto a qué variables deben ser incluidas en el modelo. Entonces, la probabilidad de que la variable FGR tome un valor significativo (por debajo de 60) viene dado por la ecuación:

$$\Phi(\pi(x)) = -2.9522 + 0.0434Leucocitos + 0.0169V1HORA + 0.1071HBA1C \quad (ec. 19)$$

La segunda columna nos muestra las desviaciones estándar de las estimaciones. El test de Wald explica las razones de excluir unas variables en el modelo y de incluir otras.

Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 98: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1045.181	1009.167
SC	1051.446	1034.228
-2 Log L	1043.181	1001.167

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 99: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	42.0134	3	<.0001
Score	47.8772	3	<.0001
Wald	45.3649	3	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 100: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	389	1	5.62	388	383.38
2	389	6	6.68	383	382.32
3	389	9	7.45	380	381.55
4	389	6	8.20	383	380.80
5	389	7	8.98	382	380.02
6	389	7	9.84	382	379.16
7	389	15	11.02	374	377.98
8	389	21	12.85	368	376.15
9	389	20	15.84	369	373.16
10	385	24	29.36	361	355.64

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
15.1557	8	0.0562

A partir de estas tablas, se ve que el modelo no ajusta bien con los datos, con lo que vamos a descartar este modelo como adecuado. El p-valor no es excesivamente pequeño, pero tampoco es tan grande como el que hemos encontrado en otros modelos.

Modelo log-log

La información muestral viene dada por

Tabla 101: Numero de muestras con FGR<60 y FGR>60

Number of Observations Read	3886
Number of Observations Used	3886

Response Profile		
Ordered Value	FGR_	Total Frequency
1	enfermo(<60)	116
2	sano(>=60)	3770

Probability modeled is FGR_='enfermo(<60)'.

La tabla 102 nos muestra las estimaciones de los parámetros del modelo:

Tabla 102: Análisis de ajuste de los datos al modelo

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	COL	1	6	0.0641	0.8002	COL
2	GGT	1	5	1.3087	0.2526	GGT
3	GPT	1	4	3.0746	0.0795	GPT
4	GOT	1	3	3.5634	0.0591	GOT

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.3246	0.3823	193.9721	<.0001
Leucocitos	1	0.1005	0.0291	11.9172	0.0006
V1HORA	1	0.0293	0.00664	19.4084	<.0001
HBA1C	1	0.1484	0.0661	5.0381	0.0248

Véase que se han descartado las mismas variables que en los otros dos modelos. Entonces, la probabilidad de que se la variable FGR tome un valor significativo viene dado por la ecuación:

$$\log(-\log(\pi(x))) = -5.3246 + 0.1005Leucocitos + 0.0293V1HORA + 0.1484HBA1C \quad (ec. 20)$$

Las variables han sido descartadas o incluidas a partir de los valores del test de Wald. Si ahora observamos la bondad de ajuste del modelo se obtiene la siguiente tabla:

Tabla 103: Bondad del ajuste del modelo

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1045.181	1013.501
SC	1051.446	1038.562
-2 Log L	1043.181	1005.501

Se observan los valores de los estadísticos de los contrastes, en el caso en que se considera solamente el término independiente (primera columna) y el modelo con todos los parámetros del modelo (segunda columna); como los valores en la primera columna son mayores que los que aparecen en la segunda columna, esto es indicativo de que el modelo necesita los distintos parámetros.

A continuación, se estudia para este modelo el contraste de hipótesis nula global

$$H_0: \beta = 0 \text{ frente a } H_0: \beta \neq 0$$

La tabla que se obtiene es:

Tabla 104: Contraste de hipótesis nula global

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	37.6797	3	<.0001
Score	59.8616	3	<.0001
Wald	60.8704	3	<.0001

Y como en todos los casos se obtiene un p-valor muy pequeño, se concluye que se rechaza la hipótesis nula y que es necesario incluir los demás parámetros del modelo.

Finalmente, nos queda observar si el modelo se ajusta bien a los datos. Para ello consideramos el contraste de Hosmer y Lemeshow, cuyos resultados aparecen en la siguiente tabla:

Tabla 105: Contraste de Hosmer y Lemeshow

Partition for the Hosmer and Lemeshow Test					
Group	Total	FGR_ = enfermo(<60)		FGR_ = sano(>=60)	
		Observed	Expected	Observed	Expected
1	389	3	6.72	386	382.28
2	389	7	7.62	382	381.38
3	389	9	8.27	380	380.73
4	389	4	8.90	385	380.10
5	389	6	9.54	383	379.46
6	389	10	10.23	379	378.77
7	390	15	11.19	375	378.81
8	389	17	12.51	372	376.49
9	389	23	14.79	366	374.21
10	384	22	25.66	362	358.34

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
14.6151	8	0.0671

A partir de estas tablas, se ve que el modelo no ajusta tan bien los datos como otros modelos, con lo que vamos a descartar este modelo como adecuado. Como en el caso anterior, no es un valor excesivamente bajo, pero hemos encontrado otros modelos mejores.

En este último caso podemos decir que la Bondad de Ajuste de Hosmer y Lemeshow para los datos correspondientes a las mujeres es difícilmente aceptable, aunque quizá los modelos logit y probit estarían al límite de lo aceptable con un nivel de significación de 0.05

- $\text{logit}(\text{FGR}_) = \text{Leucocitos, V1HORA, HBA1C}$; p-valor=**0.0933**.
- $\text{probit}(\text{FGR}_) = \text{Leucocitos, V1HORA, HBA1C}$; p-valor=0.0562.
- $\text{cloglog}(\text{FGR}_) = \text{Leucocitos, V1HORA, HBA1C}$; p-valor=0.0671.

Una vez realizados los estudios con los tres modelos y en las distintas variantes de las muestras, podemos resumir lo siguiente:

Se ha llevado a cabo un estudio de tres modelos diferentes:

- Modelo de regresión logística o Modelo logit;
- Modelo probit;
- Modelo loglog complementario (cloglog).

El uso de estos modelos es, porque como hemos comentado anteriormente el método de regresión logística busca explicar el resultado de una variable binaria Y en función de un grupo de variables explicativas, hemos encontrado otro caso en la bibliografía en el que se utilizaba para ver la asociación entre enfermedad renal y enfermedad cardiovascular en Corea, con una población de 53.221 hombres con edades entre 21 y 77 años (*Ryoo JH et al, 2011*) y además se usan muy frecuentemente en estudios médicos (*Bagley SC et al, 2001*) en los que se busca explicar la probabilidad de enfermar a partir de unas variables, y los otros dos (probit y log-log) nos los ha proporcionado el programa estadístico SAS (*SAS Institute Inc. 2014*), por ser similares al de regresión logística, ser modelos que ya se han usado en otros estudios médicos (*Pourhoseingholi, A et al, 2010*) y nos ha permitido compararlos para ver cuál era el que mejor se ajustaba.

Para las tres modelizaciones se ha utilizado el Procedimiento LOGISTIC haciendo uso de la función enlace correspondiente. Como variable respuesta se ha tomado la variable FGR categorizada, FGR_, según sea menor a 60 mL/min/1.73 m² (individuo enfermo crónico renal) o mayor o igual a 60 mL/min/1.73 m² (individuo sano, sin enfermedad crónica renal). El resto de las variables (Leucocitos, V1HORA, COL, GOT, GGT, GPT, HBA1C), se han tomado como variables explicativas.

La variable FGR lleva implícita la información relativa al sexo de los individuos, así que en un principio no sería necesario realizar un estudio separado para hombres y mujeres. En una fase inicial se han tomado 9.750 individuos (hombres y mujeres entre 2 y 99 años), tanto agregados como desagregados por sexo y se han ajustado los tres modelos. A continuación se ha realizado el mismo estudio pero solamente con la población con una edad comprendida entre los 18 y los 70 años, siguiendo las recomendaciones del National Kidney Disease Education Program (NKDEP) de Estados Unidos advierte que el valor de FGR no debe ser usada en edades menores a 18 ni mayores a 70 años. (*Myers GL et al, 2006*).

El estudio se ha realizado únicamente con población del corredor del Henares, por lo que tenemos que tener claro, que los resultados podrían variar fuera de esta población, ya que es posible que cualquier otra población frente a la población de estudio no fueran homogéneas respecto a las variables del estudio.

De todos los estudios realizados el que mejor resultados ha ofrecido es el modelo probit para la población conjunta entre 18 y 70 años, ya que es el que ha obtenido un p-valor más alto (0.8014), y explica bastante bien la probabilidad de éxito a partir de unas variables explicativas, y además concuerda con la recomendación del NKDEP de no utilizar el valor de FGR en población fuera de ese rango de edad.

Además se han obtenido valores también bastante aceptables para los modelos probit y logit, cuando se ha hecho el estudio sobre la población exclusivamente de hombres, con aquellos con una edad comprendida entre 18 y 70 años. Ninguno de los modelos utilizados en el estudio con mujeres exclusivamente se ha mostrado válido, y es que aunque hay un estudio reciente que indica las diferencias específicas de sexo en la susceptibilidad a enfermedades renales (*Seppi T et al, 2016*), indicando que las hormonas femeninas podrían hacer a las mujeres menos susceptibles a desarrollar insuficiencia renal, no está totalmente demostrado esa variabilidad, por mucho que sea cierto que la enfermedad renal crónica afecta más a hombres que a mujeres.

A partir de los gráficos que se incluyen a continuación, puede observarse que hay una observación anómala u outlier, que sigue patrones distintos al resto de los individuos ya que de acuerdo al modelo probit es el individuo con mayor probabilidad predicha de padecer enfermedad crónica renal, pero sin embargo está sano. Los datos correspondientes a este individuo son los que se muestran en la tabla 106.

Tabla 106: Datos del individuo con la HBA1C errónea

id	Leucocitos	V1HORA	COL	GOT	GPT	GGT	HBA1C	FGR_	Sexo
2026	8.42	4	169	20	21	41	51	sano(>=60)	hombre

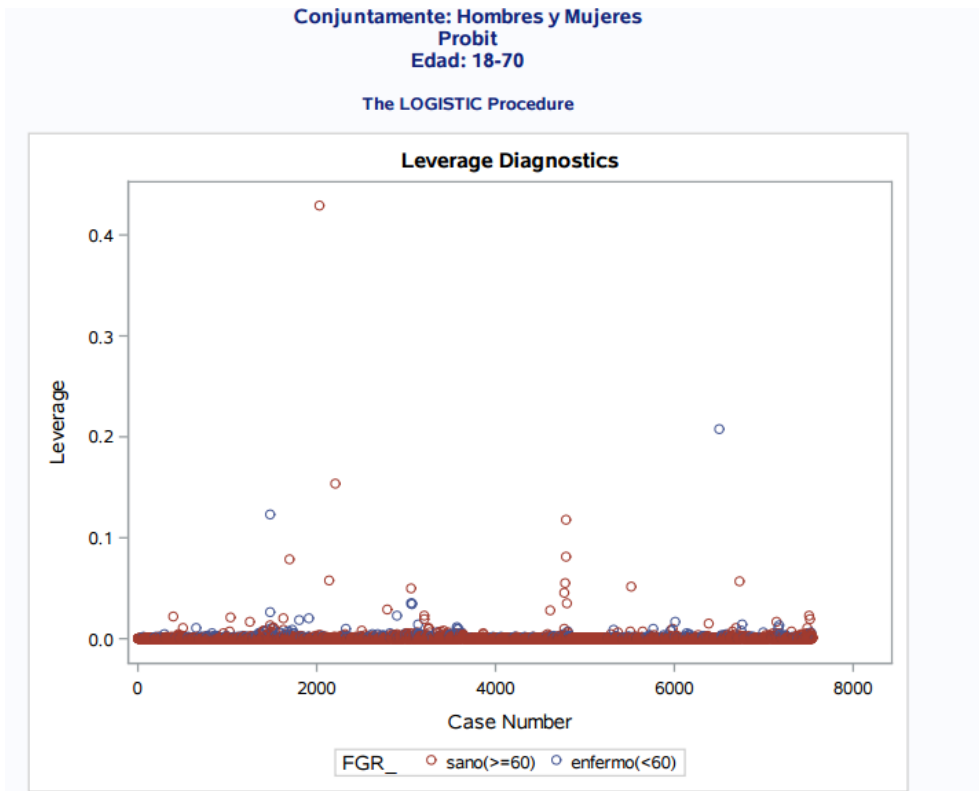


Figura 14: Influencia de cada dato en los valores estimados por el modelo

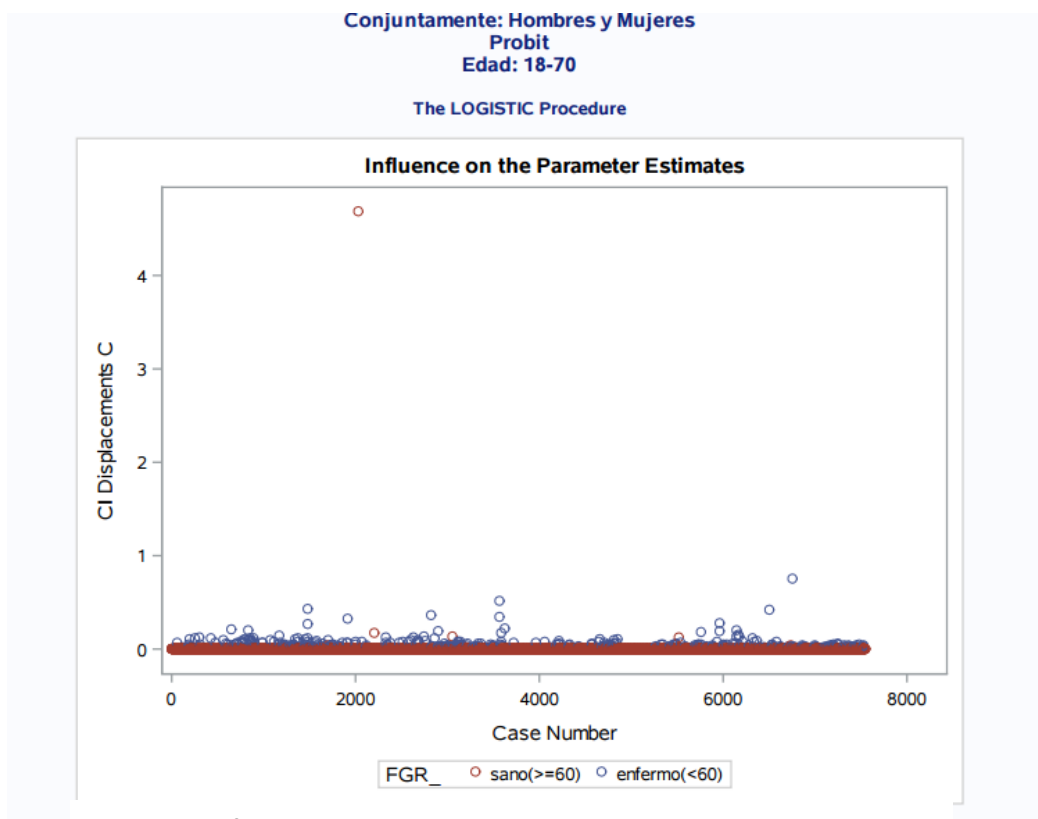


Figura 15: Influencia de cada dato en los valores estimados por el modelo.

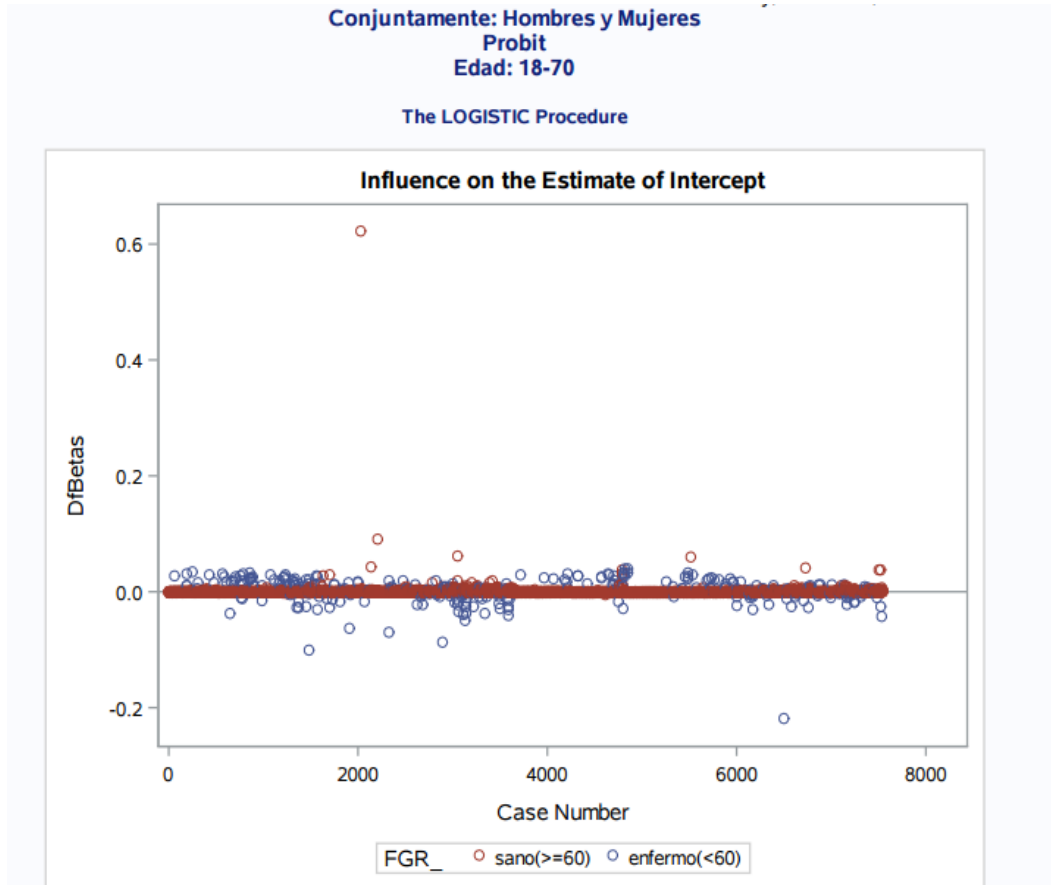


Figura 16: Influencia de cada dato en los valores estimados por el modelo.

Las figuras 14,15 y 16 muestran de varias maneras la influencia que tienen los distintos de la muestra analizada en las estimaciones de los parametros del modelo y podemos observar en todas ellas la influencia que tiene la observación anómala, por el error de HbA1C, que se ha medido como 51%, cuando la realidad es que debería ser 5,1%, lo que claramente es un error de transcripción en la BBDD del laboratorio. Es un dato que influye mucho en las estimaciones, y que no cuadra con el modelo por lo que intenta cambiar el modelo para ajustarlo mejor, aunque eso implique llegar a tener más error en el resto de los datos. Esto nos permite determinar que el modelo elegido es lo suficientemente sensible para detectar cualquier anomalía que se produzca en los datos.

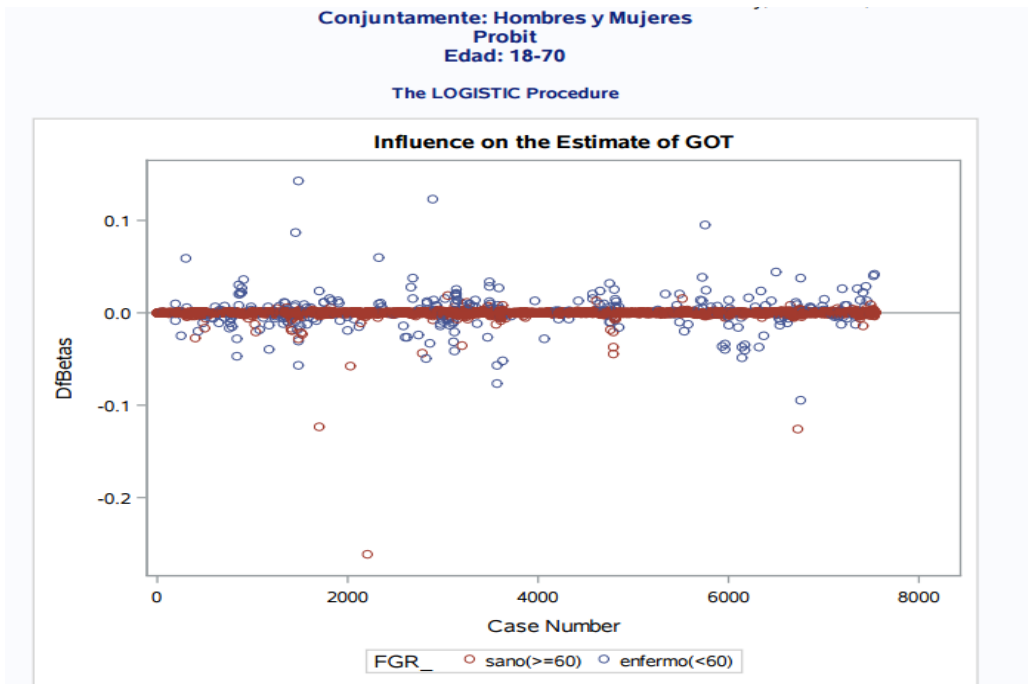


Figura 17: Influencia de GOT en los datos.

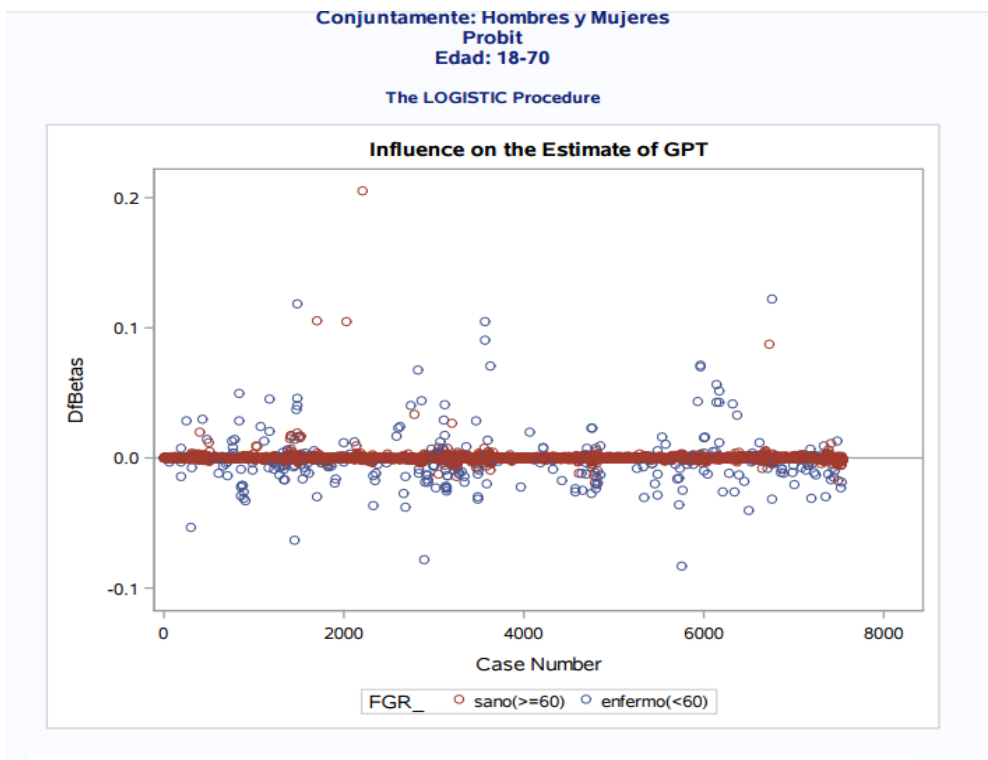


Figura 18: Influencia de GPT en los datos.

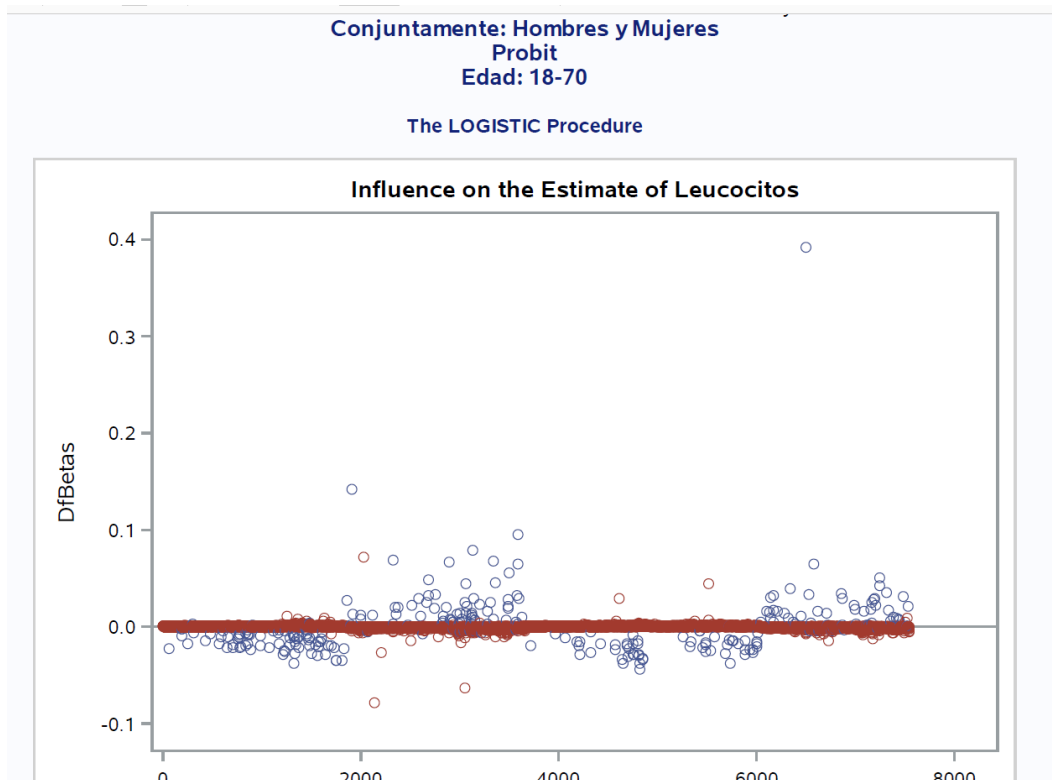


Figura 19: Influencia de Leucocitos en los datos.

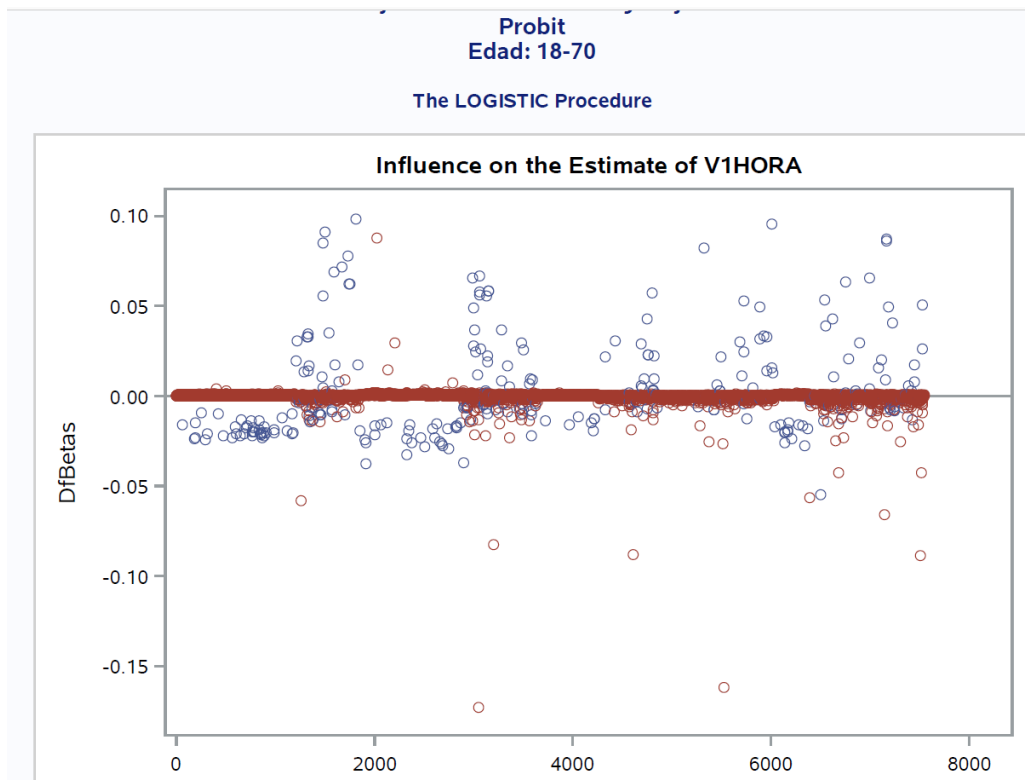


Figura 20: Influencia de V1HORA en los datos.

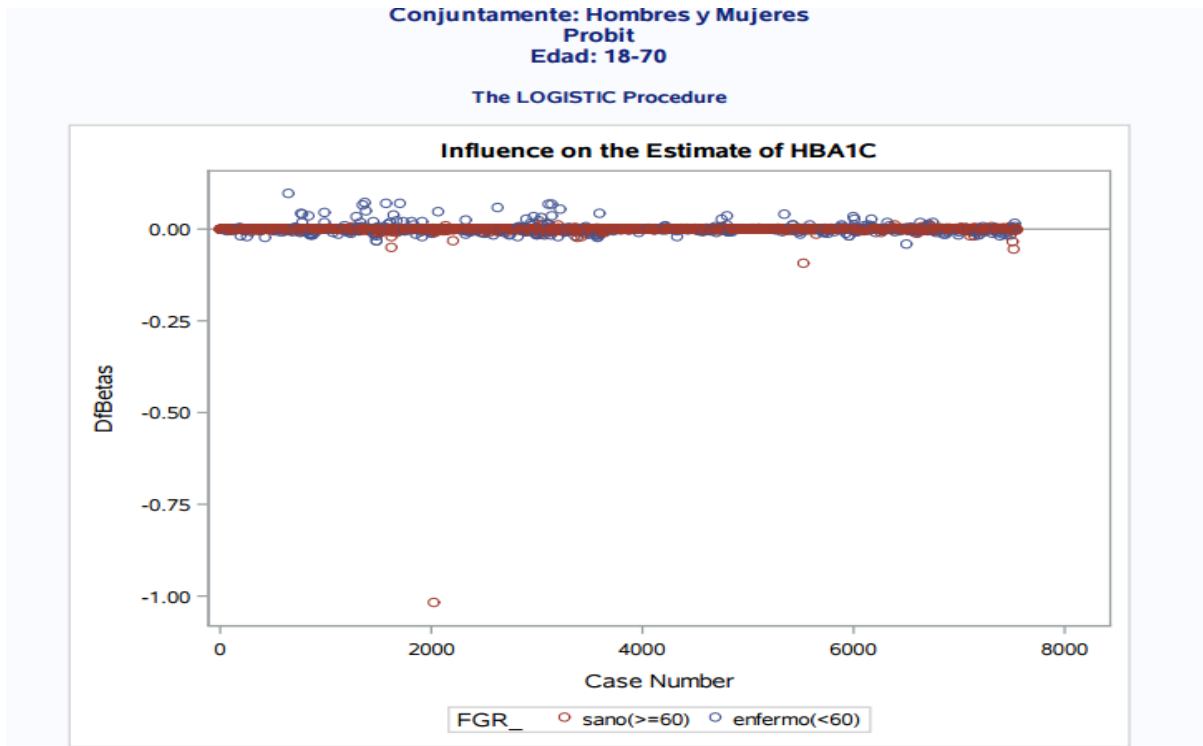


Figura 21: Influencia de HbA1C en los datos.

Las figuras 17, 18, 19, 20 y 21 muestran como influyen las variables explicativas en los datos, en el eje horizontal aparecen todos los datos analizados, y en el eje vertical los valores de la variable explicativa correspondiente. Los valores en rojo corresponden a los individuos que están tienen un valor de FGR mayor de 60 y los que están en azul son los individuos con un valor de FGR menor de 60. A partir de la gráfica se puede observar que en general los individuos sanos tienen valores cercanos a 0 y los individuos enfermos tienen valores que en general están más separados de 0. En la gráfica de la hemoglobina glicosilada (HbA1C), se observa que el gráfico es más pequeño, pero es debido a que hay un individuo con un FGR mayor de 60, pero tiene un valor extraordinariamente negativo, lo que afecta a la escala de la gráfica.

En cuanto a la capacidad de predicción del modelo probit(FGR_) basado en las variables Leucocitos, V1HORA, GOT, GPT, HBA1C, medida mediante el área de la curva ROC es de 0.66, lo cual es un valor aceptable.

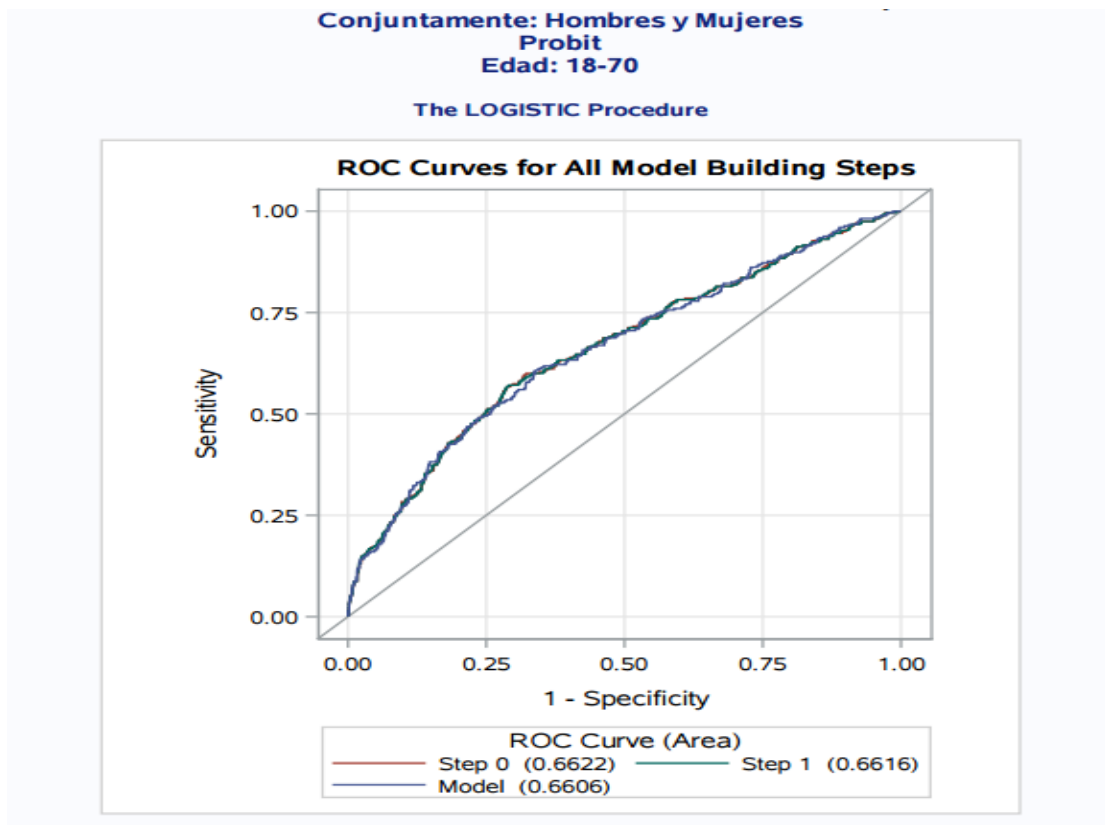


Figura 20: Área bajo la curva.

La figura 20 es la curva ROC o curva de rendimiento diagnóstico. La sensibilidad es el porcentaje de verdaderos positivos y 1-especificidad es la proporción de falsos positivos. Si el modelo es muy malo, acierta al azar (la mitad de las veces) y entonces obtenemos la diagonal. Si el modelo fuese perfecto, entonces no hay falsos negativos, es decir, la sensibilidad sería 100% (no habría falsos negativos) y 1-especificidad sería 0% (no habría falsos positivos). Esto daría el valor (1,0). En nuestro caso se va individuo a individuo, de forma que se va viendo la sensibilidad y la 1-especificidad que se va acumulando, de esta forma en el modelo perfecto sería una línea que toma el valor 1 siempre (el techo de la gráfica). Lo que se hace para ver si el modelo es bueno es calcular el área bajo la curva. En el modelo inútil este área vale 0.5 y en el modelo perfecto vale 1 (es un cuadrado). Cuanto más grande sea el valor mejor será el modelo. En nuestro caso está sobre 0.66, que es un valor bastante aceptable.

La figura 21 nos muestra los datos de sensibilidad y especificidad (Argimon JM et al, 2000). La sensibilidad es la proporción de individuos enfermos (con FGR menor de 60) que serían clasificados como enfermos; y la especificidad es la proporción de individuos sanos (con FGR mayor de 60) que son clasificados como tales.

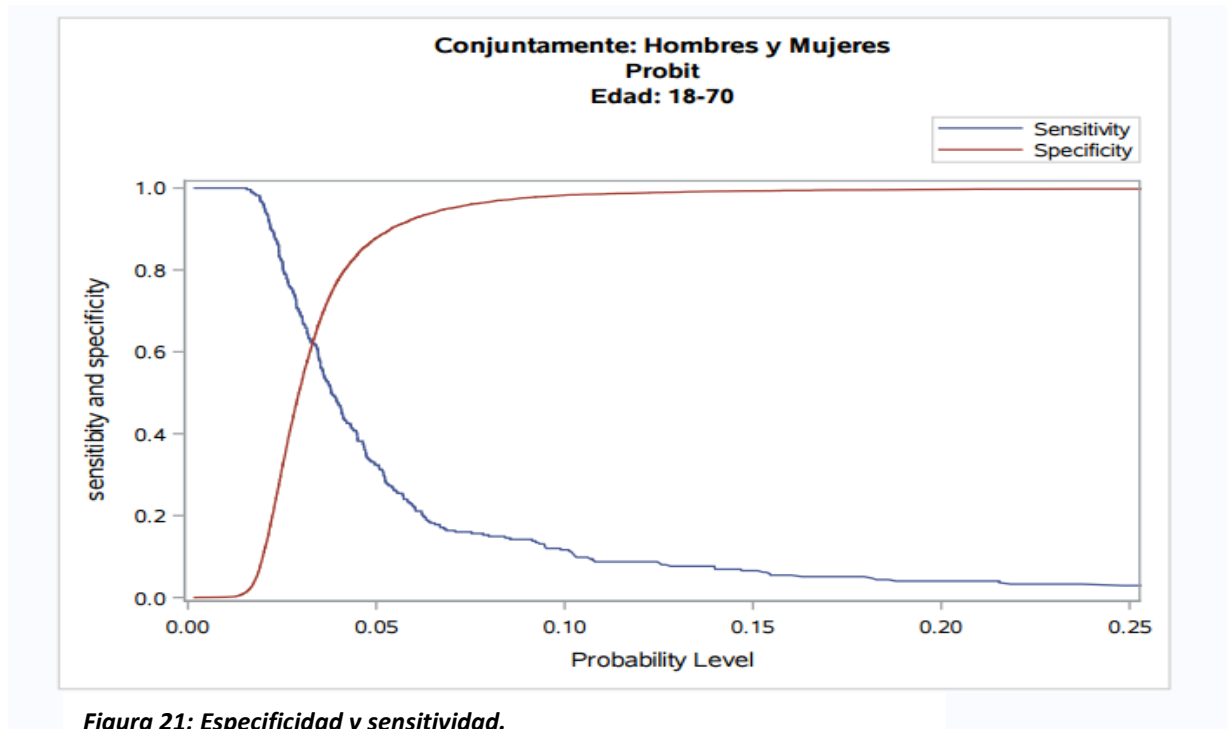


Figura 21: Especificidad y sensibilidad.

Dado un individuo, este individuo puede estar sano o enfermo. De este individuo tenemos unos datos, que nos lleva a un valor de probabilidad a partir de nuestro modelo, sin embargo, es posible equivocarse. La curva azul nos dice qué porcentaje de individuos enfermos se clasificarían mal en función de la probabilidad que obtiene nuestro modelo. Cuando es una probabilidad muy pequeña, entonces es muy probable que lo clasifiquemos mal si está enfermo. Y esta probabilidad va disminuyendo cuando va aumentando la probabilidad. Con la especificidad va al revés y por eso sale así la curva roja; nos dice qué porcentaje de individuos sanos se clasificarían mal. Cuando es una probabilidad muy pequeña, entonces es muy probable que lo clasifiquemos bien. Y esta probabilidad va aumentando cuando va aumentando la probabilidad. Las dos curvas se cortan en el valor de probabilidad 0.033, lo que nos dice el punto a partir del cual hay que clasificar un individuo como sano o enfermo.

6.- CONCLUSIONES

1. Se ha encontrado un modelo, que es capaz de explicar o predecir la variabilidad del filtrado glomerular, en función de la variabilidad de una serie de analitos (variables explicativas) como son: Leucocitos, Velocidad de sedimentación, Hemoglobina glicosilada, glutamato oxalacetato transaminasa (GOT), y Glutamato piruvato transaminasa. Habiendo descartado el modelo los analitos de Colesterol y Gamma glutamil transpeptidasa.
2. El modelo es válido cuando se estudia sobre una población con edades comprendidas entre los 18 y los 70 años, lo cual reafirma que el estudio del filtrado glomerular no es recomendable en pacientes fuera de ese rango de edad.
3. El modelo no se ajusta cuando analizamos únicamente población de mujeres, lo cual abre un campo de estudio para analizar si las hormonas femeninas juegan un papel importante en la prevención del desarrollo de enfermedad renal crónica en mujeres.
4. El modelo propuesto como válido, lo es en una población del Corredor del Henares, pero no tiene por qué serlo en otra población, puesto que podrían no ser homogéneas.
5. La población estudiada tiene un promedio de las variables explicativas, que se encuentra dentro de la normalidad según los valores de referencia, por lo que es una población perfectamente válida para el estudio.
6. Dada la influencia que tiene la observación anómala o outlier, podría ser recomendable repetir el estudio una vez corregido el valor erróneo, pero siendo una única muestra frente a las casi 10.000 del estudio, no creemos que los resultados fueran a variar significativamente.
7. Podría ser interesante en un futuro estudio incluir como variable explicativa la fracción de colesterol LDL, en vez de la variable colesterol total, por la importancia que tiene dicha fracción del colesterol como factor de riesgo para desarrollar una enfermedad cardiovascular.

8. Como conclusión final podemos decir, que dada la importancia que tiene la detección precoz de una posible enfermedad renal, para la optimización del tratamiento de la misma, los resultados de este trabajo muestran que la utilización de modelos matemáticos basados en el análisis de variables explicativas de fácil acceso, pueden ser de utilidad para predecir la posibilidad de llegar a desarrollar una enfermedad renal crónica.

7.- BIBLIOGRAFIA

Alcázar R, Egocheaga MI, Orte L, Lobos JM, Gonzalez Parra E, Alvarez Guisasola F, et al. [SEN-SEMFYC consensus document on chronic kidney disease]. *Nefrología : publicación oficial de la Sociedad Española Nefrología*. 2008;28(3):273-82.

Argimon JM, Jiménez J. *Métodos de investigación clínica y epidemiológica*. 2ª ed Barcelona: Harcourt; 2000.

Arrieta F, Iglesias P, Pedro-Botet J, Tébar FJ, Ortega E, Nubiola A, et al. [Diabetes mellitus and cardiovascular risk: Working group recommendations of Diabetes and Cardiovascular Disease of the Spanish Society of Diabetes (SED, 2015)]. *Atención primaria*. 2016;48(5):325-36.

Bagley SC, White H, Golomb BA. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *Journal of clinical epidemiology*. 2001;54(10):979-85.

Bagshaw SM, Cruz DN, Aspromonte N, Daliento L, Ronco F, Sheinfeld G, et al. Epidemiology of cardio-renal syndromes: workgroup statements from the 7th ADQI Consensus Conference. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*. 2010;25(5):1406-16.

Bachorick PS. *Lipids and dyslipoproteinemia*. Bernard J. Clinical diagnosis and management by laboratory methods. 23rd ed. Philadelphia, PA: Elsevier Saunders; 2016

Bender R, Grouven U. Ordinal logistic regression in medical research. *Journal of the Royal College of Physicians of London*. 1997;31(5):546-51.

Bliss CI. The Method of Probits. *Science*. 1934;79:38-9.

Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, Jr., et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *Jama*. 2003;289(19):2560-72.

Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron*. 1976;16(1):31-41.

Paton D, Nuñez J, Bao D. Forage biomass of 22 shrub species from Monfragüe Natural Park (SW Spain) assessed by log-log regression models *Journal of Arid Environments*. 2002;52(2):223-31.

Dalmau Llorca MR, Boira Costa M, Lopez Pablo C, Pepio Vilaubi JM, Aguilar Martin C, Forcadell Drago E. [Differences between MDRD-4 and CG in the prevalence of renal failure and its associated variables in type2 diabetic patients]. *Atención primaria*. 2016;48(9):596-603.

Dawber TR, Moore FE, Mann GV. II. Coronary Heart Disease in the Framingham Study. *International journal of epidemiology*. 2015;44(6):1767-80.

Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Houston Miller N, Hubbard VS, et al. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology*. 2014;63(25 Pt B):2960-84.

Fagerland MW, Hosmer DW. A goodness-of-fit test for the proportional odds regression model. *Statistics in medicine*. 2013;32(13):2235-49.

Fox CS, Matsushita K, Woodward M, Bilo HJ, Chalmers J, Heerspink HJ, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. *Lancet*. 2012;380(9854):1662-73.

Fu Q, Cao L, Li H, Wang B, Li Z. Cardiorenal syndrome: pathophysiological mechanism, preclinical models, novel contributors and potential therapies. *Chinese medical journal*. 2014;127(16):3011-8.

Gnanaraj J, Radhakrishnan J. Cardio-renal syndrome. *F1000Research*. 2016;5.

Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *The New England journal of medicine*. 2004;351(13):1296-305.

Goldman L. Approach to the patient with possible cardiovascular disease. In: Goldman L, Schafer AI, eds. *Goldman-Cecil Medicine*. 25th ed. Philadelphia, PA: Elsevier Saunders; 2016:chap 51.

Gómez Carracedo A, Baztan Cortes JJ. [Renal function evaluation methods in the elderly: reliability and clinical implications]. *Revista española de geriatría y gerontología*. 2009;44(5):266-72.

Gonzalez de Buitrago JM. *Técnicas y métodos del laboratorio clínico*. 2ª ed. Salamanca, Ed Masson; 2.004

Gowans EM, Fraser CG. Biological variation of serum and urine creatinine and creatinine clearance: ramifications for interpretation of results and patient care. *Annals of clinical biochemistry*. 1988;25 (Pt 3):259-63.

Gracia S, Montanes R, Bover J, Cases A, Deulofeu R, Martin de Francisco AL, et al. Recommendations for the use of equations to estimate glomerular filtration rate in adults. *Spanish Society of Nephrology. Nefrología : publicación oficial de la Sociedad Española Nefrología*. 2006;26(6):658-65.

Harbert KR. Venipuncture. In: Dehn RW, Asprey DP, eds. *Essential Clinical Procedures*. 3rd ed. Philadelphia, PA: Elsevier Saunders; 2013:chap 5

Henny J. [Determining and verifying reference intervals in clinical laboratories]. *Annales de biologie clinique*. 2011;69(2):229-37.

Inal BB, Oguz O, Emre T, Usta M, Inal H, Altunoglu E, et al. Evaluation of MDRD, Cockcroft-Gault, and CKD-EPI formulas in the estimated glomerular filtration rate. *Clinical laboratory*. 2014;60(10):1685-94.

Inker LA, Fan L, Levey AS. Assessment of renal function. In: Johnson RJ, Feehally J, Floege J. *Comprehensive Clinical Nephrology*. 5th ed. Philadelphia, PA: Elsevier Saunders; 2015:chap 3

Kagan A, Dawber TR, Kannel WB, Revotskie N. The Framingham study: a prospective study of coronary heart disease. *Federation proceedings*. 1962;21(4)Pt 2:52-7.

Kilbride HS, Stevens PE, Eaglestone G, Knight S, Carter JL, Delaney MP, et al. Accuracy of the MDRD (Modification of Diet in Renal Disease) study and CKD-EPI (CKD Epidemiology Collaboration) equations for estimation of GFR in the elderly. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2013;61(1):57-66.

Landry DW, Bazari H. Approach to the patient with renal disease. In: Goldman L, Schafer AI, eds. *Goldman's Cecil Medicine*. 25th ed. Philadelphia, PA: Elsevier Saunders; 2016:chap 114

Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. *Modification of Diet in Renal Disease Study Group. Annals of internal medicine*. 1999;130(6):461-70.

Levey AS, Coresh J, Balk E, Kausz AT, Levin A, Steffes MW, et al. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Annals of internal medicine*. 2003;139(2):137-47.

Lorenzo-Medina M, Uranga M, Rus A, Dominguez I. Comparación entre 2 sistemas analíticos para la determinación de la hemoglobina A1c: inmunoturbidimetría versus cromatografía líquida de alta eficiencia. *Laboratorio Clínico*. 2013;6(4):145-50.

Mann JF, Gerstein HC, Pogue J, Bosch J, Yusuf S. Renal insufficiency as a predictor of cardiovascular outcomes and the impact of ramipril: the HOPE randomized trial. *Annals of internal medicine*. 2001;134(8):629-36.

Marín R. [Kidney and cardiovascular disease]. *Nefrología : publicación oficial de la Sociedad Española Nefrología*. 2004;24 Suppl(6):17-20.

Marrugat J, Elosua R, Icaza G, Morales-Salinas A, Degano IR. [Practicality of cardiovascular risk functions]. *Medwave*. 2016;16(Suppl4):e6792.

McCullagh P, Nelder J. *Generalized Linear Models*. London: Chapman and Hall, 1989

Milton J S. *Estadística para biología y ciencias de la salud*. 3.^a ed, Mc Graw-Hill, 2001

Morris M. *Blood basic exam*. Bernard J. *Clinical diagnosis and management by laboratory methods*. 23rd ed. Philadelphia, PA: Elsevier Saunders; 2016

Myers GL, Miller WG, Coresh J, Fleming J, Greenberg N, Greene T, et al. Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program. *Clinical chemistry*. 2006;52(1):5-18.

National Kidney F. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2002;39(2 Suppl 1):S1-266.

Newby DE, Grubb Nr, Bradbury A. *Cardiovascular disease*. In: Walker BR, Colledge NR, Ralston SH, Perman ID, eds. *Davidson's Principles and Practice of Medicine*. 22nd ed. Philadelphia, PA: Elsevier Churchill Livingstone; 2012:chap 18

Norte Navarro AI, Sansano Perea M, Martinez Sanz JM, Sospedra Lopez I, Hurtado Sanchez JA, Ortiz Moncada R. Estudio de factores de riesgo de enfermedad cardiovascular en trabajadores universitarios españoles. *Nutrición hospitalaria*. 2016;33(3):273.

Nuñez J, Minana G, Santas E, Bertomeu-Gonzalez V. Cardiorenal Syndrome in Acute Heart Failure: Revisiting Paradigms. *Revista española de cardiología*. 2015;68(5):426-35.

Otero A, Gayoso P, Garcia F, de Francisco AL, group Es. Epidemiology of chronic renal disease in the Galician population: results of the pilot Spanish EPIRCE study. *Kidney international Supplement*. 2005(99):S16-9.

Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. *Statistics in medicine*. 2013;32(1):67-80.

Payne RB. Creatinine clearance and glomerular filtration rate. *Annals of clinical biochemistry*. 2000;37 (Pt 1):98-9.

Pourhoseingholi A, Pourhoseingholi MA, Rostami-Nejad M, Rostami K, Mirsatari D, Zojaji H, et al. Implementation of statistical analysis in the clinical research of coeliac disease: use of probit and logit analysis. *East African journal of public health*. 2010;7(2):168-70.

Ronco C, McCullough P, Anker SD, Anand I, Aspromonte N, Bagshaw SM, et al. Cardio-renal syndromes: report from the consensus conference of the acute dialysis quality initiative. *European heart journal*. 2010;31(6):703-11.

Ruilope LM, Salvetti A, Jamerson K, Hansson L, Warnold I, Wedel H, et al. Renal function and intensive lowering of blood pressure in hypertensive participants of the hypertension optimal treatment (HOT) study. *Journal of the American Society of Nephrology : JASN*. 2001;12(2):218-25.

Ryoo JH, Kim SG, Suh BS, Kim DI, Park SK. Relationship between chronic kidney disease and risk of coronary heart disease in Korean men. *Journal of Korean medical science*. 2011;26(6):753-8.

Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Kirkman MS, et al. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Diabetes care*. 2011;34(6):e61-99.

Santamaria Olmo R, Gorostidi Perez, M. Síndrome Cardiorenal. *Nefrología al día*. 2012.

Sarnak MJ, Levey AS. Cardiovascular disease and chronic renal disease: a new paradigm. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2000;35(4 Suppl 1):S117-31.

Schumann G, Bonora R, Ceriotti F, Ferard G, Ferrero CA, Franck PF, et al. IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 degrees C. International Federation of Clinical Chemistry and Laboratory Medicine. Part 5. Reference procedure for the measurement of catalytic concentration of aspartate aminotransferase. *Clinical chemistry and laboratory medicine*. 2002;40(7):725-33.

Schumann G, Bonora R, Ceriotti F, Ferard G, Ferrero CA, Franck PF, et al. IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 degrees C. International Federation of Clinical Chemistry and Laboratory Medicine. Part 4. Reference procedure for the measurement of catalytic concentration of alanine aminotransferase. *Clinical chemistry and laboratory medicine*. 2002;40(7):718-24.

Schumann G, Bonora R, Ceriotti F, Ferard G, Ferrero CA, Franck PF, et al. IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 degrees C. International Federation of Clinical Chemistry and Laboratory Medicine. Part 6. Reference procedure for the measurement of catalytic concentration of gamma-glutamyltransferase. *Clinical chemistry and laboratory medicine*. 2002;40(7):734-8.

Seppi T, Prajczek S, Dorler MM, Eiter O, Hekl D, Nevinny-Stickel M, et al. Sex Differences in Renal Proximal Tubular Cell Homeostasis. *Journal of the American Society of Nephrology : JASN*. 2016;27(10):3051-62.

Siest G, Henny J, Grasbeck R, Wilding P, Petitclerc C, Queralto JM, et al. The theory of reference values: an unfinished symphony. *Clinical chemistry and laboratory medicine*. 2013;51(1):47-64.

Siu AL, Force USPST. Screening for high blood pressure in adults: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine*. 2015;163(10):778-86.

Soares AA, Eyff TF, Campani RB, Ritter L, Camargo JL, Silveiro SP. Glomerular filtration rate measurement and prediction equations. *Clinical chemistry and laboratory medicine*. 2009;47(9):1023-32.

Soriano Cabrera S, Sociedad Española de N. [Definition, classification of stages, and prevalence of chronic kidney disease. Guidelines for early detection. Risk factors for chronic kidney disease]. *Nefrología : publicación oficial de la Sociedad Española Nefrología*. 2004;24 Suppl(6):27-34, 187-235.

Stevens LA, Coresh J, Feldman HI, Greene T, Lash JP, Nelson RG, et al. Evaluation of the modification of diet in renal disease study equation in a large diverse population. *Journal of the American Society of Nephrology : JASN*. 2007;18(10):2749-57.

Swan SK. The search continues--an ideal marker of GFR. *Clinical chemistry*. 1997;43(6 Pt 1):913-4.

Swedko PJ, Clark HD, Paramsothy K, Akbari A. Serum creatinine is an inadequate screening test for renal failure in elderly patients. *Archives of internal medicine*. 2003;163(3):356-60.

The HOPE (Heart Outcomes Prevention Evaluation) Study: the design of a large, simple randomized trial of an angiotensin-converting enzyme inhibitor (ramipril) and vitamin E in patients at high risk of cardiovascular events. The HOPE study investigators. *The Canadian journal of cardiology*. 1996;12(2):127-37.

The Hypertension Optimal Treatment Study (the HOT Study). *Blood pressure*. 1993;2(1):62-8.

The Modification of Diet in Renal Disease Study: design, methods, and results from the feasibility study. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 1992;20(1):18-33.

Thomas B, Matsushita K, Abate KH, Al-Aly Z, Arnlov J, Asayama K, et al. Global Cardiovascular and Renal Outcomes of Reduced GFR. *Journal of the American Society of Nephrology : JASN*. 2017.

Tonelli M, Wiebe N, Culeton B, House A, Rabbat C, Fok M, et al. Chronic kidney disease and mortality risk: a systematic review. *Journal of the American Society of Nephrology : JASN*. 2006;17(7):2034-47.

Toth PP, Shamma NW, Foreman Bj, Byrd JB. Cardiovascular disease. In: Rakel RE, Rakel D, eds. *Textbook of Family Medicine*. 9th ed. Philadelphia, PA: Elsevier Saunders; 2016:chap 27

Verschuren WM, Jacobs DR, Bloemberg BP, Kromhout D, Menotti A, Aravanis C, et al. Serum total cholesterol and long-term coronary heart disease mortality in different cultures. Twenty-five-year follow-up of the seven countries study. *Jama*. 1995;274(2):131-6.

Victor RG. Arterial hypertension. In: Goldman L, Schafer AI, eds. *Goldman's Cecil Medicine*. 25th ed. Philadelphia, PA: Elsevier Saunders; 2016:chap 67

Wayne P. Defining, establishing and verifying reference intervals in the clinical laboratory; approved guideline. CLSI Document C28-A3. 2008.

Weiner DE, Tighiouart H, Stark PC, Amin MG, MacLeod B, Griffith JL, et al. Kidney disease as a risk factor for recurrent cardiovascular disease and mortality. American journal of kidney diseases : the official journal of the National Kidney Foundation. 2004;44(2):198-206.