

# Facultad de Ciencias

## Departamento de Estadística

Carrera Ingeniería Estadística

# PROYECTO DE TÍTULO II

- Asignatura : Proyecto de Título II (220061)
- Título : Análisis y Modelo de Predicción para el Riesgo Cardiovascular Global en pacientes con Artritis Reumatoide considerando Biomarcadores de Inflamación Celular y Disfunción Endotelial.
- Realizado por : Luciano Toledo Conejeros.  
Ingeniería Estadística.
- Profesor Guía : Christian Caamaño Carrillo.  
Departamento de Estadística.
- Patrocinante : Dra. Liliana Lamperti Fernández.  
Departamento de Bioquímica Clínica e Inmunología,  
Facultad de Farmacia, Universidad de Concepción.
- Semestre : Primer semestre 2012.
- Fecha : Junio 2012.



*Versión N° 1*

UNIVERSIDAD DEL BÍO BÍO  
[www.ubiobio.cl](http://www.ubiobio.cl)

Universidad del Bío-Bío  
Facultad de Ciencias  
Departamento de Estadística



UNIVERSIDAD DEL BÍO-BÍO

**“ANÁLISIS Y MODELO DE PREDICCIÓN PARA EL RIESGO  
CARDIOVASCULAR GLOBAL EN PACIENTES CON ARTRITIS  
REUMATOIDE CONSIDERANDO BIOMARCADORES DE INFLAMACIÓN  
CELULAR Y DISFUNCIÓN ENDOTELIAL”**

Luciano Alejandro Toledo Conejeros  
***Estudiante de Ingeniería Estadística***

Christian Caamaño Carrillo  
***Profesor Guía***  
Departamento de Estadística  
Facultad de Ciencias  
Universidad del Bío-Bío

Liliana Lamperti Fernández  
***Patrocinante***  
Departamento de Bioquímica Clínica e Inmunología  
Facultad de Farmacia  
Universidad de Concepción

Concepción, Chile  
2012

# 1. Índice General

1. ÍNDICE GENERAL	3
1.1 ÍNDICE DE FIGURAS	4
1.2 ÍNDICE DE TABLAS	5
2. RESUMEN	7
3. INTRODUCCIÓN	8
4. ORIGEN DE TEMA	11
5. JUSTIFICACIÓN	11
6. OBJETIVO DEL ESTUDIO	12
6.1 OBJETIVO GENERAL	12
6.2 OBJETIVOS ESPECÍFICOS	12
7. ALCANCES	14
8. METODOLOGÍA	16
9. ANTECEDENTES GENERALES	27
10. RESULTADOS	28
11. CONCLUSIÓN	91
12. BIBLIOGRAFÍA	95
13. ANEXOS	96

## 1.1 Índice de Figuras

1. Descripción de Artritis Reumatoide.	8
2. Descripción de una placa aterosclerótica dañada.	8
3. Evolución de una placa aterosclerótica.	9
4. Factores de riesgo cardiovascular.	9
5. Estimación del riesgo cardiovascular global.	10
6. Análisis de Correspondencia múltiple para todas las variables de acuerdo a la condición AR.	29
7. Análisis de correspondencia múltiple para variables que dieron dependencia con la variable Condición AR.	32
8. Análisis de Correspondencia múltiple para todas las variables de acuerdo al Tipo de RCV Global.	34
9. Análisis de correspondencia múltiple para variables que dieron dependencia con el Tipo de RCV Global.	37
10. Curva ROC para el biomarcador ROS.	40
11. Análisis de Correspondencia múltiple para todos los biomarcadores de acuerdo a la condición AR.	43
12. Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con la condición AR de los pacientes .	45
13. Análisis de Correspondencia múltiple para todos los biomarcadores de acuerdo al Tipo de RCV Global.	47
14. Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con el Tipo de RCV Global de los pacientes.	49
15. Herramienta interactiva original para un modelo de predicción logístico.	86
16. Imagen 1 de herramienta interactiva para un modelo de predicción logístico para la condición AR de los pacientes.	87
17. Imagen 2 de herramienta interactiva para un modelo de predicción logístico para la condición AR de los pacientes.	87
18. Imagen 1 de herramienta interactiva para un modelo de predicción logístico para el tipo de RCG de los pacientes.	88
19. Imagen 2 de herramienta interactiva para un modelo de predicción logístico para el tipo de RCG de los pacientes.	88

## 1.2 Índice de Tablas

1. Clasificación del tamaño de la muestra.	15
2. Categorización de todas las variables.	18
3. ANOVA para un Diseño en Bloques Aleatorizados.	22
4. Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todas las variables con respecto a la condición AR.	30
5. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y la condición AR de los pacientes.	31
6. Medidas de discriminación de pesos del Análisis de correspondencia múltiple para variables que dieron dependencia con la variable Condición AR.	33
7. Perfil de todas las variables asociadas con la condición AR de los pacientes.	33
8. Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todas las variables con respecto al Tipo de RCV Global.	38
9. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el Tipo de riesgo cardiovascular global de los pacientes.	36
10. Medidas de discriminación de pesos del Análisis de correspondencia múltiple para variables que dieron dependencia con la variable tipo de RCV Global.	38
11. Perfil de todas las variables asociadas con el Tipo de RCG de los pacientes.	39
12. Valores de corte para el biomarcador ROS.	42
13. Categorización de todos los biomarcadores.	43
14. Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todos los biomarcadores con respecto a la condición AR.	44
15. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada biomarcador y la condición AR de los pacientes.	44
16. Medidas de discriminación de pesos del Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con la variable Condición AR.	46
17. Perfil de todos los biomarcadores asociados con la condición AR de los pacientes.	46
18. Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todos los biomarcadores con respecto al Tipo de RCV Global.	48
19. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada biomarcador y el Tipo de riesgo cardiovascular global de los pacientes.	48

20. Medidas de discriminación de pesos del Análisis de correspondencia múltiple. para biomarcadores que dieron dependencia con la variable tipo de RCG .	50
21. Perfil de todos los biomarcadores asociados con el Tipo de RCV Global de los pacientes.	50
22. Contraste de perfiles con respecto a variables.	51
23. Contraste de perfiles con respecto a biomarcadores.	52
24. Contraste de perfiles combinados con respecto a variables y biomarcadores.	53
25. Primera ANOVA para la comparación del RCV Global.	55
26. ANOVA final para la comparación del RCV Global.	56
27. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el biomarcador TNFa.	58
28. ANOVA para la comparación del TNFa.	59
29. Simplificación de ANOVA para la comparación del TNFa.	60
30. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el biomarcador PCR.	61
31. ANOVA para la comparación de la PCR.	62
32. Primera simplificación de ANOVA para la comparación de la PCR.	63
33. Segunda simplificación de ANOVA para la comparación de la PCR.	63
34. ANOVA final para la comparación de la PCR.	64
35. Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el biomarcador ROS.	65
36. ANOVA para la comparación del ROS.	66
37. Simplificación de ANOVA para la comparación de la ROS.	67
38. Estadísticos descriptivos del modelo logístico de prueba.	69
39. Parámetros estimados del modelo logístico de prueba.	70
40. Análisis de desviación del modelo logístico de prueba.	70
41. Medidas de bondad y ajuste del modelo logístico de prueba.	71
42. Prueba de Hosmer y Lemeshow del modelo logístico de prueba.	71
43. Significación de variables en modelo logístico de prueba.	75
44. Modelo logístico para la condición AR de los pacientes.	76
45. Variables dummy para la obesidad central.	76
46. Variables dummy para la hipertensión y el tratamiento antihipertensivo.	77
47. Modelo logístico para el tipo de RCG de los pacientes.	81
48. Variables dummy para la hipertrigliceridemia.	82

## 2. Resumen

La Artritis Reumatoide (AR) es una enfermedad sistémica autoinmune, que afecta a múltiples articulaciones. Las comorbilidades asociadas a la AR toman importancia durante el curso de la enfermedad, siendo las enfermedades cardiovasculares las más prevalentes en estos pacientes y son la primera causa de muerte en los individuos que padecen de AR. En pacientes con AR, la estimación del Riesgo Cardiovascular (RC) es poco utilizada; sin embargo, el cálculo del RC, aplicando las tablas de Framingham y la determinación de nuevos biomarcadores, permitirían estimar el Riesgo Cardiovascular Global (RCG). Este estimador otorgaría mayor precisión sobre el estado proaterogénico de los pacientes con AR, clasificándolos de acuerdo a su riesgo cardiovascular en forma estratificada desde riesgo bajo, moderado y alto. El objetivo del estudio es analizar el riesgo cardiovascular global y proponer un modelo de predicción en pacientes con artritis reumatoide y sujetos controles considerando biomarcadores de inflamación celular y disfunción endotelial. Para alcanzar este objetivo se reclutaron pacientes con AR en el Hospital Regional de Concepción. Ellos fueron caracterizados en cuanto a la actividad de su enfermedad y tratamiento farmacológico. Sumado a esto se midieron biomarcadores de disfunción endotelial, de inflamación celular y marcadores bioquímicos clásicos de enfermedad cardiovascular. Hacer una nueva estratificación de los pacientes con AR, en cuanto a su RC Global, puede posicionar a los pacientes con bajo riesgo en un riesgo cardiovascular muy alto, siendo esta información importante para aplicar nuevos enfoques terapéuticos en los pacientes con AR, en que se considere de forma integrada la enfermedad cardiovascular y permitan prevenir un evento cardiovascular fatal.

Los resultados de los análisis expuestos en este informe muestran que hay ciertas variables a considerar para mejorar la condición de riesgo de los pacientes, sin embargo a la hora de comparar entre el grupo de pacientes con AR y sujetos controles se demuestra que si hay diferencias entre ellos en cuanto al valor numérico de su riesgo cardiovascular global y además en cuanto al valor numérico de los biomarcadores que se expusieron en este trabajo. Con respecto a los modelos de regresión logística se puede apreciar que serán de gran utilidad para predecir la probabilidad de un alto RC Global en los pacientes. Estos resultados entregarán información relevante para los Servicios de Salud y en especial como apoyo a las guías clínicas para un manejo y seguimiento más eficiente de los pacientes con Artritis Reumatoide.

### 3. Introducción

La Artritis Reumatoide (AR) es una enfermedad autoinmune de etiología desconocida que se caracteriza por una inflamación sistémica, daño y deformaciones articulares irreversibles. Produce daño irreversible al cartílago y tejido óseo subyacente, manifestándose con deformidades



Figura 1: Descripción de Artritis Reumatoide.

y rigidez de las articulaciones durante el progreso de la enfermedad cardiovascular, deteriorando la calidad de vida de los pacientes (Figura 1). La enfermedad se caracteriza por un estado inflamatorio de carácter sistémico, por lo tanto estos pacientes tienen serias complicaciones debido a las comorbilidades asociadas. (1)

La Enfermedad Cardiovascular (ECV) es una de las principales comorbilidades de país. Considerando las características de la enfermedad, en cuanto a las citoquinas proinflamatorias que se producen localmente en las articulaciones afectadas y que se liberan a circulación, se ha demostrado una progresión de la aterosclerosis más acelerada en estos pacientes, y consecuentemente mayor riesgo de muerte por ECV (Figura 2 y Figura 3). (2)

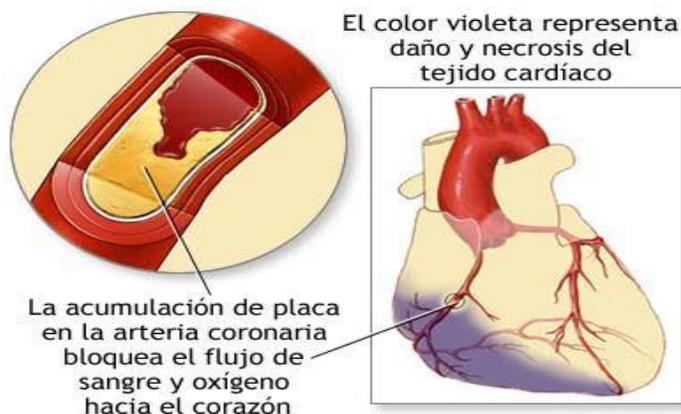


Figura 2: Descripción de una placa aterosclerótica dañada.

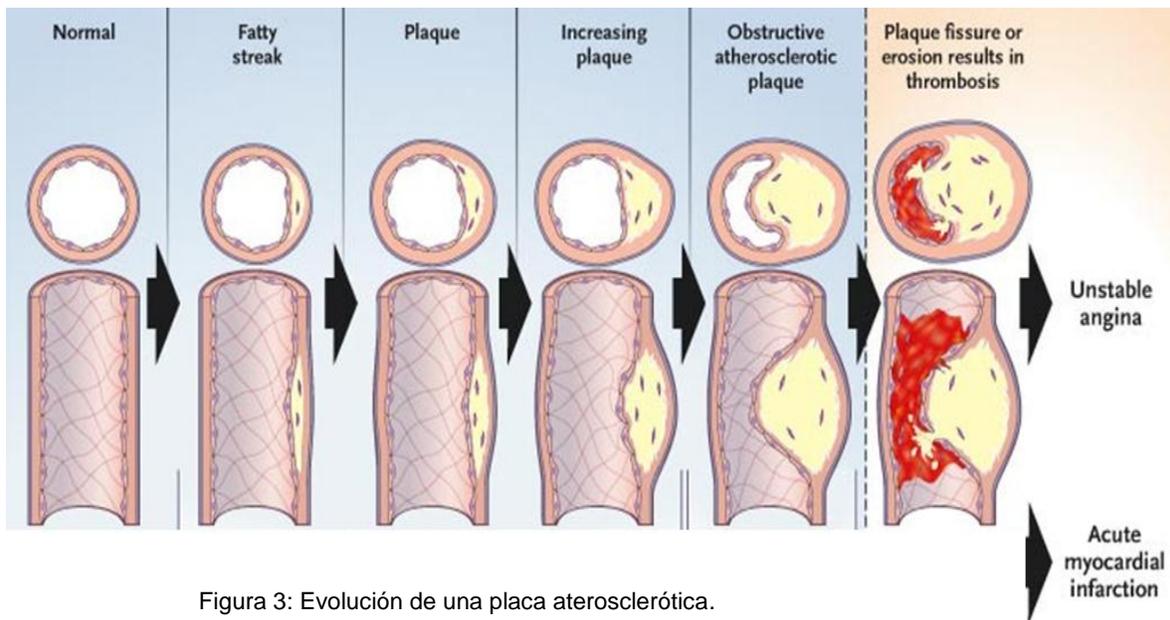


Figura 3: Evolución de una placa aterosclerótica.

Los factores de riesgo clásicos para ECV, como los que se describen en la Figura 4, son comunes de encontrarlos en los pacientes con AR y serían en parte responsables del mayor riesgo de ECV; sin embargo, la determinación de nuevos **biomarcadores** de riesgo no clásico de ECV, permitirían una mejor estimación del verdadero riesgo de desarrollar un evento cardiovascular en los pacientes con AR. (2)

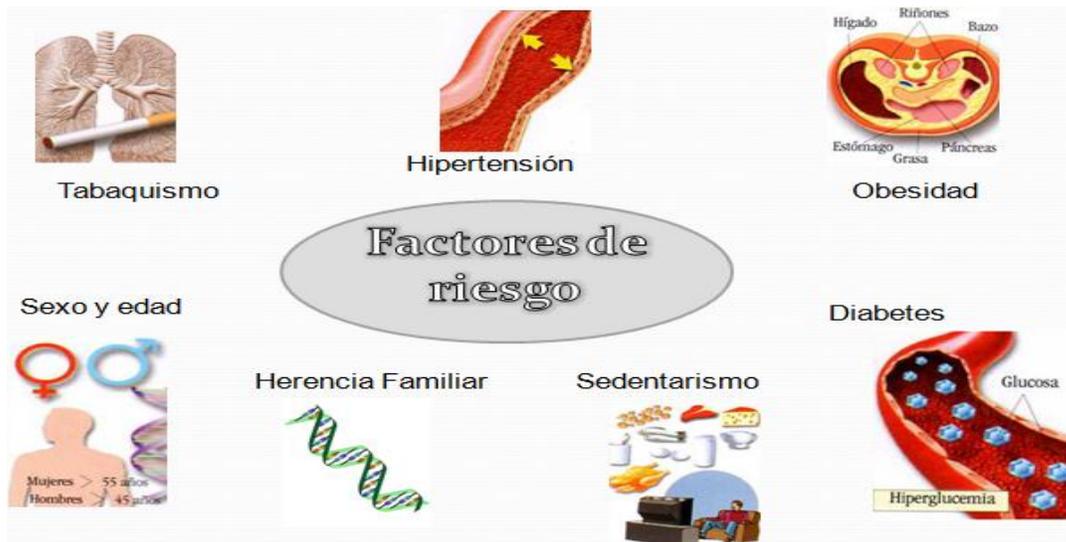


Figura 4: Factores de riesgo cardiovascular.

Durante el tratamiento y seguimiento de la enfermedad, y teniendo en cuenta el elevado riesgo de estos pacientes, se hace necesario evaluar en conjunto las características clínicas y bioquímicas propias de la AR y el Riesgo Cardiovascular Global (RCG).

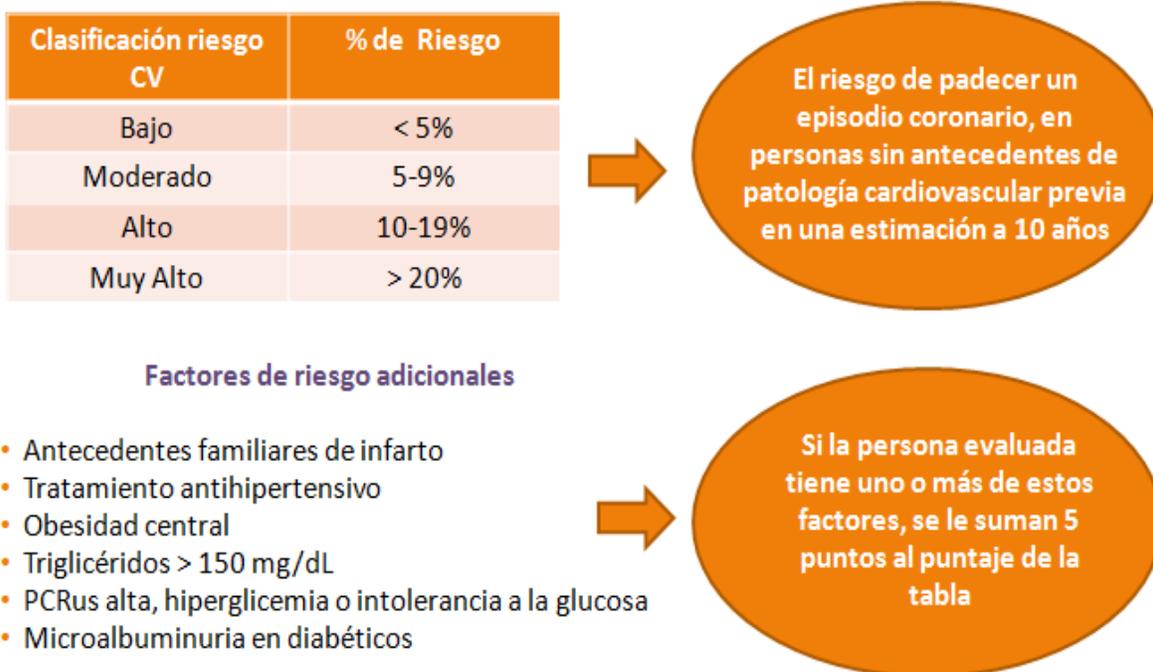


Figura 5: Estimación del riesgo cardiovascular global.

Es por ello que este estudio propone caracterizar una población de pacientes con Artritis Reumatoide en base a su Riesgo Cardiovascular Global, a través de la medición de biomarcadores de inflamación celular y de disfunción endotelial. Se sumarán los marcadores bioquímicos clásicos y antecedentes familiares de enfermedad cardiovascular de los pacientes con AR. De este modo, se podrá estimar el riesgo de presentar un evento cardiovascular a 10 años, utilizando las tablas de Framingham (Ver primera parte de este proyecto de título), adaptadas a la población chilena y entonces, estratificar a los pacientes de acuerdo a la puntuación obtenida en pacientes con riesgo bajo, moderado, alto y muy alto.

## 4. Origen del tema

Actualmente la Enfermedad Cardiovascular es una de las principales comorbilidades de los pacientes con Artritis Reumatoide, siendo la primera causa de muerte en Chile y una de las más importantes a nivel mundial.

En particular hace años el laboratorio del Grupo de Investigación en Función Endotelial (GIFE) del departamento de Bioquímica Clínica e Inmunología de Facultad de Farmacia de la Universidad de Concepción, ha realizado una serie de investigaciones y trabajos relacionadas con biomarcadores tempranos de disfunción endotelial.

## 5. Justificación

El Ministerio de Salud en Chile (MINSAL) realizó en el año 1996 un estudio(1, 3), sobre el impacto de diferentes enfermedades a nivel de la población general, considerando varios aspectos, y además empleando el método de estimación AVISA (años de vida ajustados por discapacidad). En dicho estudio, la AR se encuentra dentro de las 15 patologías que poseen una mayor carga de enfermedad. Por esto, a partir del año 2007, la AR se incluyó dentro de las patologías más relevantes y a las vez preocupantes para el MINSAL.

## 6. Objetivos

### 6.1 Objetivo General.

**Analizar el riesgo cardiovascular global y proponer un modelo de predicción en pacientes con artritis reumatoide y sujetos controles considerando biomarcadores de inflamación celular y disfunción endotelial.**

### 6.2 Objetivos Específicos.

Este estudio está conformado por dos partes, en donde en la primera de ellas, Proyecto de Título I, se desarrolló todo el marco teórico de la investigación, para luego continuar con este Proyecto de Título II que acontece, con la finalidad de resolver finalmente el objetivo General en cuestión.

Los objetivos de ambas partes se muestran a continuación:

#### 1. Proyecto de Título I

A1) Comprender los conceptos teóricos y conceptuales básicos que se incluyen en el estudio.

A2) Conocer la población objetivo y el contenido de la base de datos de los pacientes y validarlo.

A3) Diseñar la metodología a utilizar en una segunda parte de este trabajo, correspondiente a los objetivos que se muestran a continuación.

## 2. Proyecto de Título II

B1) Analizar la relación de parámetros antropométricos, clínicos, bioquímicos e inmunológicos y biomarcadores de disfunción endotelial e inflamación celular de acuerdo a la condición AR de los pacientes y Tipo de Riesgo cardiovascular Global.

B2) Determinar y contrastar perfiles para cada una de las relaciones anteriores.

B3) Identificar variables y biomarcadores modificables para disminuir un evento cardiovascular a futuro en los pacientes.

B4) Comparar el riesgo cardiovascular global entre una población de pacientes con artritis reumatoide y sujetos controles, considerando los parámetros o variables en estudio.

B5) Comparar los biomarcadores de inflamación celular (TNF- $\alpha$ , y proteína C reactiva ultrasensible (PCRus)) y disfunción endotelial (especies reactivas de oxígeno (ROS)) entre una población de pacientes con artritis reumatoide y sujetos controles, considerando los parámetros o variables en estudio.

B6) Crear un modelo de predicción de regresión logística interactivo, a través de Microsoft Excel, para el Riesgo Cardiovascular Global de pacientes con artritis reumatoide y sujetos controles, considerando biomarcadores.

## 7. Alcance

### POBLACIÓN.

#### ➤ PACIENTES.

Los sujetos que participaron en este estudio fueron reclutados desde el universo de pacientes con Artritis Reumatoide (AR), atendidos en el Policlínico de Reumatología (Servicio de Medicina Interna) del Hospital Regional de Concepción, de acuerdo a los siguientes criterios:

#### *Criterios de Inclusión:*

- Paciente con diagnóstico de AR según los criterios ACR (American College Rheumatology).
- Rango etario de adulto, sin discriminación de sexo.
- Bajo tratamiento con fármacos modificadores de la enfermedad como; Metotrexato, Leflunomida, Sulfazalacina, Hidrocloroquina, según la indicación del médico en su ficha clínica.
- Con o sin presencia de Artrosis.

#### *Criterios de Exclusión:*

- Pacientes con otra patología autoinmune (excepto Síndrome de Sjögren secundario).
- Paciente con Artritis Seronegativa con características de Pelviespondilopatías (Psoriática, Espondiloartritis Anquilosante, Artritis Reactivas, etc.)
- Pacientes que presentan otras patologías inflamatorias crónicas, como asma o enfermedades inflamatorias intestinales.
- Pacientes que reciban terapias biológicas (Ej. Anti TNF- $\alpha$ ).
- Pacientes con antecedentes de cardiopatía coronaria conocida.

➤ SUJETOS CONTROL.

Los sujetos control que participaron en este estudio, fueron reclutados desde el universo de alumnos o funcionarios de la Universidad de Concepción, que aceptaron participar de forma voluntaria, firmando el consentimiento informado respectivo y de acuerdo a los siguientes criterios:

*Criterios de Inclusión:*

- Sujetos sanos no fumadores
- Edad entre 35 y 65 años, sin discriminación de sexo.
- Que no presenten un cuadro infeccioso o inflamatorio.
- Que no hayan sufrido un evento cardiovascular menor o mayor.

*Criterios de Exclusión:*

- Que presenten alguna enfermedad de la piel o alergia a los metales.
- Que sufran de hipotensión
- Que estén en tratamiento anticoagulante o consumiendo algún tipo de antiinflamatorio no esterooidal.
- Que hayan sido donante de sangre durante la semana anterior a la toma de muestra.
- Que consuman exceso de suplementos vitamínicos.
- Embarazo

En resumen la muestra de pacientes se muestra en la siguiente tabla de clasificación.

	PACIENTE CON AR	SUJETOS CONTROLES	TOTAL TIPO DE RCG
TIPO DE RCG BAJO	5	14	19
TIPO DE RCG MODERADO	48	26	74
TIPO DE RCG ALTO	4	2	6
TOTAL CONDICIÓN AR	57	42	TOTAL MUESTRA = 99

Tabla 1: Clasificación del tamaño de la muestra.

Observación: Se excluyeron 2 pacientes por falta de su riesgo cardiovascular global.

## 8. Metodología

Para llegar a comprender los conceptos teóricos que se incluyen en la investigación, se realizará una revisión bibliográfica a través de libros, documentos y páginas Web, mostradas en la parte de Bibliografía.

Luego para conocer la población objetivo y el contenido de la base de datos de los pacientes y validarlo, primero se estudiarán las diferentes variables y después se revisará la base de datos adjunta que contiene a dichas variables con el fin de validarla, dichas variables son las que se muestran a continuación.

### **Variables de Interés.**

- Tipo de Riesgo Cardiovascular Global
- Condición AR (Artritis Reumatoide)

### **Variables Clínicas.**

- Diabetes
- Fumador
- Tratamiento Antihipertensivo
- Antecedentes Familiares de Infarto

### **Variables Antropométricas.**

- Sexo
- Edad
- Hipertensión
- Obesidad Central
- Índice De Masa Corporal (IMC)

### **Variables Bioquímicas**

- Hipercolesterolemia
- Colesterol HDL
- Colesterol LDL
- Colesterol VLDL
- Colesterol TOTAL/HDL
- Creatinina
- Homa -IR
- Hiperinsulinemia
- Tasa de filtración Glomerular(TFG)
- Triglicéridos > 150 mg/dL (Hipertrigliceridemia)
- Glicemia > 100 mg/dL (Hiperglicemia)
- Proteína C Reactiva Ultrasensible (PCRus)

### **Biomarcadores**

- Interleuquina-6 (IL6)
- Factor de Necrosis Tumoral (TNF- $\alpha$ )
- Proteína C Reactiva Ultrasensible (PCRus)
- Especies Reactivas de Oxígeno (ROS)
- Oxido Nítrico (NO)

Para lograr el objetivo general de la investigación, se aplicarán todas las herramientas ya aprendidas por los cursos anteriores tales como: Inferencia Estadística, Regresión Lineal, Métodos Multivariados, Diseño de Experimentos, entre otros.

Así para analizar y comparar entre ambos grupos de pacientes la relación del riesgo cardiovascular global y los diferentes niveles de biomarcadores, se utilizarán técnicas tales como: tabulaciones cruzadas, pruebas chi-cuadrado, análisis de correspondencia múltiple, análisis de varianzas unifactorial y multifactorial, regresión Logística, entre otros; con el fin de determinar perfiles que mejoren la condición de estos dos grupos de personas. Además se realizarán pruebas de Normalidad para evaluar previamente a la población de estudio.

Los datos se analizarán principalmente con los software estadísticos: R, SPSS, Statgraphics y PSPP.

Para estos análisis será necesario antes categorizar cada una de las principales variables a considerar en esta investigación, tal como se muestra en la siguiente tabla.

Tabla 2: Categorización de todas las variables.

VARIABLE	CATEGORIZACIÓN
<b>Tipo de Riesgo Cardiovascular</b>	1: Bajo                      3: Alto 2: Moderado
<b>Artritis Reumatoide</b>	1: Paciente con AR 2: Sujeto Control
<b>Diabetes</b>	1: si 2: no
<b>Edad</b>	1: ≤ 54 años      3: ≥75 2: 54-75 años
<b>Género</b>	1: Hombre 2: Mujer
<b>Índice de masa corporal (IMC)</b>	1: normal (18,5 - 24,9 kg/m <sup>2</sup> ) 2: sobrepeso (25 - 29,9 kg/m <sup>2</sup> ) 3: obeso (≥ 30 kg/m <sup>2</sup> )
<b>Fumador</b>	1: si 2: no
<b>Hipertrigliceridemia</b>	1: si ≥ 150 mg/dL 2: no < 150 mg/dL
<b>Hipercolesterolemia</b>	1: si ≥ 200 mg/dL 2: no < 200 mg/dL
<b>Hiperglicemia</b>	1: si ≥ 100 mg/dL 2: no < 100 mg/dL
<b>Colesterol LDL alto</b>	1: si ≥ 130 mg/dL 2: no < 130 mg/dL
<b>Colesterol HDL bajo</b>	1: si < 35 mg/dL 2: no ≥ 35 mg/dL
<b>Colesterol VLDL alto</b>	1: si ≥ 40 mg/dL 2: no < 40 mg/dL
<b>Colesterol TOTAL / HDL</b>	1: si ≥ 5 2: no < 5
<b>Obesidad central</b>	1: si (hombres ≥102cm, Mujer ≥88cm) 2: no (hombres <102cm, Mujer <88cm)
<b>Hipertensión Arterial</b>	1: si 2: no
<b>Tratamiento antihipertensivo</b>	1: si 2: no
<b>Tasa de Filtración Glomerular (TFG)</b>	1: Baja < 60 mL/min 2: Normal ≥60 mL/min
<b>Antecedentes familiares de infarto</b>	1: si 2: no
<b>Creatinina alta</b>	1: si ≥ 1,1 2: no < 1.1
<b>Homa-IR</b>	1: si > 2,5 2: no ≤ 2,5
<b>Hiperinsulinemia</b>	1: si ≥ 14 2: no <14
<b>Proteína C Reactiva Ultrasensible (PCRus Alta)</b>	1: si > 3 mg/L 2: no ≤ 3 mg/L

A continuación se procede a revisar el procedimiento de algunas de las técnicas estadísticas a utilizar en esta segunda parte de este proyecto de título.

- **Análisis Multivariado: Análisis de Correspondencia Múltiple (ACM)**

El Análisis de Correspondencias es una técnica estadística que se aplica al análisis de tablas de contingencia y construye un diagrama cartesiano basado en la asociación entre las variables analizadas. En dicho gráfico se representan conjuntamente las distintas modalidades de la tabla de contingencia, de forma que la proximidad entre los puntos representados está relacionada con el nivel de asociación entre dichas modalidades. Es decir, esta técnica estadística permite asociar gráficamente tres o más variables y establecer grupos de casos de acuerdo al grado de cercanía que las categorías de cada variable en su conjunto establecen entre sí, cuanto más cerca estén estos puntos mayor será la afinidad entre las variables que representan, ahora bien, para que estas relaciones sean validas, dichos puntos deben estar alejados del origen, de lo contrario la posible relación que se produzca será mínima. (4)

- **Test de Independencia: Chi- Cuadrado**

Para evaluar de una manera más precisa las posibles relaciones existentes entre la variable de interés (Tipo de RCG) y todas las demás variables, incluyendo a la variable Artritis Reumatoide que denota a los dos grupos de pacientes considerados en este estudio, se realizarán entre ellos tabulaciones cruzadas (tablas de contingencia) con sus respectivos test de independencia Chi-Cuadrado, para esto se plantearán las siguientes:

**Hipótesis.**

H0: La variable y Condición AR (ó Tipo de RCG ó Biomarcador) son independientes.

H1: La variable y Condición AR (ó Tipo de RCG ó Biomarcador) no son independientes.

De esta manera, evaluando el valor-p, será posible aceptar o rechazar la hipótesis H0 considerando un nivel de confianza del 95%.

- **Diseño de Experimentos: Diseño en Bloques aleatorios Completos.**

## Introducción

Al estudiar la influencia de uno o más factor(es) sobre una variable cuantitativa (variable Riesgo Cardiovascular Global) es frecuente que aparezcan otras variables o factores que también influyan y deban ser controladas.

A estas variables se las denomina variables bloque, y se caracterizan por:

- i) No son el motivo del estudio sino que aparecen de forma natural
- ii) Se asume que no tienen interacción con el factor en estudio.

Para efecto de esta técnica estadística, definiremos a continuación algunos de los conceptos básicos más comunes que son propios de estos diseños.

Modelo de Diseño Experimental: Corresponde, en la mayor parte de los casos a un modelo de tipo lineal, en el cual encontramos tres elementos básicos, una variable respuesta, un conjunto de parámetros y un término de error experimental.

Tratamientos: Corresponde a un procedimiento, generalmente de tipo experimental, cuyo efecto queremos medir.

Unidad Experimental: Es la unidad muestral básica sobre la que se aplica el tratamiento y por lo tanto nos permite medir la respuesta al tratamiento.

Error Experimental: corresponde a una medida de las distintas respuestas que dan unidades experimentales equivalentes y que son tratadas en forma similar (con un mismo tratamiento).

El modelo se dice que es de bloques aleatorizados completos cuando en cada bloque se presentan todos los posibles tratamientos (o un múltiplo de ese número) y dentro de cada bloque se asignan los tratamientos de forma aleatoria.

**Modelo.**

El modelo lineal que describe una observación de un modelo de diseño en bloque aleatorizado, en el cual el análisis está basado, es

$$Y_{ij} = \mu + \beta_j + \tau_i + \varepsilon_{ij}$$

$Y_{ij}$  = rendimiento del i-ésimo tratamiento en el j-ésimo bloque.

$\mu$  = es la media de todas las observaciones.

$\beta_j$  = adiciona una característica del efecto en el j-ésimo bloque.

$\tau_i$  = una característica del efecto en el i-ésimo tratamiento.

$\varepsilon_{ij}$  = error aleatorio asociado con la unidad en el i-ésimo tratamiento en el j-ésimo bloque.

Lo anterior, bajo los siguientes supuestos.

**Supuestos.**

- Normalidad:  $\varepsilon_{ij}$  sigue una distribución normal. Esto es equivalente a que  $Y_{ij}$  sigue una distribución normal.
- Linealidad:  $E(\varepsilon_{ij})=0$ : Esto es equivalente a que  $E(Y_{ij})=\mu + \beta_j + \tau_i$ .
- Homocedasticidad:  $\text{Var}(\varepsilon_{ij})=\sigma^2$ . Esto es equivalente a que  $\text{Var}(Y_{ij})=\sigma^2$ .
- Independencia:  $\varepsilon_{ij}$  son independientes entre sí. Esto es equivalente a que  $Y_{ij}$  son independientes entre sí.

Tabla 3: ANOVA para un Diseño en Bloques Aleatorizados.

F. de Variación	g.l.	SC	CM	F
Bloques	$r-1$	SCB	CMB	$F_B$
Tratamiento	$t-1$	SCTR	CMTR	$F_{TR}$
Error	$(r-1)(t-1)$	SCE	CME	
tTotal	$rt-1$	SCT		

$$F_{TR} = \frac{CMTR}{CME} \approx f_{(t-1, (t-1)(r-1))}.$$

Las hipótesis a probar por  $F_{RT}$  dependen de si los tratamientos son fijos o efectos aleatorio. Por lo general en diseño de experimento los tratamientos son fijos.

En este caso  $F_{RT}$  prueba la hipótesis:

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_t = 0$$

$$H_1 : \tau_i \neq 0 \text{ para algún } i$$

Esta es una prueba de significación de la diferencia entre las medias de los tratamientos (variable Artritis Reumatoide), es decir, nos permitirá determinar si hay diferencias o no entre ambos grupos de pacientes tanto para el riesgo cardiocascular Global como para los diferentes Biomarcadores, considerando las diferentes otras variables que puedan influir en la interpretación de los resultados finales (Bloques).

La tabla de ANOVA anterior nos permite probar la hipótesis de equivalencia entre los efectos de los tratamientos. Sin embargo, es posible además extraer algunas conclusiones a partir de la construcción de una razón del tipo  $\frac{CMB}{CME}$ .

Para este efecto analicemos la razón F siguiente, que determinará si los bloques son o no significativos.

$$F_B = \frac{CMB}{CME} \approx f_{(r-1, (t-1)(r-1))}.$$

En primer lugar debemos notar que la aleatorización se realiza sólo al interior de los bloques y no entre ellos. Esto significa que en rigor no estamos en condiciones de probar la hipótesis de equivalencia entre los bloques.

En efecto, esta hipótesis no ha sido planteada y si queremos considerar al factor de bloque como a un factor que debe ser analizado, deberíamos considerarlo en el momento de proceder a la aleatorización de los tratamientos, luego las siguientes consideraciones deben ser tenidas en cuenta para el análisis de la información. Por lo tanto, las siguientes consideraciones no tienen validez desde un punto de vista inferencial sino deben ser consideradas sólo referenciales.

Debemos recordar que la definición de los bloques se basó en el supuesto que ellos agrupaban a unidades reconocidamente semejantes pero diferentes para cada bloque, por lo que se espera que las diferencias entre bloques sean grandes. Esto significa que valores altos de la estadística F, es decir, valores pequeños de valor- p en la tabla ANOVA, asociada con los bloques serían una indicación que efectivamente el bloqueo fue bien realizado.(6, 7)

- **Regresión No Lineal: Modelos Logísticos**

### **Introducción**

Cuando tengamos una variable Y dependiente dicotómica, es decir, una variable respuesta que pueda tomar dos valores como: 0/1, si/no, sano/enfermo, paga/no paga etc.; que deseemos predecir, o para la que queramos evaluar la asociación o relación con otras (más de una) variables independientes y de control, el procedimiento a realizar es una REGRESIÓN LOGÍSTICA (RL) BINARIA MULTIVARIANTE.

El modelo de regresión logística se utiliza cuando estamos interesados en pronosticar la **probabilidad** de que ocurra o no un suceso determinado. Por ejemplo, a la vista de un conjunto de pruebas médicas, que una persona tenga una determinada enfermedad. (8)

Acá se podrá estudiar el impacto que tiene cada una de las variables explicativas X (Parámetros ó Variables y Biomarcadores) en la probabilidad de que ocurra el suceso en estudio Y (Tipo de RCV Global)

El análisis de regresión logística es una herramienta muy flexible en cuanto a la naturaleza de las variables explicativas, pues éstas pueden ser cuantitativas y categóricas.

***La Regresión Logística es probablemente el tipo de análisis multivariante más empleado en Ciencias de la Vida.***

### **Modelo**

El modelo de regresión logística parte de la hipótesis de que los datos siguen el siguiente modelo:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_k * x_k + \mu = x * b + \mu$$

Con el fin de simplificar la notación, definimos Z:

$$z = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_k * x_k$$

Por lo tanto el modelo, se puede representar como:

$$\ln\left(\frac{p}{1-p}\right) = z + \mu$$

Donde  $p$  es la probabilidad de que ocurra el suceso de estudio.

Operando algebraicamente sobre el modelo:

$$\ln\left(\frac{p}{1-p}\right) = z$$

$$\frac{p}{1-p} = e^z$$

$$p = (1-p) * e^z$$

$$p = e^z - p * e^z$$

$$p + p * e^z = e^z$$

$$p(1 + e^z) = e^z$$

$$p = \frac{e^z}{1 + e^z}$$

Como la función de distribución logística es:

$$F(x) = \frac{e^x}{1 + e^x}$$

Por tanto, podemos reescribir el modelo de forma mucho más compacta:

$$p = \frac{e^z}{1 + e^z} = F(z) = F(x * b)$$

De donde se deduce que el modelo de regresión logística es, en principio, un modelo de regresión **no lineal**, pero es lineal en escala logarítmica atendiendo a su definición original:

$$\ln\left(\frac{p}{1-p}\right) = z$$

$$\ln(p) - \ln(1-p) = z$$

$$\ln(p) - \ln(1-p) = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_k * x_k$$

Es decir, la diferencia de la probabilidad de que ocurra un suceso respecto de que no ocurra es lineal pero en escala logarítmica. Por tanto, el significado de los coeficientes, aunque guardando una cierta relación con el modelo de regresión lineal, va a ser algo más complejo de interpretar.

Por lo tanto recordemos las dos formas más importantes de expresar el modelo de regresión logística

$$- \ln(p) - \ln(1-p) = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_k * x_k$$

$$- \frac{P}{1-p} = e^{b_0} * e^{b_1 * x_1} * e^{b_2 * x_2} \dots e^{b_k * x_k}$$

Donde la primera expresión se llama **Logit** y la segunda **Odds Ratio** o cociente de probabilidad y nos serán útil para interpretar los coeficientes de regresión logística.

### Supuestos

La naturaleza dicotómica de la variable dependiente impide la utilización de métodos tradicionales, tales como el método de los mínimos cuadrados, en la estimación de parámetros. Modelos de este tipo se pueden resolver mediante métodos de máxima verosimilitud. (9, 10) Además el análisis de regresión logística sólo precisa del principio de monotonía, es decir, si el suceso A es que una determinada persona padezca de artritis y X representa la edad, deberá de ocurrir:

$$x_i \geq x_j \Rightarrow P(A / x_i) \geq P(A / x_j)$$

## 9. Antecedentes Generales

- **Análisis Estadístico**

Para el análisis estadístico se escogerán los estadígrafos más pertinentes a utilizar directamente relacionado a los objetivos planteados. Los datos serán procesados utilizando principalmente los software SPSS y Statgraphics, y los resultados serán expresados en el desarrollo de este informe.

- **Aspectos éticos del estudio**

El proyecto será presentado ante el Comité de Ética del Hospital Regional de Concepción y el Comité de Ética de la Universidad de Concepción, quienes evaluarán su aprobación. Todos los pacientes y sujetos controles deberán firmar el consentimiento informado, basado en los principios enumerados en la declaración de Helsinki. No existen criterios para discontinuar la investigación, ya que este estudio sólo contempla la toma de una muestra de sangre bajo normas estandarizadas para medir parámetros bioquímicos.

- **Compensación económica**

Los participantes de este proyecto renuncian a compensación económica porque no existe riesgo mayor al mínimo y los fondos solicitados a la Dirección de Investigación sólo cubrirán los gastos de los insumos que son requeridos en este estudio. Sólo se cubrirán los pasajes de los pacientes al momento de asistir a la toma de muestra

## 10. Resultados.

Dado que hay dos variables de interés en el estudio, Condición de Artritis Reumatoide (AR) y Tipo de riesgo cardiovascular (RCV) Global de los pacientes, los resultados obtenidos a continuación serán primeramente para una variable y luego para la otra con la finalidad de comparar los perfiles combinados que se obtengan en ambos casos.

### 10.1 Relación y dependencia de *variables* con respecto a la *condición AR* de los pacientes.

#### 10.1.1 Descripción de la población.

Para analizar la homogeneidad y evaluar las relaciones de las variables con respecto a la condición AR de los pacientes, con la finalidad de considerar aquellas que puedan influir en la interpretación de los resultados en cuanto al riesgo cardiovascular global, se realizó un análisis de correspondencia múltiple.

A continuación se muestra el gráfico con el Análisis de correspondencia múltiple entre todas las variables evaluadas en la población incluyendo la variable Condición AR que denota a ambos grupos de pacientes, es decir, pacientes con AR y sujetos controles.



Esto se puede corroborar con la siguiente tabla de discriminación de pesos para todas las variables que se presenta a continuación.

**Tabla 4:** Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todas las variables con respecto a la condición AR.

	Ponderación de la variable	Dimensión		Media
		1	2	
CONDICIÓN AR	2	,133	,180	,157
GÉNERO	2	,001	,000	,000
EDAD	3	,237	,505	,371
IMC	3	,351	,149	,250
OBESIDAD CENTRAL	2	,452	,013	,232
HIPERTENSIÓN	2	,453	,164	,309
TRATAMIENTO ANTIHIPERTENSIVO	2	,390	,249	,319
FUMADOR	2	,010	,007	,008
DIABETES	2	,238	,088	,163
HIPERCOLESTEROLEMIA	2	,058	,138	,098
HIPERGLICEMIA	2	,142	,078	,110
HIPERTRIGLICERIDEMIA	2	,175	,047	,111
COLESTEROL LDL ALTO	2	,035	,217	,126
COLESTEROL HDL BAJO	2	,029	,061	,045
COLESTEROL VLDL ALTO	2	,130	,042	,086
COLESTEROL TOTAL / HDL	2	,155	,003	,079
CREATININA ALTA	2	,001	,083	,042
TFG	2	,002	,263	,133
HOMA-IR	2	,315	,114	,214
HIPERINSULINEMIA	2	,220	,081	,150
ANTECEDENTES FAMILIARES	2	,079	,029	,054
PCRus	2	,242	,016	,129
TIPO RCV GLOBAL	3	,645	,044	,345
Total activo <sup>a</sup>		10,214	5,841	8,028

Para determinar cual eje tiene más peso en cada una de las variables se adjunta esta tabla de medidas de discriminación para las dos dimensiones o ejes: dimensión 1 (eje X) y dimensión 2 (eje Y). Acá se ve el mayor valor entre las 2 dimensiones, por lo tanto ese sería el mayor peso para cada variable. Por ejemplo para la variable Condición AR, la dimensión 2 es mayor, entonces se divide el eje Y en dos partes, del eje X hacia arriba y del eje X hacia abajo, donde todas las variables que caigan en un lado y en el otro se podrían asociar, tal como se ha comentado anteriormente.

Para evaluar de una manera más precisa estas relaciones, se realizaron tabulaciones cruzadas (tablas de contingencia) con sus respectivos test de independencia Chi-Cuadrado para cada una de las variables con la condición AR de los pacientes.

## 10.1.2 Test de Independencia.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

**Tabla 5:** Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y la condición AR de los pacientes.

VARIABLE	VALOR-P	HIPOTESIS H <sub>0</sub>	DEPENDENCIA
<b>Tipo de RCG</b>	<b>0,0103 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Diabetes	0,1373 > 0,05	ACEPTADA	NO
<b>Edad</b>	<b>0,0010 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Género	0,2016 > 0,05	ACEPTADA	NO
IMC	0,7561 > 0,05	ACEPTADA	NO
Fumador	0,5983 > 0,05	ACEPTADA	NO
Hipertrigliceridemia	0,5137 > 0,05	ACEPTADA	NO
Hipercolesterolemia	0,6473 > 0,05	ACEPTADA	NO
Hiperglicemia	0,3938 > 0,05	ACEPTADA	NO
Colesterol LDL alto	0,4625 > 0,05	ACEPTADA	NO
Colesterol HDL bajo	0,0820 > 0,05	ACEPTADA	NO
Colesterol VLDL alto	0,2034 > 0,05	ACEPTADA	NO
Colesterol TOTAL /HDL	0,6907 > 0,05	ACEPTADA	NO
<b>Obesidad central</b>	<b>0,0001 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Hipertensión Arterial</b>	<b>0,0118 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Tratamiento Antihipertensivo</b>	<b>0,0001 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
TFG	0,4690 > 0,05	ACEPTADA	NO
Creatinina alta	0,3866 > 0,05	ACEPTADA	NO
Homa-IR	0,5520 > 0,05	ACEPTADA	NO
Hiperinsulinemia	0,6836 > 0,05	ACEPTADA	NO
Antecedentes familiares de infarto	0,4229 > 0,05	ACEPTADA	NO
<b>PCRus</b>	<b>0,0001 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>

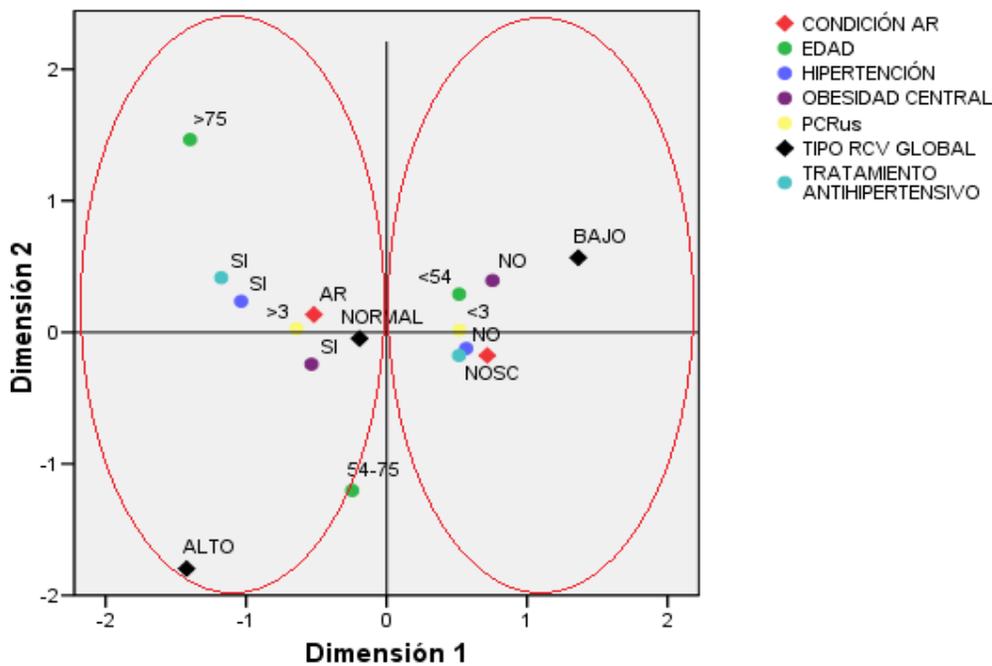
Al analizar cada uno de los valores de la tabla anterior, se observa que seis variables presentan un valor  $p < 0,05$ , siendo el mayor de estos de 0,0118, lo que significa que dichas variables presentan gran dependencia con respecto a las diferentes categorías de la variable Condición AR, estas variables son las siguientes:

- Edad
- Obesidad Central
- Tratamiento Antihipertensivo
- PCRus
- Hipertensión Arterial
- **Tipo de RCG**

Por lo tanto, deberán considerarse más adelante estos factores como las posibles variables que puedan influir en la comparación del valor numérico del riesgo cardiovascular global y en el modelo de regresión logística final, entre la población con Artritis Reumatoide y sujetos controles.

Adicionalmente, si consideramos solo estas variables en el análisis de correspondencia múltiple, es posible observar más claramente la relación de estas variables con respecto a ambas categorías de la variable Condición AR en la siguiente figura.

**Diagrama conjunto de puntos de categorías**



**Figura 7:** Análisis de correspondencia múltiple para variables que dieron dependencia con el porcentaje de obstrucción

Las líneas rojas denotan la agrupación de variables que se relacionan entre sí, respecto a cada una de las categorías de la variable Condición AR, que lo ha determinado la tabla de discriminación de pesos que se muestra a continuación.

**Tabla 6:** Medidas de discriminación de pesos del Análisis de correspondencia múltiple para variables que dieron dependencia con la variable Condición AR.

	Ponderación de la variable	Dimensión		Media
		1	2	
CONDICIÓN AR	2	,370	,023	,197
EDAD	3	,451	,808	,629
OBESIDAD CENTRAL	2	,378	,092	,235
HIPERTENSIÓN	2	,589	,029	,309
TRATAMIENTO ANTIHIPERTENSIVO	2	,609	,074	,342
PCRus	2	,329	,000	,165
TIPO RCV GLOBAL	3	,497	,254	,376
Total activo <sup>a</sup>		7,394	3,624	5,509

Acá solo basta mirar el peso de la variable CONDICIÓN AR para que se vean claramente las relaciones que existen para cada uno de los grupos de pacientes con respecto a las demás variables, es decir, dado que el peso para la dimensión 1 es mayor al peso de la dimensión 2, se tiene que el eje X se divide en dos partes, o sea del eje Y hacia la derecha y del eje Y hacia la izquierda.

### 10.1.3 Obtención de primer perfil.

Así las relaciones para ambos grupos de pacientes son evidentes. Por lo tanto, así se obtiene entonces el primer perfil de esta investigación que corresponde a la asociación de los **variables** que presentan dependencia con respecto a la variable **condición AR**.

**Tabla 7:** Perfil de todas las variables asociadas con la condición AR de los pacientes.

RELACIÓN DE LOS PACIENTES	
CON ARTRITIS REUMATOIDE	SUJETOS CONTROLES
TIPO DE RCV GLOBAL NORMAL Y ALTO	TIPO DE RCV GLOBAL BAJO
EDAD > 54 AÑOS	EDAD < 54 AÑOS
CON HIPERTENSIÓN	SIN HIPERTENSIÓN
CON OBESIDAD CENTRAL	SIN OBESIDAD CENTRAL
CON TRATAMIENTO ANTIHIPERTENSIVO	SIN TRATAMIENTO ANTIHIPERTENSIVO
CON PCR > 3 mg/L	CON PCR < 3 mg/L

## 10.2 Relación y dependencia de variables con respecto al Tipo de riesgo cardiovascular Global de los pacientes.

### 10.2.1 Descripción de la población.

Ahora para analizar la homogeneidad y evaluar las relaciones de las variables con respecto a las tres categorías de la variable Tipo de RCV Global de los pacientes, se realizó un nuevo análisis de correspondencia múltiple.

A continuación se muestra el gráfico con el Análisis de correspondencia múltiple entre todas las variables evaluadas en la población incluyendo la variable Tipo de RCV Global que denota a los tres grupos de pacientes, es decir, pacientes con un Tipo de RCV Global Bajo, Moderado y Alto.

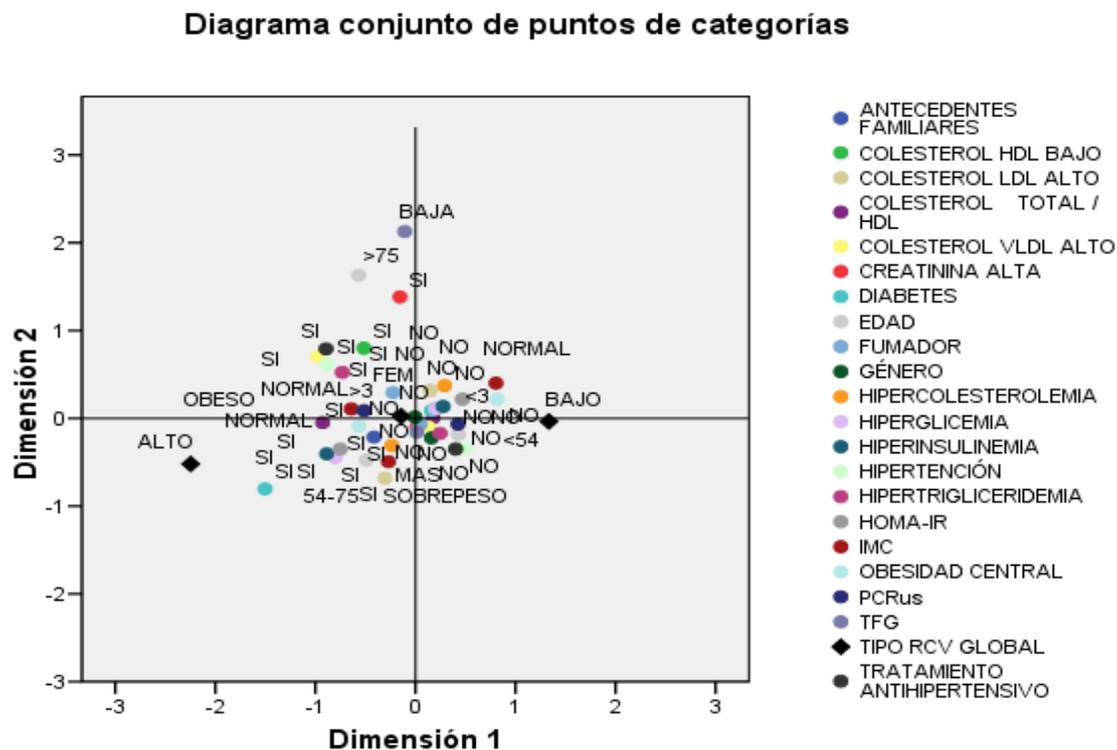


Figura 8: Análisis de Correspondencia múltiple para todas las variables de acuerdo al Tipo de RCV Global.

Del diagrama de Análisis de correspondencia múltiple, se puede observar que a simple vista la población de estudio nuevamente se comporta bastante homogénea para las tres categorías de riesgo cardiovascular global (denotados por rombos negros), eso sí que la relación que existe entre la mayoría de las variables con respecto a estos grupos es bastante pequeña ya que también se aproximan bastante al origen por ambos ejes de coordenadas, aun así, es posible distinguir algunos grupos de variables relacionadas: Por ejemplo para las personas con un Riesgo Cardiovascular Moderado, en particular, se puede ver que están casi en el centro de la nube de puntos (origen de gráfico) por lo que no es posible distinguir muy bien a que categorías del resto de variables se asocia más. Por su parte, personas con Riesgo Cardiovascular Bajo, se caracterizan más por tener un IMC Normal, sin Hipercolesterolemia, sin obesidad central y sin Hipertensión. Finalmente encontramos que pacientes con un Riesgo Cardiovascular Alto se relacionan más a tener Diabetes, Hiperглиcemia, Hiperinsulinemia y Homa-IR.

Esto se puede corroborar con la siguiente tabla de discriminación de pesos para todas las variables que se presenta a continuación.

**Tabla 8:** Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todas las variables con respecto al Tipo de RCV Global.

	Ponderación de la variable	Dimensión		Media
		1	2	
TIPO RCV GLOBAL	3	,649	,017	,333
GÉNERO	2	,002	,003	,002
EDAD	3	,220	,481	,350
IMC	3	,372	,144	,258
OBESIDAD CENTRAL	2	,434	,023	,228
HIPERTENSIÓN	2	,436	,213	,325
TRATAMIENTO ANTIHIPERTENSIVO	2	,357	,277	,317
FUMADOR	2	,011	,020	,016
DIABETES	2	,247	,071	,159
HIPERCOLESTEROLEMIA	2	,069	,116	,092
HIPERGLICEMIA	2	,161	,049	,105
HIPERTRIGLICERIDEMIA	2	,178	,090	,134
COLESTEROL LDL ALTO	2	,044	,216	,130
COLESTEROL HDL BAJO	2	,023	,055	,039
COLESTEROL VLDL ALTO	2	,128	,068	,098
COLESTEROL TOTAL / HDL	2	,162	,000	,081
CREATININA ALTA	2	,002	,142	,072
TFG	2	,001	,337	,169
HOMA-IR	2	,347	,075	,211
HIPERINSULINEMIA	2	,243	,052	,147
ANTECEDENTES FAMILIARES	2	,082	,022	,052
PCRus	2	,216	,006	,111
Total activo <sup>a</sup>		10,008	5,597	7,803

Para determinar cual eje tiene más peso para cada variable se adjunta esta tabla de medidas de discriminación para las dos dimensiones o ejes: dimensión 1 (eje x) y dimensión 2 (eje y). Acá se ve el mayor valor entre las 2 dimensiones, por lo tanto ese sería el mayor peso para cada variable. Por ejemplo para la variable Tipo RCV Global, la dimensión 1 es mayor entonces se divide el eje x en dos partes, del eje y hacia la derecha y del eje y hacia la izquierda, donde todas las variables que caigan en un lado y en el otro se podrían asociar, tal como se ha comentado anteriormente.

## 10.2.2 Test de Independencia.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

Tabla 9: Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el Tipo de riesgo cardiovascular global de los pacientes.

VARIABLE	VALOR-P	HIPOTESIS H <sub>0</sub>	DEPENDENCIA
<b>Artritis Reumatoide</b>	<b>0,0103 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Diabetes</b>	<b>0,0000 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Edad	0,0687 > 0,05	ACEPTADA	NO
Género	0,7855 > 0,05	ACEPTADA	NO
<b>IMC</b>	<b>0,0008 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Fumador	0,1443 > 0,05	ACEPTADA	NO
<b>Hipertrigliceridemia</b>	<b>0,0099 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Hipercolesterolemia	0,2040 > 0,05	ACEPTADA	NO
<b>Hiperglicemia</b>	<b>0,0000 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Colesterol LDL alto	0,3731 > 0,05	ACEPTADA	NO
Colesterol HDL bajo	0,0655 > 0,05	ACEPTADA	NO
Colesterol VLDL alto	0,1903 > 0,05	ACEPTADA	NO
<b>Colesterol TOTAL /HDL</b>	<b>0,0363 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Obesidad central</b>	<b>0,0000 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Hipertensión Arterial</b>	<b>0,0002 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Tratamiento Antihipertensivo</b>	<b>0,0017 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
TFG	0,6278 > 0,05	ACEPTADA	NO
Creatinina alta	0,6278 > 0,05	ACEPTADA	NO
<b>Homa-IR</b>	<b>0,0010 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Hiperinsulinemia</b>	<b>0,0079 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Antecedentes familiares de infarto</b>	<b>0,0151 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>PCRus</b>	<b>0.0010 &lt; 0.05</b>	<b>RECHAZADA</b>	<b>SI</b>



Las líneas rojas denotan la agrupación de variables que se relacionan entre sí, respecto a cada una de las categorías de la variable Tipo de RCG, que lo ha determinado la tabla de discriminación de pesos que se muestra a continuación.

Tabla 10: Medidas de discriminación de pesos del Análisis de correspondencia múltiple para variables que dieron dependencia con la variable tipo de RCV Global.

	Ponderación de la variable	Dimensión		Media
		1	2	
TIPO RCV GLOBAL	3	,664	,588	,626
OBESIDAD CENTRAL	2	,466	,071	,269
HIPERTENSIÓN	2	,442	,032	,237
TRATAMIENTO ANTIHIPERTENSIVO	2	,374	,053	,213
DIABETES	2	,277	,206	,241
HIPERGLICEMIA	2	,184	,177	,181
HIPERTRIGLICERIDEMIA	2	,139	,122	,130
COLESTEROL TOTAL / HDL	2	,141	,000	,070
IMC	3	,374	,247	,310
HOMA-IR	2	,363	,141	,252
HIPERINSULINEMIA	2	,266	,070	,168
ANTECEDENTES FAMILIARES	2	,069	,017	,043
PCRus	2	,255	,004	,129
CONDICIÓN AR	2	,117	,136	,126
Total activo <sup>a</sup>		9,300	4,560	6,930

Acá solo basta mirar el peso de la variable TIPO RCV GLOBAL para que se vean claramente las relaciones que existen para cada uno de los grupos de pacientes con respecto a las demás variables, es decir, dado que el peso para la dimensión 1 es mayor al peso de la dimensión 2, se tiene que el eje X se divide en dos partes, o sea del eje Y hacia la derecha y del eje Y hacia la izquierda.

### 10.2.3 Obtención de segundo perfil.

Así las relaciones para los tres grupos de pacientes son evidentes. Por lo tanto, así se obtiene entonces el segundo perfil de esta investigación que corresponde a la asociación de las **variables** que presentan dependencia con respecto a la variable **Tipo de RCV Global**.

Tabla 11: Perfil de todas las variables asociadas con el Tipo de RCV Global de los pacientes.

RELACIÓN DE LOS PACIENTES	
CON RCV GLOBAL BAJO	CON RCV GLOBAL MODERADO - ALTO
PACIENTES CONTROLES	PACIENTES CON ARTRITIS REUMATOIDE
SIN ANTECEDENTES FAMILIARES	CON ANTECEDENTES FAMILIARES
CON COLESTEROL TOTAL/HDL < 5	CON COLESTEROL TOTAL/HDL > 5
SIN DIABETES	CON DIABETES
SIN HIPERGLICEMIA	CON HIPERGLICEMIA
SIN HIPERTRIGLICERIDEMIA	CON HIPERTRIGLICERIDEMIA
SIN HIPERTENSIÓN	CON HIPERTENSIÓN
SIN TRATAMIENTO ANTIHIPERTENSIVO	CON TRATAMIENTO ANTIHIPERTENSIVO
CON HIPERINSULINEMIA < 14	CON HIPERINSULINEMIA > 14
CON HOMA-IR < 2,5	CON HOMA-IR >2,5
CON IMC NORMAL	CON SOBREPESO Y OBESIDAD
SIN OBESIDAD CENTRAL	CON OBESIDAD CENTRAL
CON PCRus < 3 mg/L	CON PCRus > 3 mg/L

Así cada vez se hace más interesante la idea de comparar los perfiles ya obtenidos de **variables** y por obtener a continuación de **biomarcadores** entre ambas variables de interés (Condición AR y Tipo de RCV Global), ya que tanto en el punto 10.1.2 y 10.2.2 que corresponden a los test chi-cuadrados de independencia de dichas variables, se muestra una evidente dependencia significativa entre las variables CONDICION AR Y TIPO DE RCV GLOBAL, con un fuerte valor-p igual a 0.0103 al 95% de confianza.

### 10.3 Relación y dependencia de *Biomarcadores* con respecto a la *Condición AR* de los pacientes.

Para analizar y evaluar las relaciones de estos biomarcadores con respecto a la condición AR de los pacientes, se realizó un nuevo análisis de correspondencia múltiple, pero dado que estas variables son numéricas, fue necesario antes categorizarlas.

Para categorizar estos biomarcadores se hizo uso de las curvas ROC, ya que no existen valores de corte para determinar dichas categorías en algunos casos, como se muestra a continuación:

Las Curvas ROC nos permite realizar pruebas de diagnostico para una cierta variable con la finalidad de determinar un punto de corte óptimo y así poder clasificar en categorías dicha variable.

A continuación se muestra el desarrollo para el biomarcador ROS, el cual se obtuvo a través del software PSPP.

Presentación e interpretación de los resultados:

1) Lo primero que aparece es un resumen del procesamiento de los casos indicándonos cuantos casos tenemos con el problema (los que tienen la enfermedad) y cuantos sin el problema, en nuestro caso son 38 sujetos controles y 57 pacientes con AR (en total son 95, y no los 101 originales ya que faltan 6 datos del biomarcador ROS). Además, se nos recuerda que se han hecho todos los cálculos suponiendo que los casos con la enfermedad son aquellos que tienen el valor 1 en la variable COND\_AR (CONDICION ARTRITIS REUMATOIDE).

ROC  
ROC ROS BY COND\_AR (1)  
/PLOT CURVE (REFERENCE)  
/PRINT SE COORDINATES.

COND_AR	N Válido (listwise)	
	No ponderado	Ponderado
Positivo	57	57,00
Negativo	38	38,00

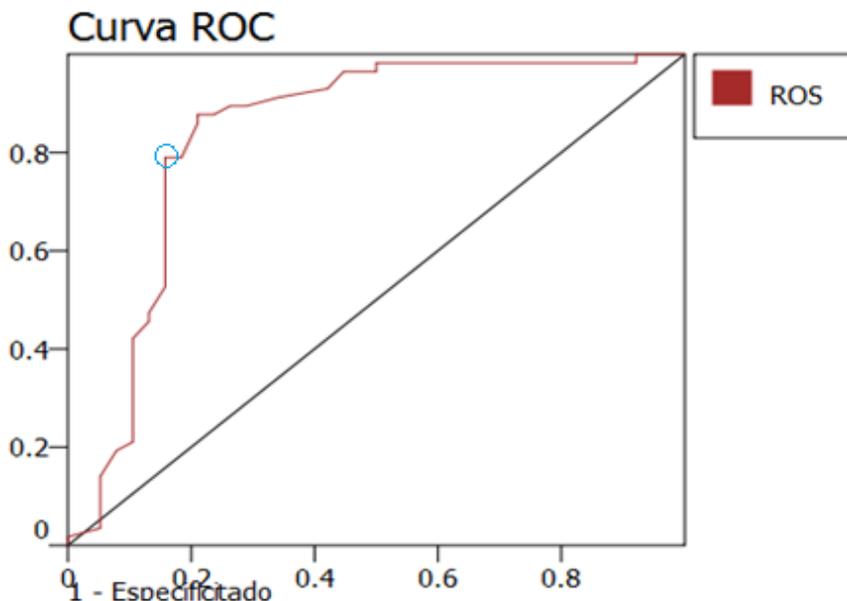


Figura 10: Curva ROC para el biomarcador ROS.

2) Tras ello, aparece el dibujo de la curva ROC ajustada a nuestros datos, en donde, cuanto más alejada esté la curva ROC de la diagonal principal mejor es el método de diagnóstico y cuanto más cercana esté a dicha diagonal peor será el método de diagnóstico; recuérdese que la diagonal principal es la que corresponde al peor test diagnóstico y que tienen un área bajo ella de 0.5.

Área	Error Estándar	Sig. Asintótica	Asymp. 95% Intervalo de Confianza	
			Límite Inferior	Límite Superior
,84	,05	,00	,76	,92

3) La estimación puntual del área bajo la curva es de 0.84 que como se ve difiere demasiado de 0.5 que sería el mínimo exigible para un método de diagnóstico. El error estándar de esa estimación es bastante pequeño y vale 0.05, valor que multiplicado por 1,96 (para una confianza del 95%) y sumado y restado de 0,84 nos da aproximadamente el intervalo de confianza que figura al final y que da de límite inferior 0.76 y de límite superior 0.92. Como el intervalo no contiene al valor 0.5 podemos afirmar que el área bajo la curva ROC para este ejemplo es significativamente mayor que lo mínimo exigible 0.5; eso se ve confirmado con la significación asintótica que aparece en la tabla y que no es más que el valor de P del test, que siendo A el área bajo la curva ROC, tiene de hipótesis nula que  $A=0.5$  y como alternativa que  $A \neq 0.5$ ; como en nuestro caso el valor-P (sig. asintótica) es aproximadamente 0 podemos rechazar la hipótesis nula y acabar concluyendo que el área bajo la curva ROC es significativamente distinta de 0.5.

4) Lo último que aparece en los resultados es la lista de coordenadas de la curva ROC, esta tabla serviría para decidir puntos de corte para una sensibilidad o una especificidad fijada de antemano. (Anexo 4)

Tabla 12: Valores de corte para el biomarcador ROS.

Positivo si es mayor o igual a	Sensibilidad	1 - Especificidad
-,88	1,00	1,00
,31	1,00	,92
,33	1,00	,92
,36	,98	,92
,39	,98	,87
,41	,98	,84
,43	,98	,82
,48	,98	,76
,55	,98	,76
,55	,98	,76
,62	,98	,68
,70	,98	,58
,73	,98	,55
,78	,98	,50
,80	,98	,50
,81	,98	,50
,85	,98	,50
,89	,98	,50
,90	,96	,50
,92	,96	,50
,95	,96	,45
,99	,96	,45
1,13	,93	,42
1,21	,91	,34
1,25	,89	,29
1,25	,89	,29
1,36	,89	,26
1,39	,89	,26
1,54	,88	,24
1,59	,88	,21
1,65	,88	,21
1,87	,86	,21
1,89	,86	,21
2,12	,79	,18
2,13	,79	,16
2,15	,75	,16
2,21	,72	,16
2,27	,70	,16
2,31	,68	,16
2,31	,68	,16
2,42	,65	,16
2,49	,63	,16
2,59	,61	,16
2,60	,61	,16
2,67	,58	,16
2,71	,58	,16

Se ha demostrado que una buena prueba de este tipo es que tenga un punto de corte tal que la sensibilidad (S) y especificidad (E) sean altas (dígase mayores de 0,8) tendrá una curva con concavidad más afilada, y la curva ROC de una prueba perfecta (S y E iguales a 1) serán las líneas formadas por los propios ejes izquierdo y superior del recuadro que sirve de marco a la curva en el grafico anterior.

Observación: se recomienda una especificidad cercana a 0.8 pero dado que la tabla entrega el complemento de este valor, es decir, 1 – especificidad un valor cercano a 0.2 sería el más recomendable (<0.2).

Por lo tanto con la ayuda del gráfico y tabla anterior, se propone 2.12 como el punto de corte para el biomarcador ROS, el cual se denota con una pequeña circunferencia en la grafica anterior y con una especificidad de 0.82 y sensibilidad de 0.79 en la tabla.

Así al obtener el valor de corte antes señalado se puede proceder a categorizar este biomarcador, y en general el procedimiento es el mismo para todos los biomarcadores (ANEXOS 1 y 2), por lo que a continuación se presenta una tabla de resumen con los puntos de corte y las categorizaciones ya realizadas para el resto de biomarcadores.

Tabla 13: Categorización de todos los biomarcadores.

BIOMARCADOR	VALOR DE CORTE	CATEGORIZACIÓN
ROS	2.12	1: SI > 2.12 IF/ ug proteína 2: NO < 2.12 IF/ ug proteína
TNFa	4.65	1: SI > 4.65 pg/mL 2: NO < 4.65 pg/mL
PCR	3.19	1: SI > 3.19 mg/L 2: NO < 3.19 mg/L

### 10.3.1. Descripción de la población.

A continuación se muestra el gráfico con el Análisis de correspondencia múltiple entre todos los **biomarcadores** evaluados en la población incluyendo la variable Condición AR que denota a ambos grupos de pacientes, es decir, pacientes con Artritis Reumatoide y sujetos controles.

Diagrama conjunto de puntos de categorías

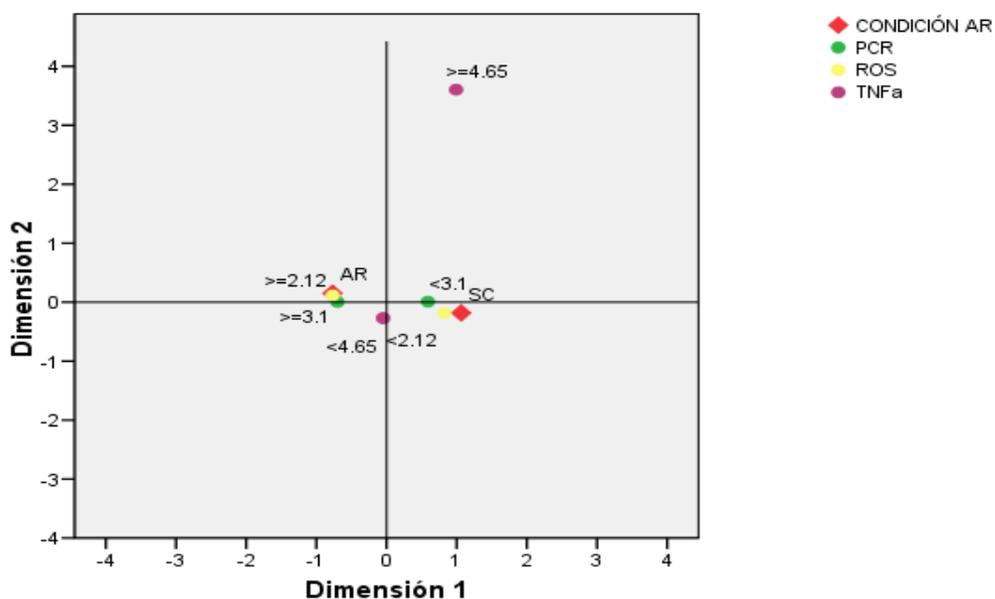


Figura 11: Análisis de Correspondencia múltiple para todos los biomarcadores de acuerdo a la condición AR.

Del diagrama de Análisis de correspondencia múltiple, se puede observar que a simple se conforman dos grupos de puntos por separado.

Por ejemplo, se observa que pacientes controles (SC en el gráfico) se caracterizan por tener una PCR < 3,1 mg/L y un ROS < 2.12 mg/l. Si ahora analizamos los puntos cercanos a la población de pacientes con Artritis Reumatoide (AR en el gráfico), se observa que están relacionadas con una PCR Alta, es decir, > ó igual a 3,1 mg/L y un ROS > ó igual a 2.12 mg/l. Por otra parte el comportamiento del biomarcador TNFa es indiferente para ambos grupos de pacientes, es decir, se asocia muy similarmente en ambos casos. Además se debe considerar que estas relaciones son bastante pequeñas por estar cerca del origen.

Esto se puede corroborar con la siguiente tabla de discriminación de pesos para todas las variables que se presenta a continuación.

Tabla 14: Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todas los biomarcadores con respecto a la condición AR.

**Medidas de discriminación**

	Ponderación de la variable	Dimensión		Media
		1	2	
CONDICIÓN AR	2	,816	,027	,422
TNFa	2	,070	,966	,518
PCR	2	,407	,000	,203
ROS	2	,593	,022	,308
Total activo <sup>a</sup>		3,772	2,032	2,902

a. Las ponderaciones de las variables están incorporadas en los estadísticos de Total activo.

### 10.3.2 Test de Independencia.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

Tabla15: Resumen de valores -p obtenidos mediante una prueba de independencia entre cada biomarcador y la condición AR de los pacientes.

BIOMARCADOR	VALOR-P	HIPOTESIS H <sub>0</sub>	DEPENDENCIA
TNFa	0,41 > 0,05	ACEPTADA	NO
PCR	0,0001 < 0,05	RECHAZADA	SI
ROS	0,0000 < 0,05	RECHAZADA	SI

Al analizar cada uno de los valores de la tabla anterior, se observa que solo dos de los tres biomarcadores presentan un valor  $p < 0,05$ , siendo el mayor de estos de 0,001, lo que significa que dichas variables presentan gran dependencia con respecto a las diferentes categorías de la variable Condición AR, estas variables son las siguientes:

- Proteína C reactiva (PCR)
- Especies Reactivas de Oxígeno (ROS)

Por lo tanto, deberán considerarse más adelante estos factores como las posibles variables que puedan influir en el modelo de regresión logística final.

Adicionalmente, si consideramos solo estas variables en un análisis de correspondencia múltiple, es posible observar más claramente la relación de estas con ambas categorías de la variable Condición AR en la siguiente figura.

### Diagrama conjunto de puntos de categorías

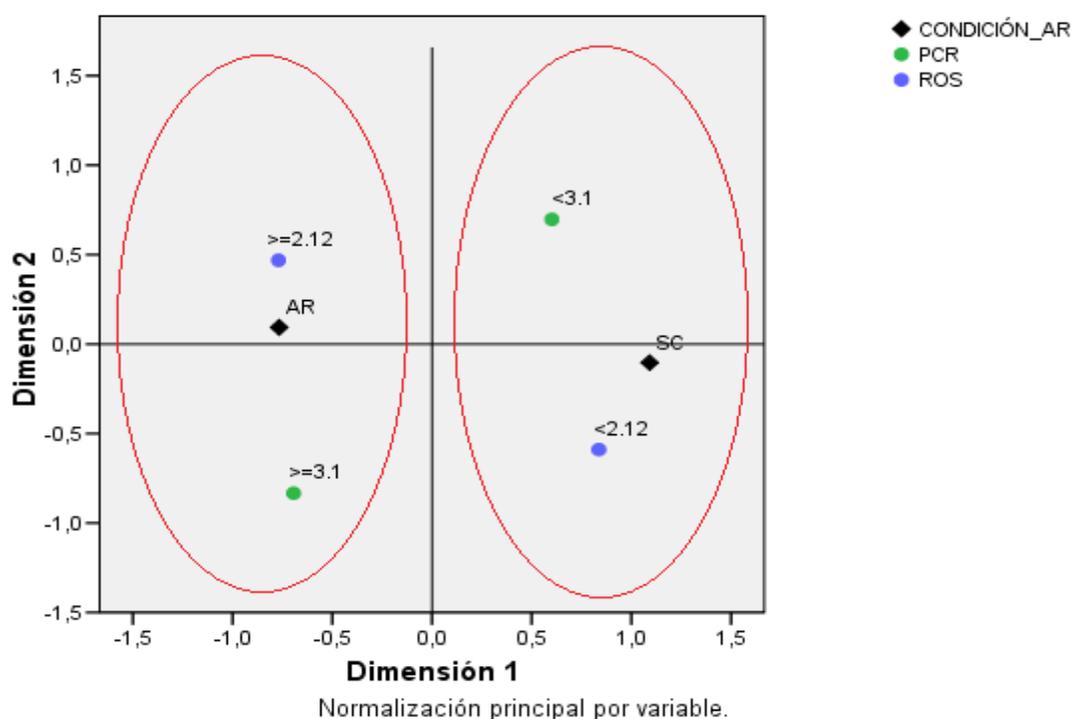


Figura 12: Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con la condición AR de los pacientes

Las líneas rojas denotan la agrupación de variables que se relacionan entre sí, respecto a cada una de las categorías de la variable Condición AR, que lo ha determinado la tabla de discriminación de pesos que se muestra a continuación.

Tabla 16: Medidas de discriminación de pesos del Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con la variable Condición AR.

**Medidas de discriminación**

	Ponderación de la variable	Dimensión		Media
		1	2	
PCR	2	,411	,574	,493
ROS	2	,604	,262	,433
CONDICIÓN_AR	2	,838	,010	,424
Total activo <sup>a</sup>		3,708	1,692	2,700

a. Las ponderaciones de las variables están incorporadas en los estadísticos de Total activo.

Acá solo basta mirar el peso de la variable Condición AR para que se vean claramente las relaciones que existen para cada uno de los grupos de pacientes con respecto a las demás biomarcadores, es decir, dado que el peso para la dimensión 1 es mayor al peso de la dimensión 2, se tiene que el eje X se divide en dos partes, o sea del eje Y hacia la derecha y del eje Y hacia la izquierda.

### 10.3.3 Obtención de Tercer perfil.

Así las relaciones para ambos grupos de pacientes son evidentes. Por lo tanto, así se obtiene entonces el tercer perfil de esta investigación que corresponde a la asociación de **biomarcadores** tanto de inflamación celular como de disfunción endotelial que a simple vista presentan dependencia con respecto a la variable **condición AR**.

Tabla 17: Perfil de todos los biomarcadores asociados con la condición AR de los pacientes.

RELACIÓN DE LOS PACIENTES	
SUJETOS CONTROLES	CON ARTRITIS REUMATOIDE
CON PCR < 3,1 mg/l	CON PCR > ó = 3,1 mg/l
CON ROS < 2,12 mg/l	CON ROS > ó = 2,12 mg/l

## 10.4 Relación y dependencia de *Biomarcadores* con respecto al *Tipo de RCV Global* de los pacientes.

### 10.4.1 Descripción de la población.

Para evaluar las relaciones de estos biomarcadores ahora con respecto al **Tipo de RCV Global** de los pacientes, se realizó un nuevo análisis de correspondencia múltiple.

A continuación se muestra el gráfico con el Análisis de correspondencia múltiple entre todos los **biomarcadores** evaluados en la población incluyendo la variable Tipo de RCV Global que denota a los tres grupos de pacientes, es decir, pacientes con un RCV Global Bajo y Moderado-Alto.

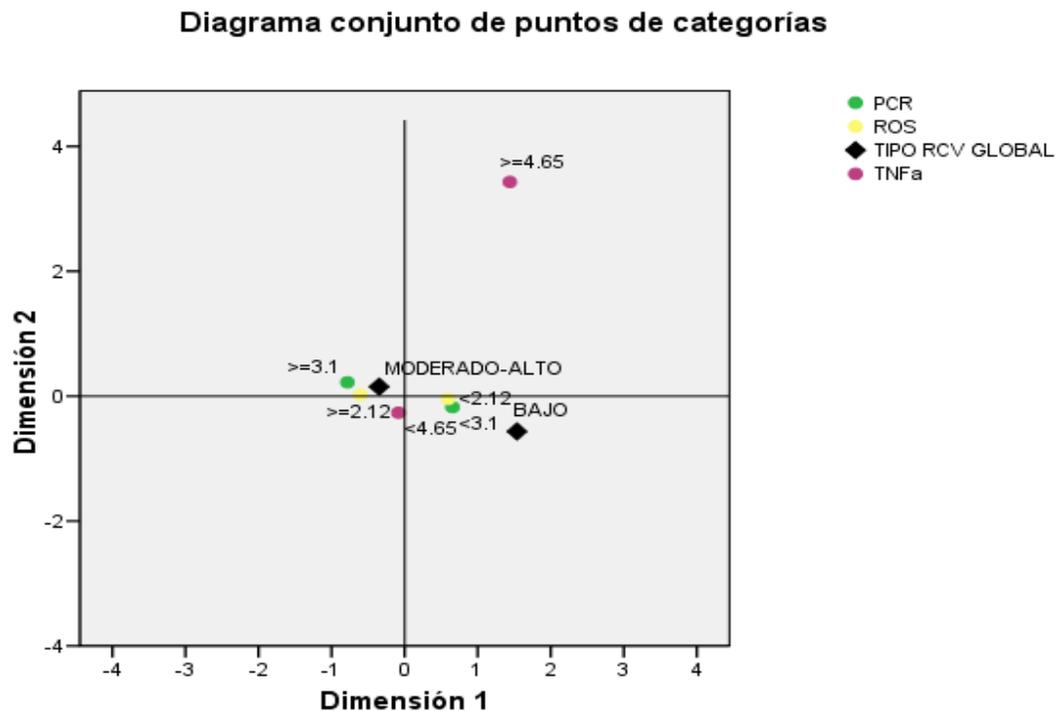


Figura 13: Análisis de Correspondencia múltiple para todos los biomarcadores de acuerdo al Tipo de RCV Global.

Del diagrama de Análisis de correspondencia múltiple, se puede observar que a simple se conforman dos grupos de puntos por separado.

Por ejemplo, se observa que muy similarmente al caso de la asociación de biomarcadores con respecto a la condición AR de los pacientes, personas con un RCV Global bajo se relacionan a tener una PCR < 3,1 mg/L y un ROS < 2.12 mg/l y por su parte si ahora analizamos los puntos cercanos a la población de pacientes con un RCV Global Moderado-Alto se observa que están relacionadas con una PCR Alta, es decir, > ó igual a 3,1 mg/L y un ROS > ó igual a 2.12 mg/l. Por otra parte nuevamente el comportamiento del biomarcador TNFa es indiferente para ambos grupos de pacientes, es decir, se asocia muy similarmente en ambos casos. Además se debe considerar que estas relaciones son bastante pequeñas por estar cerca del origen.

Esto se puede corroborar con la siguiente tabla de discriminación de pesos para todas las variables que se presenta a continuación.

Tabla 18: Medidas de discriminación de pesos del Análisis de Correspondencia múltiple para todos los biomarcadores con respecto al Tipo de RCV Global.

**Medidas de discriminación**

	Ponderación de la variable	Dimensión		Media
		1	2	
TIPO RCV GLOBAL	3	,820	1,004	,912
TNFa	2	,075	,002	,039
PCR	2	,457	,021	,239
ROS	2	,187	,021	,104
Total activo <sup>a</sup>		3,898	3,098	3,498

a. Las ponderaciones de las variables están incorporadas en los estadísticos de Total activo.

## 10.4.2 Test de In dependencia.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

Tabla 19: Resumen de valores -p obtenidos mediante una prueba de independencia entre cada biomarcador y el Tipo de RCG de los pacientes.

BIOMARCADOR	VALOR-P	HIPOTESIS H <sub>0</sub>	DEPENDENCIA
TNFa	0,54 > 0,05	ACEPTADA	NO
PCR	0,007 < 0,05	RECHAZADA	SI
ROS	0,04 < 0,05	RECHAZADA	SI

Al analizar cada uno de los valores de la tabla anterior, se observa que nuevamente dos de los tres biomarcadores presentan un valor  $p < 0,05$ , siendo el mayor de estos de 0,04, lo que significa que dichas variables presentan cierta dependencia con respecto a las diferentes categorías de la variable Tipo de RCG, estas variables son las siguientes:

- Proteína C reactiva (PCR)
- Especies Reactivas de Oxígeno (ROS)

Por lo tanto, deberán considerarse más adelante estos factores, que coinciden con el punto anterior, como las posibles variables que puedan influir en el modelo de regresión logística final.

Adicionalmente, si consideramos solo estas variables en un análisis de correspondencia múltiple, es posible observar más claramente la relación de estas con ambas categorías de la variable Tipo de RCG en la siguiente figura.

### Diagrama conjunto de puntos de categorías

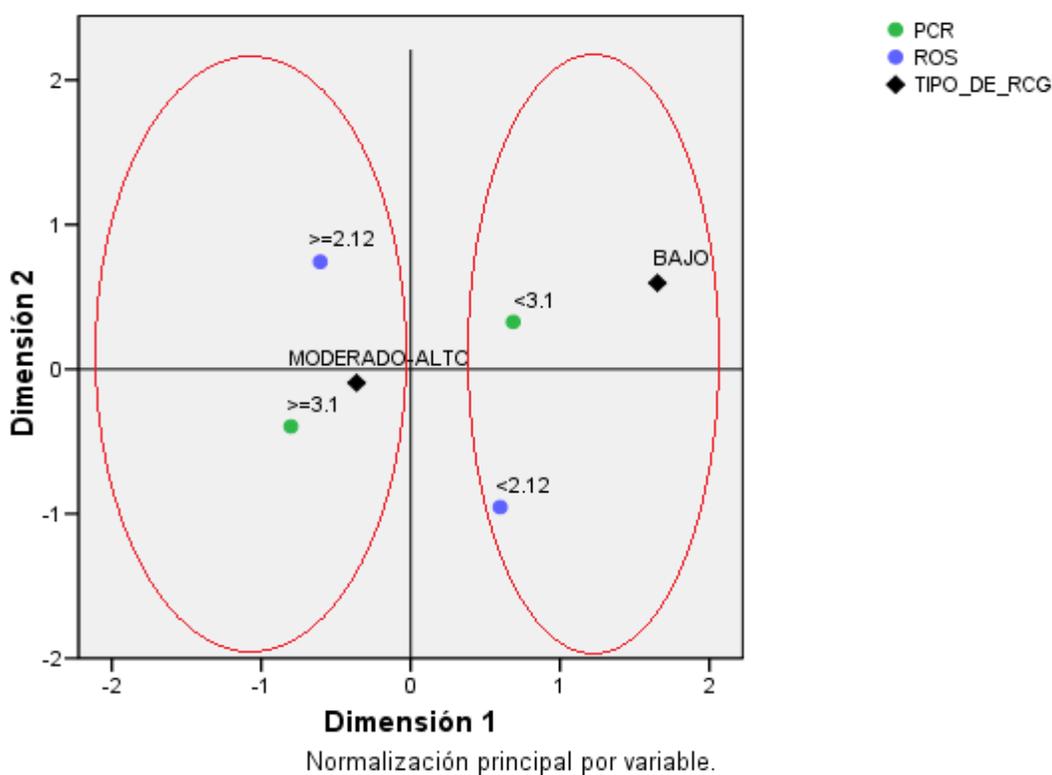


Figura 14: Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con el Tipo de RCG de los pacientes

Las líneas rojas denotan la agrupación de variables que se relacionan entre sí, respecto a cada una de las categorías de la variable Tipo de RCG, que lo ha determinado la tabla de discriminación de pesos que se muestra a continuación.

Tabla 20: Medidas de discriminación de pesos del Análisis de correspondencia múltiple para biomarcadores que dieron dependencia con el Tipo de RCG de los pacientes.

**Medidas de discriminación**

	Ponderación de la variable	Dimensión		Media
		1	2	
PCR	2	,543	,128	,335
ROS	2	,341	,676	,508
TIPO_DE_RCG	2	,615	,074	,345
Total activo <sup>a</sup>		2,998	1,755	2,377

a. Las ponderaciones de las variables están incorporadas en los estadísticos de Total activo.

Acá solo basta mirar el peso de la variable Tipo de RCG para que se vean claramente las relaciones que existen para cada uno de los grupos de pacientes con respecto a las demás biomarcadores, es decir, dado que el peso para la dimensión 1 es mayor al peso de la dimensión 2, se tiene que el eje X se divide en dos partes, o sea del eje Y hacia la derecha y del eje Y hacia la izquierda.

### 10.4.3 Obtención de cuarto perfil.

Así las relaciones para los dos grupos de pacientes son evidentes. Por lo tanto, así se obtiene entonces el cuarto perfil de esta investigación que corresponde a la asociación de los **biomarcadores** que presentan dependencia con respecto a la variable **Tipo de RCV Global**.

Tabla 21: Perfil de todos los biomarcadores asociados con el Tipo de RCV Global de los pacientes.

<b>RELACIÓN DE LOS PACIENTES</b>	
<b>CON RCV GLOBAL BAJO</b>	<b>CON RCV GLOBAL MODERADO - ALTO</b>
CON PCR < 3,1 mg/l	CON PCR > ó = 3,1 mg/l
CON ROS < 2,12 mg/l	CON ROS > ó = 2,12 mg/l

## 10.5. Variables a considerar para mejorar la condición de riesgo cardiovascular de los pacientes en función de los perfiles combinados.

### 10.5.1 Contraste de perfiles con respecto a variables

Ahora al comparar el perfil 1 de los pacientes que corresponde a la asociación de todas las variables dependientes a la condición AR con el perfil 2, que ha asociado a variables que dieron dependencia con el tipo de riesgo cardiovascular global de los pacientes, se observa que estos son muy similares tal como se ve en la figura siguiente, donde se puede analizar relacionando el perfil de los pacientes controles versus los pacientes con RCV Global bajo y el perfil de los pacientes con Artritis Reumatoide versus los pacientes con RCV Global moderado y alto, ya que estas variables presentan una gran dependencia entre ellas.

Tabla 22: Contraste de perfiles con respecto a variables.

RELACIÓN DE LOS PACIENTES		
<b>SUJETOS CONTROLES</b>	<b>CON ARTRITIS REUMATOIDE</b>	<b>P</b>
<b>TIPO DE RCV GLOBAL BAJO</b>	<b>TIPO DE RCV GLOBAL NORMAL Y ALTO</b>	<b>E</b>
EDAD < 54 AÑOS	EDAD > ó = 54 AÑOS	<b>R</b>
SIN HIPERTENSIÓN	CON HIPERTENSIÓN	<b>F</b>
SIN OBESIDAD CENTRAL	CON OBESIDAD CENTRAL	<b>I</b>
SIN TRATAMIENTO ANTIHIPERTENSIVO	CON TRATAMIENTO ANTIHIPERTENSIVO	<b>L</b>
CON PCR < 3 mg/L	CON PCR > ó = 3 mg/L	<b>1</b>
<b>CON RCV GLOBAL BAJO</b>	<b>CON RCV GLOBAL MODERADO - ALTO</b>	
<b>PACIENTES CONTROLES</b>	<b>PACIENTES CON ARTRITIS REUMATOIDE</b>	
SIN ANTECEDENTES FAMILIARES	CON ANTECEDENTES FAMILIARES	<b>P</b>
CON COLESTEROL TOTAL/HDL < 5	CON COLESTEROL TOTAL/HDL > ó = 5	<b>E</b>
SIN DIABETES	CON DIABETES	<b>R</b>
SIN HIPERGLICEMIA	CON HIPERGLICEMIA	<b>F</b>
SIN HIPERTRIGLICERIDEMIA	CON HIPERTRIGLICERIDEMIA	<b>I</b>
SIN HIPERTENSIÓN	CON HIPERTENSIÓN	<b>L</b>
SIN TRATAMIENTO ANTIHIPERTENSIVO	CON TRATAMIENTO ANTIHIPERTENSIVO	<b>2</b>
CON HIPERINSULINEMIA < 14	CON HIPERINSULINEMIA > ó = 14	
CON HOMA-IR < 2,5	CON HOMA-IR > ó = 2,5	
CON IMC NORMAL	CON SOBREPESO Y OBESIDAD	
SIN OBESIDAD CENTRAL	CON OBESIDAD CENTRAL	
CON PCRus < 3 mg/L	CON PCRus > ó = 3 mg/L	

Acá se observa que en el caso de los perfiles de pacientes controles y de RCV Global bajo, en ambos se asocian personas que no tienen Hipertensión, sin una obesidad central significativa, con una PCR baja, es decir, < 3 mg/l y sin un tratamiento antihipertensivo, en comparación al grupo con Artritis Reumatoide y con riesgo moderado-alto que se asocian a personas que si tienen Hipertensión, una obesidad central significativa, una PCR alta, es decir, > ó = 3 mg/l y que siguen un tratamiento antihipertensivo.

## 10.5.2 Contraste de perfiles con respecto a biomarcadores

Ahora al comparar el perfil 3 de los pacientes que corresponde a la asociación de todas los biomarcadores dependientes a la condición AR con el perfil 4, que ha asociado a biomarcadores que dieron dependencia con el tipo de riesgo cardiovascular global de los pacientes, se observa que estos son muy similares tal como se ve en la figura siguiente, donde se puede analizar, igualmente al contraste anterior, relacionando el perfil de los pacientes controles versus los pacientes con RCV Global bajo, y el perfil de los pacientes con Artritis Reumatoide versus los pacientes con RCV Global moderado y alto.

Tabla 23: Contraste de perfiles con respecto a biomarcadores.

RELACIÓN DE LOS PACIENTES		
<b>SUJETOS CONTROLES</b>	<b>CON ARTRITIS REUMATOIDE</b>	<b>P E R F I L 3</b>
<b>TIPO DE RCV GLOBAL BAJO</b>	<b>TIPO DE RCV GLOBAL NORMAL Y ALTO</b>	
CON PCR < 3,1 mg/l	CON PCR > ó = 3,1 mg/l	
CON ROS < 2,12 mg/l	CON ROS > ó = 2,12 mg/l	
<b>CON RCV GLOBAL BAJO</b>	<b>CON RCV GLOBAL MODERADO - ALTO</b>	<b>P E R F I L 4</b>
<b>PACIENTES CONTROLES</b>	<b>PACIENTES CON ARTRITIS REUMATOIDE</b>	
CON PCR < 3,1 mg/l	CON PCR > ó = 3,1 mg/l	
CON ROS < 2,12 mg/l	CON ROS > ó = 2,12 mg/l	

### 10.5.3 Perfiles Combinados

Para identificar **variables** y **biomarcadores** modificables con la finalidad de disminuir un evento cardiovascular a futuro en los pacientes, se analizaron los perfiles combinados obtenidos, es decir, pacientes con RCV bajo / sujetos controles y pacientes con riesgo moderado-alto/con artritis reumatoide tal como se muestra a continuación:

Tabla 24: Contraste de perfiles combinados con respecto a variables y biomarcadores.

RELACIÓN DE LOS PACIENTES		
<b>SUJETOS CONTROLES</b>	<b>CON ARTRITIS REUMATOIDE</b>	P E R F I L 1
TIPO DE RCV GLOBAL BAJO	TIPO DE RCV GLOBAL NORMAL Y ALTO	
EDAD < 54 AÑOS	EDAD > ó = 54 AÑOS	
SIN HIPERTENSIÓN	CON HIPERTENSIÓN	
SIN OBESIDAD CENTRAL	CON OBESIDAD CENTRAL	
SIN TRATAMIENTO ANTIHIPERTENSIVO	CON TRATAMIENTO ANTIHIPERTENSIVO	P E R F I L 2
CON PCR < 3 mg/L	CON PCR > ó = 3 mg/L	
<b>CON RCV GLOBAL BAJO</b>	<b>CON RCV GLOBAL MODERADO - ALTO</b>	
<b>PACIENTES CONTROLES</b>	<b>PACIENTES CON ARTRITIS REUMATOIDE</b>	
SIN ANTECEDENTES FAMILIARES	CON ANTECEDENTES FAMILIARES	
CON COLESTEROL TOTAL/HDL < 5	CON COLESTEROL TOTAL/HDL > ó = 5	P E R F I L 3
SIN DIABETES	CON DIABETES	
SIN HIPERGLICEMIA	CON HIPERGLICEMIA	
SIN HIPERTRIGLICERIDEMIA	CON HIPERTRIGLICERIDEMIA	
SIN HIPERTENSIÓN	CON HIPERTENSIÓN	
SIN TRATAMIENTO ANTIHIPERTENSIVO	CON TRATAMIENTO ANTIHIPERTENSIVO	P E R F I L 4
CON HIPERINSULINEMIA < 14	CON HIPERINSULINEMIA > ó = 14	
CON HOMA-IR < 2,5	CON HOMA-IR > ó = 2,5	
CON IMC NORMAL	CON SOBREPESO Y OBESIDAD	
SIN OBESIDAD CENTRAL	CON OBESIDAD CENTRAL	
CON PCRus < 3 mg/L	CON PCRus > ó = 3 mg/L	
<b>SUJETOS CONTROLES</b>	<b>CON ARTRITIS REUMATOIDE</b>	P E R F I L 3
TIPO DE RCV GLOBAL BAJO	TIPO DE RCV GLOBAL NORMAL Y ALTO	
CON PCR < 3,1 mg/l	CON PCR > ó = 3,1 mg/l	
CON ROS < 2,12 mg/l	CON ROS > ó = 2,12 mg/l	
<b>CON RCV GLOBAL BAJO</b>	<b>CON RCV GLOBAL MODERADO - ALTO</b>	
<b>PACIENTES CONTROLES</b>	<b>PACIENTES CON ARTRITIS REUMATOIDE</b>	P E R F I L 3
CON PCR < 3,1 mg/l	CON PCR > ó = 3,1 mg/l	
CON ROS < 2,12 mg/l	CON ROS > ó = 2,12 mg/l	

En el caso de las **variables** que se pueden modificar para mejorar la condición de riesgo de los pacientes, son la **hipertensión**, un **tratamiento antihipertensivo**, que está totalmente relacionado a la variable anterior, la **obesidad central** y la **Proteína C reactiva**, que son las variables que más determinan una condición de artritis de otra y también una condición de tipo de riesgo cardiovascular global de las otras.

En el caso de los **biomarcadores** que más determinan una condición de artritis de otra y también una condición de tipo de riesgo cardiovascular global de las otras, son las **especies reactivas de oxígeno** y la **Proteína C reactiva**, que en este caso está considerada como biomarcador y corrobora que anteriormente como variable también es determinante en las condiciones para este grupo pacientes.

También se puede ver de los perfiles que la edad es una variable determinante en que una persona tenga o no la condición de artritis reumatoide.

Por otra parte los antecedentes familiares, la diabetes, el colesterol total/HDL, la hiperglicemia, la hipertrigliceridemia, la hiperinsulinemia, el HOMA-IR y el IMC de los pacientes es determinante en que una persona tenga un riesgo bajo, moderado o alto; y no así es resto de variables que no son parte de estos perfiles.

## 10.6 Comparación del **Riesgo Cardiovascular Global** entre sujetos controles y pacientes con AR considerando variables

Para comparar el valor real cuantitativo del riesgo cardiovascular global entre ambas poblaciones de pacientes se realizó un diseño en bloques completamente aleatorio, considerando las variables que anteriormente dieron dependencia con la variable Condición AR, las cuales pueden influir en la interpretación de los resultados finales.

Para este análisis de varianza se propusieron las siguientes hipótesis:

H0: La media del riesgo cardiovascular global para pacientes con Artritis reumatoide es igual a la media del riesgo cardiovascular global para sujetos controles.

H1: La media del riesgo cardiovascular global para pacientes con Artritis reumatoide es distinta a la media del riesgo cardiovascular global para sujetos controles.

Acá se pretende determinar el efecto que tienen los dos grupos de pacientes, es decir, personas que efectivamente tienen Artritis reumatoide y sujetos controles, sobre el valor real de su riesgo cardiovascular Global, considerando una serie de otras variables que no interesa determinar si hay diferencia entre sus categorías (bloques), pero sí, probablemente puedan influir en estas comparaciones. Es por esto que se decide realizar este diseño de experimentos y no otro.

Observación: Acá es importante mencionar, que para este análisis inicialmente se consideraron los factores que dieron dependencia con la variable Condición AR (edad, hipertensión, tratamiento antihipertensivo, PCRus y obesidad central).

La Tabla ANOVA resultante de este análisis se muestra a continuación:

Tabla 25: Primera ANOVA para la comparación del RCV Global.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: RIESGO GLOBAL

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	3490,494 <sup>a</sup>	7	498,642	144,235	,000
CONDICIÓNAR	314,620	2	157,310	45,503	,000
EDAD	9,928	1	9,928	2,872	,094
OBESIDADCENTRAL	130,794	1	130,794	37,833	,000
HIPERTENCIÓN	28,271	1	28,271	8,177	,005
TRATAMIENTOANTIHIPERTENSIVO	7,071	1	7,071	2,045	,156
PCRus	15,594	1	15,594	4,511	,037
Error	297,316	86	3,457		
Total	3787,810	93			

a. R cuadrado = ,922 (R cuadrado corregida = ,915)

De todas las variables (bloques), solo 3 dieron una diferencia significativa con un 95 % de confianza (valor-p < 0.05), que fueron obesidad central, Hipertensión y PCRus, por lo tanto el análisis de varianza final que se muestra a continuación se realizo solo considerando estas tres variables como bloqueos.

Tabla 26: ANOVA final para la comparación del RCV Global.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: RIESGO GLOBAL

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	3474,022 <sup>a</sup>	5	694,804	194,854	,000
CONDICIÓNAR	879,912	2	439,956	123,383	,000
OBESIDADCENTRAL	133,633	1	133,633	37,476	,000
HIPERTENSIÓN	51,797	1	51,797	14,526	,000
PCRus	23,111	1	23,111	6,481	,013
Error	313,788	88	3,566		
Total	3787,810	93			

a. R cuadrado = ,917 (R cuadrado corregida = ,912)

Acá se corrobora que los bloques que quedan están bien aplicados ya que siguen siendo muy significativos, es decir, que presentan diferencias estadísticamente significativas en sus categorías con respecto al riesgo cardiovascular global, pero recordemos que no son estas comparaciones las de interés, sino las diferencias de la variable CONDICION AR, que en este caso también es significativa ya que presenta una significación (valor-p) aproximadamente cero, es decir, se rechaza la hipótesis nula que antes se planteó y podemos afirmar que efectivamente hay diferencias estadísticamente significativas de riesgo cardiovascular global entre pacientes con artritis reumatoide y sujetos controles, pero además se debe considerar que hay otras tres variables ( bloques) que influyen en la comparación de estos dos grupos de pacientes, por lo que se hace muy recomendable separar previamente para efecto de estas comparaciones entre las categorías de la variable hipertensión (si/no), obesidad central (si/no) y PCRus (si/no).

## 10.7 Comparación de **Biomarcadores** entre sujetos controles y pacientes con AR considerando variables

Primero cabe destacar que en este punto se considerará a los biomarcadores como variables respuestas o dependientes Y.

## 10.7.1 Biomarcadores de Inflamación Celular

Para comparar el valor real cuantitativo de estos tipos de Biomarcadores entre ambas poblaciones de pacientes se realizó nuevamente un diseño en bloques completamente aleatorio, pero antes se debieron realizar los test de independencia correspondientes entre cada biomarcador de inflamación celular y todas las variables o parámetros para determinar cuáles de estas pueden influir en la interpretación de los resultados finales y se deban considerar en estos análisis de varianza.

### 10.7.1.1 Factor de Necrosis Tumoral (TNF $\alpha$ )

Las hipótesis asociadas, en este caso a todas las pruebas de independencia realizadas, son las que se exponen en la metodología de este informe.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

Tabla 27: Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el biomarcador TNFa.

VARIABLE	VALOR-P	HIPOTESIS H <sub>0</sub>	DEPENDENCIA
Tipo de RCG	0,68 > 0,05	ACEPTADA	NO
Condición AR	0,41 > 0,05	ACEPTADA	NO
Edad	0,98 > 0,05	ACEPTADA	NO
Género	0,48 > 0,05	ACEPTADA	NO
IMC	0,65 > 0,05	ACEPTADA	NO
Fumador	0,17 > 0,05	ACEPTADA	NO
Hipertrigliceridemia	0,48 > 0,05	ACEPTADA	NO
Hipercolesterolemia	0,15 > 0,05	ACEPTADA	NO
Hiperglicemia	0,68 > 0,05	ACEPTADA	NO
Colesterol LDL alto	0,87 > 0,05	ACEPTADA	NO
Colesterol HDL bajo	0,44 > 0,05	ACEPTADA	NO
Colesterol VLDL alto	0,30 > 0,05	ACEPTADA	NO
Colesterol TOTAL /HDL	0,88 > 0,05	ACEPTADA	NO
Obesidad central	0,07 > 0,05	ACEPTADA	NO
Hipertensión Arterial	0,71 > 0,05	ACEPTADA	NO
Tratamiento Antihipertensivo	0,49 > 0,05	ACEPTADA	NO
Diabetes	0,35 > 0,05	ACEPTADA	NO
Creatinina alta	0,44 > 0,05	ACEPTADA	NO
Homa-IR	0,16 > 0,05	ACEPTADA	NO
Hiperinsulinemia	0,11 > 0,05	ACEPTADA	NO
<b>Antecedentes familiares de infarto</b>	<b>0,02 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
PCRus	0,38 > 0,05	ACEPTADA	NO
TFG	0,44 > 0,05	ACEPTADA	NO
Biomarcador PCR	0,38 > 0,05	ACEPTADA	NO
Biomarcador ROS	0,30 > 0,05	ACEPTADA	NO

Al analizar cada uno de los valores de la tabla anterior, se observa que solo una variable presenta un valor-p de  $0,02 < 0,05$ , lo que significa que dicha variable presenta cierta dependencia con respecto a las diferentes categorías de este biomarcador, esta variable es la siguiente:

- **Antecedentes Familiares**

Por lo tanto, deberá considerarse este factor como la posible variable que pueda influir en la comparación de este biomarcador, entre la población con Artritis Reumatoide y sujetos controles.

Así para este análisis de varianza se propusieron las siguientes hipótesis:

H0: La media del TNFa para pacientes con Artritis reumatoide es igual a la media del TNFa para sujetos controles.

H1: La media del TNFa para pacientes con Artritis reumatoide es distinta a la media del TNFa para sujetos controles.

**Acá se pretende determinar el efecto que tienen los dos grupos de pacientes, es decir, personas que efectivamente tienen Artritis reumatoide y sujetos controles, sobre el valor real del TNFa, considerando otras variables que no interesa determinar si hay diferencia entre sus categorías (bloques), pero sí, probablemente puedan influir en estas comparaciones. Es por esto que se decide realizar este diseño de experimentos y no otro.**

La Tabla ANOVA resultante de este análisis se muestra a continuación:

Tabla 28: ANOVA para la comparación del TNFa.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: TNFa

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	6141,148 <sup>a</sup>	3	2047,049	5,096	,003
CONDICIÓNAR	22,258	1	22,258	,055	,814
ANTECEDENT ESFAMILIARES	715,729	1	715,729	1,782	,185
Error	38163,162	95	401,717		
Total	44304,310	98			

a. R cuadrado = ,139 (R cuadrado corregida = ,111)

Se puede ver que el bloque no es significativo al 95% de confianza, ya que su valor-p es > a 0.05, por lo tanto el análisis de varianza final que se muestra a continuación se realizo sin ninguna otra variable que pueda influir (bloque), o sea, solo se determinarán las diferencias del biomarcador entre cada una de las categorías de la variable condición AR.

Tabla 29: Simplificación de ANOVA para la comparación del TNFa.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: TNFa

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	5425,700 <sup>a</sup>	2	2712,850	6,766	,002
CONDICIÓNAR	5425,700	2	2712,850	6,766	,002
Error	38894,610	97	400,975		
Total	44320,310	99			

a. R cuadrado = ,122 (R cuadrado corregida = ,104)

Recordemos que son estas comparaciones las de interés, es decir, las diferencias que se presentan en cuanto a este biomarcador para la variable CONDICION AR, que en este caso es significativa ya que presenta una significación aproximadamente cero (valor-p < 0.05), es decir, se rechaza la hipótesis nula que antes se planteó y podemos afirmar que efectivamente hay diferencias estadísticamente significativas de TNFa entre pacientes con artritis reumatoide y sujetos controles.

### 10.7.1.2 Proteína C Reactiva Ultrasensible (PCRus)

Las hipótesis asociadas, en este caso a todas las pruebas de independencia realizadas, son las que se exponen en la metodología de este informe.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

Tabla 30: Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el biomarcador PCR.

VARIABLE	VALOR-P	HIPOTESIS Ho	DEPENDENCIA
<b>Tipo de RCG</b>	<b>0,001 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Condición AR</b>	<b>0,000 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Edad</b>	<b>0,005 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Género	0,27 > 0,05	ACEPTADA	NO
IMC	0,28 > 0,05	ACEPTADA	NO
Fumador	0,81 > 0,05	ACEPTADA	NO
Hipertrigliceridemia	0,72 > 0,05	ACEPTADA	NO
Hipercolesterolemia	0,48 > 0,05	ACEPTADA	NO
<b>Hiperglicemia</b>	<b>0,04 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Colesterol LDL alto	0,86 > 0,05	ACEPTADA	NO
<b>Colesterol HDL bajo</b>	<b>0,02 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Colesterol VLDL alto	0,67 > 0,05	ACEPTADA	NO
<b>Colesterol TOTAL /HDL</b>	<b>0,03 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Obesidad central</b>	<b>0,01 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Hipertensión Arterial</b>	<b>0,004 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Tratamiento Antihipertensivo</b>	<b>0,009 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
<b>Diabetes</b>	<b>0,01 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Creatinina alta	0,09 > 0,05	ACEPTADA	NO
<b>Homa-IR</b>	<b>0,001 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Hiperinsulinemia	0,052 > 0,05	ACEPTADA	NO
Antecedentes familiares de infarto	0,77 > 0,05	ACEPTADA	NO
TFG	0,09 > 0,05	ACEPTADA	NO
Biomarcador TNFa	0,38 > 0,05	ACEPTADA	NO
Biomarcador ROS	0,07 > 0,05	ACEPTADA	NO

Al analizar cada uno de los valores de la tabla anterior, se observa que once variables presentan un valor p < 0,05, siendo el mayor de estos de 0,04, lo que significa que dichas variables presentan gran dependencia con respecto a las diferentes categorías del biomarcador PCR, estas variables son las siguientes:

- Tipo de RCG
- Condición AR
- Hiperglicemia
- Obesidad central
- Colesterol HDL bajo
- Colesterol TOTAL /HDL
- Hipertensión Arterial
- Tratamiento Antihipertensivo
- Edad
- Diabetes
- Homa-IR

Por lo tanto, deberán considerarse estos factores como las posibles variables que puedan influir en la comparación de este biomarcador, entre la población con Artritis Reumatoide y sujetos controles.

Así para este análisis de varianza se propusieron las siguientes hipótesis:

H0: La media de la PCR para pacientes con Artritis reumatoide es igual a la media de la PCR para sujetos controles.

H1: La media de la PCR para pacientes con Artritis reumatoide es distinta a la media de la PCR para sujetos controles.

**Acá se pretende determinar el efecto que tienen los dos grupos de pacientes, es decir, personas que efectivamente tienen Artritis reumatoide y sujetos controles, sobre el valor real de la PCR, considerando otras variables que no interesa determinar si hay diferencia entre sus categorías (bloques), pero sí, probablemente puedan influir en estas comparaciones. Es por esto que se decide realizar este diseño de experimentos y no otro.**

Tabla 31: ANOVA para la comparación de la PCR.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: PCR

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	7588,097 <sup>a</sup>	14	542,007	5,942	,000
CONDICIÓNAR	210,613	1	210,613	2,309	,133
TIPORCVGLOBAL	462,863	2	231,432	2,537	,086
EDAD	614,524	2	307,262	3,369	,040
OBESIDADCENTRAL	161,234	1	161,234	1,768	,188
HIPERTENCIÓN	161,802	1	161,802	1,774	,187
TRATAMIENTOANTIHIPE RTENSIVO	37,480	1	37,480	,411	,523
DIABETES	446,456	1	446,456	4,895	,030
HIPERGLICEMIA	171,812	1	171,812	1,884	,174
COLESTEROLHDLBAJO	1026,616	1	1026,616	11,255	,001
COLESTEROLTOTALHD L	287,636	1	287,636	3,153	,080
HOMAIR	18,603	1	18,603	,204	,653
Error	6841,075	75	91,214		
Total	14429,172	89			

a. R cuadrado = ,526 (R cuadrado corregida = ,437)

De todas las variables (bloques), solo 3 dieron una diferencia significativa con un 95 % de confianza (valor-p < 0.05), que fueron edad, diabetes y colesterol HDL bajo, por lo tanto el análisis de varianza que se muestra a continuación se realizó solo considerando estas tres variables como bloques.

Tabla 32: Primera simplificación de ANOVA para la comparación de la PCR.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: PCR

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	6305,263 <sup>a</sup>	6	1050,877	11,161	,000
CONDICIÓNAR	82,211	1	82,211	,873	,353
EDAD	445,994	2	222,997	2,368	,099
DIABETES	125,270	1	125,270	1,330	,252
COLESTEROLHDLBAJO	1500,726	1	1500,726	15,939	,000
Error	8567,935	91	94,153		
Total	14873,198	97			

a. R cuadrado = ,424 (R cuadrado corregida = ,386)

Pero de esta última tabla nos podemos dar cuenta que la diabetes deja de ser una variable significativa, por lo que se extrae del modelo y volvemos a realizar una nueva tabla ANOVA, tal como se muestra a continuación:

Tabla 33: Segunda simplificación de ANOVA para la comparación de la PCR.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: PCR

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	6179,993 <sup>a</sup>	5	1235,999	13,081	,000
CONDICIÓNAR	136,218	1	136,218	1,442	,233
EDAD	358,201	2	179,100	1,895	,156
COLESTEROLHDLBAJO	1541,299	1	1541,299	16,312	,000
Error	8693,205	92	94,491		
Total	14873,198	97			

a. R cuadrado = ,416 (R cuadrado corregida = ,384)

Nuevamente existe una variable que deja de ser significativa en el modelo, por lo que se vuelve a realizar el diseño de experimentos, ahora solo considerando una variable como bloque, el colesterol HDL de los pacientes.

Tabla 34: ANOVA final para la comparación de la PCR.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: PCR

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	5899,452 <sup>a</sup>	3	1966,484	20,619	,000
CONDICIÓNAR	381,515	1	381,515	4,000	,048
COLESTEROLHDLBAJO	1835,392	1	1835,392	19,244	,000
Error	9060,422	95	95,373		
Total	14959,874	98			

a. R cuadrado = ,394 (R cuadrado corregida = ,375)

Acá se corrobora que el bloque (Colesterol HDL Bajo) que queda está bien aplicado ya que sigue siendo muy significativo, es decir, que presenta diferencias estadísticamente significativas en sus categorías con respecto a la PCR, pero recordemos que no es esta comparación la de interés, sino las diferencias de la variable CONDICION AR, que en este caso también es significativa ya que presenta una significación (valor-p) menor a 0.05, es decir, se rechaza la hipótesis nula que antes se planteó y podemos afirmar que efectivamente hay diferencias estadísticamente significativas de PCR entre pacientes con artritis reumatoide y sujetos controles con un 95% de confianza, pero además se debe considerar que hay otra variable ( bloque) que influye en la comparación de estos dos grupos de pacientes, por lo que se hace muy recomendable separar previamente para efecto de estas comparaciones entre las categorías de la variable colesterol HDL bajo(SI/NO).

### 10.7.1.3 Interluquina 6 (IL-6)

Por motivos de tiempo, no se alcanzaron a reclutar todos los datos necesarios para desarrollar los análisis similares a los que se han expuesto en este informe para este biomarcador.

## 10.7.2 Biomarcadores de Disfunción Endotelial

Para el caso de estos biomarcadores el procedimiento y las hipótesis también son similares, por lo que todo se resume en los siguientes gráficos y tablas.

### 10.7.2.1 Especies Reactivas de Oxígeno (ROS)

Las hipótesis asociadas, en este caso a todas las pruebas de independencia realizadas, son las que se exponen en la metodología de este informe.

En la siguiente tabla se muestra el resumen de los valores-p para cada una de las pruebas realizadas en este punto.

Tabla 35: Resumen de valores -p obtenidos mediante una prueba de independencia entre cada variable y el biomarcador ROS.

VARIABLE	VALOR-P	HIPOTESIS H <sub>0</sub>	DEPENDENCIA
Tipo de RCG	0,11 > 0,05	ACEPTADA	NO
<b>Condición AR</b>	<b>0,000 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Edad	0,45 > 0,05	ACEPTADA	NO
Género	0,51 > 0,05	ACEPTADA	NO
IMC	0,68 > 0,05	ACEPTADA	NO
Fumador	0,90 > 0,05	ACEPTADA	NO
Hipertrigliceridemia	0,67 > 0,05	ACEPTADA	NO
Hipercolesterolemia	0,13 > 0,05	ACEPTADA	NO
Hiperglicemia	0,91 > 0,05	ACEPTADA	NO
Colesterol LDL alto	0,40 > 0,05	ACEPTADA	NO
Colesterol HDL bajo	0,32 > 0,05	ACEPTADA	NO
Colesterol VLDL alto	0,81 > 0,05	ACEPTADA	NO
Colesterol TOTAL /HDL	0,17 > 0,05	ACEPTADA	NO
<b>Obesidad central</b>	<b>0,02 &lt; 0,05</b>	<b>RECHAZADA</b>	<b>SI</b>
Hipertensión Arterial	0,17 > 0,05	ACEPTADA	NO
Tratamiento Antihipertensivo	0,056 > 0,05	ACEPTADA	NO
Diabetes	0,70 > 0,05	ACEPTADA	NO
Creatinina alta	0,55 > 0,05	ACEPTADA	NO
Homa-IR	0,56 > 0,05	ACEPTADA	NO
Hiperinsulinemia	0,12 > 0,05	ACEPTADA	NO
Antecedentes familiares de infarto	0,16 > 0,05	ACEPTADA	NO
PCRus	0,07 > 0,05	ACEPTADA	NO
TFG	0,07 > 0,05	ACEPTADA	NO
Biomarcador TNFa	0,30 > 0,05	ACEPTADA	NO
Biomarcador PCRus	0,07 > 0,05	ACEPTADA	NO

Al analizar cada uno de los valores de la tabla anterior, se observa que dos variables presentan un valor  $p < 0,05$ , siendo el mayor de estos de 0,02, lo que significa que dichas variables presentan gran dependencia con respecto a las diferentes categorías del biomarcador ROS, estas variables son las siguientes:

- Condición AR
- Obesidad central

Por lo tanto, deberán considerarse estos factores como las posibles variables que puedan influir en la comparación de este biomarcador, entre la población con Artritis Reumatoide y sujetos controles.

Así para este análisis de varianza se propusieron las siguientes hipótesis:

H0: La media del ROS para pacientes con Artritis reumatoide es igual a la media de la ROS para sujetos controles.

H1: La media de la ROS para pacientes con Artritis reumatoide es distinta a la media de la ROS para sujetos controles.

**Acá se pretende determinar el efecto que tienen los dos grupos de pacientes, es decir, personas que efectivamente tienen Artritis reumatoide y sujetos controles, sobre el valor real del ROS, considerando otras variables que no interesa determinar si hay diferencia entre sus categorías (bloques), pero sí, probablemente puedan influir en estas comparaciones.**

Es por esto que se decide realizar este diseño de experimentos y no otro.

Tabla 36: ANOVA para la comparación del ROS.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: ROS

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	1110,200 <sup>a</sup>	3	370,067	21,739	,000
CONDICIÓNAR	58,925	1	58,925	3,461	,066
OBESIDADCENTRAL	16,859	1	16,859	,990	,322
Error	1515,084	89	17,023		
Total	2625,284	92			

a. R cuadrado = ,423 (R cuadrado corregida = ,403)

Se puede ver que el bloque no es significativo al 95% de confianza, ya que su valor-p es > a 0.05, por lo tanto el análisis de varianza final que se muestra a continuación se realizó sin ninguna otra variable que pueda influir (bloque), o sea, solo se determinarán las diferencias del biomarcador entre cada una de las categorías de la variable condición AR.

Tabla 37: Simplificación de ANOVA para la comparación de la ROS.

**Pruebas de los efectos inter-sujetos**

Variable dependiente: ROS

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo	1132,674 <sup>a</sup>	2	566,337	34,122	,000
CONDICIÓNAR	1132,674	2	566,337	34,122	,000
Error	1543,575	93	16,598		
Total	2676,250	95			

a. R cuadrado = ,423 (R cuadrado corregida = ,411)

Recordemos que son estas comparaciones las de interés, es decir, las diferencias que se presentan en cuanto a este biomarcador para la variable CONDICION AR, que en este caso es muy significativa ya que presenta una significación aproximadamente cero (valor- $p < 0.05$ ), es decir, se rechaza la hipótesis nula que antes se planteó y podemos afirmar que efectivamente hay diferencias estadísticamente significativas de ROS entre pacientes con artritis reumatoide y sujetos controles.

### 10.7.2.2 Oxido Nítrico (NO)

También por motivos de tiempo, no se alcanzaron a reclutar todos los datos necesarios para desarrollar los análisis similares a los que se han expuesto en este informe para este biomarcador.

## 10.8 Modelos de Regresión Logística

Vamos a ir introduciendo los elementos de esta técnica a través de las variables que son parte de este estudio.

Para el modelo de regresión que aplicaremos en primera instancia, porque es muy interesante para este estudio, queremos que el tratamiento y pronóstico del **Tipo de riesgo cardiovascular Global** dependa solo del nivel de los distintos **biomarcadores** y de la **condición de Artritis** de los pacientes.

Sabemos que para aquellos pacientes con un riesgo muy alto se hace necesario tomar medidas más agresivas y realizar inmediatamente alguna intervención.

Si en función de una serie de pruebas médicas, que están representadas por los biomarcadores que se consideraron en este estudio, se pudiera determinar la probabilidad de que dichas personas tengan o no un alto riesgo cardiovascular global, se ahorraría tiempo y molestias a los pacientes.

Así como ya sabemos los datos a analizar pertenecen a una muestra original de 101 pacientes, de los cuales se han dejado fuera algunos por motivos de muchos o importantes datos faltantes en ciertos casos.

A cada paciente se le han medido las siguientes variables o características, las cuales serán parte en un principio de este modelo de regresión logística.

Variable Dependiente o Respuesta:

- Tipo de riesgo cardiovascular Global ( Tipo de RCV Global)

Variables Independientes:

- Factor de Necrosis Tumoral (TNFa)
- Proteína C reactiva (PCR)
- Especies Reactivas de oxígeno (ROS)
- Condición de Artritis de los pacientes (Condición AR)

**Observación:** La variable respuesta fue recodificada en solo dos categorías (ANEXO 3).

## 10.8.1 Estadísticos descriptivos

A continuación se muestran los estadísticos descriptivos de las variables involucradas en el análisis. Es de particular importancia asegurarse que las variables del tipo ausencia/presencia estén codificadas como cero y uno.

Tabla 38: Estadísticos descriptivos del modelo logístico de prueba.

**TIPO\_DE\_RCG**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
,00	19	19,6	19,6	19,6
1,00	78	80,4	80,4	100,0
Total	97	100,0	100,0	

**CONDICION\_AR**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
,00	55	56,7	56,7	56,7
1,00	42	43,3	43,3	100,0
Total	97	100,0	100,0	

**Estadísticos descriptivos**

	N	Mínimo	Máximo	Media	Dev. típ.
TNFa	97	2,70	187,30	7,4443	20,13309
PCR	97	,19	87,20	5,6566	10,98717
ROS	97	,12	31,08	3,3453	4,12316
N válido (según lista)	97				

Como vemos en todos los casos tanto para las variables categóricas como cuantitativas incluidas en el modelo, el número total de observaciones igual a 97 está correctamente reconocido por el software, al igual que la codificación de 0 y 1 en el caso de las variables categóricas.

Con las variables anteriores vamos a intentar construir un modelo de regresión logística.

## 10.8.2 Coeficientes estimados del modelo logístico

Tabla 39: Parámetros estimados del modelo logístico de prueba.

Estimated Regression Model (Maximum Likelihood)			
Parameter	Estimate	Standard Error	Estimated Odds Ratio
CONSTANT	0,764071	0,379277	
TNFa_	-0,00672946	0,011333	0,993293
PCR_	-0,0209341	0,0228997	0,979283
ROS_	0,0143173	0,0776497	1,01442
CONDICION_AR=0	1,75986	0,636684	5,81164

La salida del software muestra los resultados al ajustar un modelo de regresión logística para describir la relación entre TIPO\_DE\_RCG y 4 variables independientes.

La ecuación del modelo ajustado es

$$\text{TIPO\_DE\_RCG} = \exp(\text{zeta}) / (1 + \exp(\text{zeta})), \quad \text{donde}$$

$$z = 0.764071 - 0,00672946 \cdot \text{TNFa} - 0,0209341 \cdot \text{PCR} + 0,0143173 \cdot \text{ROS} + 1,75986 \cdot \text{CONDICION AR}$$

Tabla 40: Análisis de desviación del modelo logístico de prueba.

Analysis of Deviance			
Source	Deviance	Df	P-Value
Model	10,1031	4	0,0387
Residual	85,8555	92	0,6606
Total (corr.)	95,9587	96	

Debido a que el P-valor para el modelo en la tabla ANOVA es menor que 0,05, existe al menos una variable que influye en el RCG de los pacientes con un nivel de confianza del 95%. Además, el P-valor para los residuales es mayor que o igual a 0,10, lo que indica que el modelo no es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 90% o superior.

### 10.8.3 Medidas de bondad y ajuste

Tabla 41: Medidas de bondad y ajuste del modelo logístico de prueba.

**Resumen de los modelos**

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	85,856 <sup>a</sup>	,099	,157

a. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Acá se muestra que el porcentaje de la desviación explicada en TIPO\_DE\_RCG por el modelo es igual a 15,7% según el coeficiente de Nagelkerke. Esta estadística es similar al habitual  $R^2$  estadístico tradicional.

Aunque estos coeficientes tratan de medir la variabilidad explicada, en general, van a ser mucho más bajos que en regresión lineal y deberán de ser complementados con otras medidas de bondad de ajuste como la que se muestra a continuación.

Tabla 42: Prueba de Hosmer y Lemeshow del modelo logístico de prueba.

**Prueba de Hosmer y Lemeshow**

Paso	Chi-cuadrado	gl	Sig.
1	8,946	8	,347

Esta prueba determina si la función logística adecuada se ajusta a los datos observados. Debido a que el valor P es mayor o igual a 0,10; no hay ninguna razón para rechazar la idoneidad del modelo logístico con un nivel de confianza del 90% o superior.

**Si bien este modelo de regresión logística no es tan malo, tampoco es un buen modelo, ya sea por un  $R^2$  pequeño y principalmente porque solo la variable Condición AR es significativa. Esto implica que las predicciones que se realicen no serán buenas, pero igualmente se consideran a continuación solo con fines académicos.**

## 10.8.4 Estimación de probabilidades

Con los coeficientes ya estimados se puede predecir la probabilidad de que una persona tenga un tipo de riesgo cardiovascular global moderado-alto, afectado por el comportamiento de los diversos biomarcadores y de su condición de artritis, simplemente construyendo la siguiente función de probabilidad:

$$P(Y = TIPO\_DE\_RCG = 1 / X) = \frac{1}{1 + e^{-z}} \quad , \text{ donde:}$$

$$z = 0.764071 - 0.00672946 * TNFa - 0.0209341 * PCR + 0.014517 * ROS + 1.75986 * CONDICIÓN\_AR$$

A la vista de esta ecuación, podemos estimar la probabilidad de que un paciente con determinadas características tenga **un riesgo cardiovascular global moderado-alto ó un riesgo cardiovascular bajo**, a través de los siguientes ejemplos ilustrativos.

## 10.8.4 Ejemplos Ilustrativos

Como **primer ejemplo ilustrativo**, la probabilidad de que un paciente, con una condición de artritis reumatoide **positiva**, con un nivel de TNFa de **180 pg/mL**, un nivel de PCR de **2mg/L** y un nivel de ROS de **20 IF/ ug proteína** es de:

$$z = 0,764071 - 0,00672946 * 180 - 0,0209341 * 2 + 0,014517 * 20 + 1,75986 * 1$$

$$z = 1,5611$$

$$P(Y = TIPO\_DE\_RCG = 1 / X) = \frac{1}{1 + e^{-1,5611}} = 0,82651113$$

Por lo tanto podemos decir que la probabilidad de que un paciente tenga un riesgo cardiovascular moderado-alto para este caso particular es de un 82,65 %, la cual es bastante alta para una persona con dichas características.

Como **segundo ejemplo ilustrativo** se pretende calcular la probabilidad de la misma persona pero ahora con una condición de artritis reumatoide **negativa**, es decir, que no tenga esta enfermedad:

$$z = 0,764071 - 0,00672946 * 180 - 0,0209341 * 2 + 0,014517 * 20 + 1,75986 * 0$$

$$z = -0,19876$$

$$P(Y = TIPO\_DE\_RCG = 1 / X) = \frac{1}{1 + e^{-(-0,19876)}} = 0,45047294$$

Ahora en este otro caso particular podemos decir que la probabilidad de que un paciente tenga un riesgo cardiovascular moderado-alto es de un 45,04 %, la cual es bastante alta para una persona con dichas características, pero mucho más baja que una persona que si tiene artritis reumatoide.

**Por lo tanto podemos afirmar que la condición de artritis de un paciente influye claramente en la probabilidad de que éste tenga en su riesgo cardiovascular global.**

Como **tercer ejemplo ilustrativo** se tiene la estimación de que un paciente con similares características de biomarcadores y además con una condición de artritis reumatoide **positiva** tenga un **riesgo cardiovascular global bajo**.

Fácilmente esta probabilidad se puede calcular aplicando el complemento de la probabilidad obtenida en el primer ejemplo ilustrativo, es decir:

$$1 - P(Y = TIPO\_DE\_RCG = 1 / X) = \frac{1}{1 + e^{-1,5611}} = 0,82651113$$

Por lo tanto,

$$P(Y = TIPO\_DE\_RCG = 0 / X) = \frac{e^{-z}}{1 + e^{-z}} = 1 - 0,82651113 = 0,1734$$

Acá podemos decir que la probabilidad de que un paciente tenga un riesgo cardiovascular bajo con las características antes señaladas es de un 17,4 %, lo cual es consecuente para una persona con dichas características,

Finalmente como **cuarto ejemplo ilustrativo** se pretende al igual que en el tercer perfil calcular la probabilidad de la misma persona pero ahora con una condición de artritis reumatoide **negativa**, es decir, que no tenga esta enfermedad:

Nuevamente esta probabilidad se puede calcular fácilmente aplicando el complemento de la probabilidad obtenida en el segundo ejemplo ilustrativo, es decir:

$$1 - P(Y = TIPO\_DE\_RCG = 1 / X) = \frac{1}{1 + e^{-(-0,19876)}} = 0,45047294 \quad , \text{ Así}$$

$$P(Y = TIPO\_DE\_RCG = 0 / X) = \frac{e^{-z}}{1 + e^{-z}} = 1 - 0,45047294 = 0,5495$$

En este otro caso particular podemos decir que la probabilidad de que un paciente tenga un riesgo cardiovascular bajo es de un 55,1 %, la cual es bastante alta para una persona con dichas características, inclusive más alta que una persona que si tiene artritis reumatoide.

**Aunque las predicciones se hayan podido hacer sin ningún problema, estas no tienen mucha validez como se comentaba previamente, ya que las variables independientes asociadas a este modelo de prueba no son estadísticamente significativas, es decir, no influyen en la variable respuesta, tal como se ve a continuación, por lo que predecir no tiene mucho sentido ya que puede apuntar a falsas interpretaciones.**

Tabla 43: Significación de variables en modelo logístico de prueba.

Likelihood Ratio Tests

Factor	Chi-Square	Df	P-Value
PCR	0,70145	1	0,4023
ROS	0,0375168	1	0,8464
TNFa	0,301768	1	0,5828
CONDICION AR	8,75851	1	0,0031

## 10.8.5 Aplicación de modelos de regresión logística.

Para realizar modelos de predicción “consistentes” a continuación se realizarán dos modelos logísticos generales, los cuales están completamente relacionados a los puntos anteriores de este trabajo, ya que el primero de estos modelos se realiza considerando todos los factores que tienen cierta dependencia con la variable Condición AR, y el segundo con los factores que tienen cierta dependencia pero con la variable Tipo de RCG.

### 10.8.5.1 Modelo de predicción para la condición de artritis reumatoide de los pacientes

El modelo final que se dejó en este punto es el que se muestra a continuación, el cual fue desarrollado por el método hacia atrás de Razón de Verosimilitud, que parte ingresando todas las variables en el modelo y luego va eliminando las que son menos significativas, hasta dejar el mejor modelo en el cual todas las variables sean estadísticamente significativas.

Principalmente este modelo fue elegido primero por tener un  $R^2$  aceptable para un modelo logístico y segundo porque todos sus parámetros estimados son estadísticamente significativos, lo que implica que las variables adjuntas a estos parámetros tienen una influencia significativa en la variable respuesta y por lo tanto las predicciones son aceptables.

Tabla 44: Modelo logístico para la condición AR de los pacientes

Estimated Regression Model (Maximum Likelihood)

Parameter	Estimate	Standard Error	Estimated Odds Ratio
CONSTANT	-2,14005	1,88194	
EDAD	0,0688735	0,0305852	1,0713
PCR	0,0937055	0,0552795	1,09824
OBESIDAD=0	-1,62862	0,567799	0,196199
HIPERTENSION=0	13,1467	137,581	512292,0
TRAT. ANTIH=0	-14,2572	137,582	6,42952E-7

Analysis of Deviance

Source	Deviance	Df	P-Value
Model	49,993	5	0,0000
Residual	80,4284	89	0,7304
Total (corr.)	130,421	94	

Percentage of deviance explained by model = 38,3319  
 Adjusted percentage = 29,131

Likelihood Ratio Tests

Factor	Chi-Square	Df	P-Value
EDAD	5,63537	1	0,0176
PCR	5,21308	1	0,0224
OBESIDAD	8,61108	1	0,0033
HIPERTENSION	8,34303	1	0,0039
TRAT. ANTIH	13,2864	1	0,0003

Consideremos ahora los datos relacionados con la **condición AR** (variable respuesta en este caso) de los pacientes e ilustremos con estos datos, la siguiente regresión logística que contiene variables predictoras ó independientes cualitativas y cuantitativas. Usaremos como predictoras **la obesidad central, la hipertensión y el tratamiento antihipertensivo** (todas con dos categorías), y como variables numéricas se tiene a la edad y la Proteína c Reactiva (PCR).

La obesidad central es una variable cualitativa por lo que para sus dos categorías se usan variables dummy, como se explica a continuación:

Tabla 45: Variables dummy para la obesidad central.

$X_1$ : Obesidad central	$X_1d_1$	$X_1d_2$
0 : NO	1	0
1: SI	0	0

$X_1d_1$ : Variable dummy 1, relacionada a la categoría de **NO** tener obesidad central

$X_1d_2$ : Variable dummy 2, relacionada a la categoría de **SI** tener obesidad central

Así la variable  $X_1$  ahora se puede reescribir de la siguiente forma:  $X_1 = X_1d_1 + X_1d_2$

Si consideramos ahora un modelo logístico para calcular la probabilidad de tener una condición AR POSITIVA, sólo considerando la obesidad central como variable independiente, sería en este caso,

$$P(Y = Condición\_AR = 1 / X_1) = \frac{1}{1 + e^{-z}},$$

$z = b_0 + b_1x_1d_1$ , para una persona que **NO** tiene obesidad central (se anula  $X_1d_2$ ).

$z = b_0$ , para una persona que **SI** tiene obesidad central (se anulan ambas categorías).

Para la hipertensión y el tratamiento antihipertensivo, las variables dummy asociadas son:

Tabla 46: Variables dummy para la hipertensión y el tratamiento antihipertensivo.

$X_2$ :Hipertensión	$X_2d_1$	$X_2d_2$	$X_3$ : Trat. Antihipertensivo	$X_3d_1$	$X_3d_2$
0 : NO	1	0	0 : NO	1	0
1: SI	0	0	1: SI	0	0

Donde:

$X_2d_1$ : Variable dummy 1, relacionada a la categoría de **NO** tener hipertensión.

$X_2d_2$ : Variable dummy 2, relacionada a la categoría de **SI** tener hipertensión.

$X_3d_1$ : Variable dummy 1, relacionada a la categoría de **NO** tener Trat. Antihipertensivo.

$X_3d_2$ : Variable dummy 2, relacionada a la categoría de **SI** tener Trat. Antihipertensivo.

$$X_2 = X_2d_1 + X_2d_2 \text{ y } X_3 = X_3d_1 + X_3d_2$$

Por lo tanto el modelo general para el cálculo de la probabilidad de tener una condición AR POSITIVA, incluyendo a estas tres variables categóricas, usando variables dummy y las otras 2 variables numéricas (edad y PCR), queda expresado de la siguiente forma:

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

$$z = b_0 + b_1x_1d_1 + b_2x_1d_2 + b_3x_2d_1 + b_4x_2d_2 + b_5x_3d_1 + b_6x_3d_2 + b_7x_4 + b_8x_5$$

Así ahora se puede ir combinando todos los posibles modelos logísticos para cada variable que presente o no cierta condición. En total son 8 modelos los que se pueden obtener, los cuales se muestran a continuación:

La fórmula para la probabilidad es la misma, solo varían los coeficientes  $b_i$  en cada modelo lineal ( $\mathbf{z}$ ) a utilizar.

- 1) Para pacientes SIN Obesidad central (=0), SIN Hipertensión (=0), SIN un Tratamiento Antihipertensivo (=0) y con cierta edad y PCR,

$$z = b_0 + b_1x_1d_1 + b_3x_2d_1 + b_5x_3d_1 + b_7x_4 + b_8x_5$$

$$Z = -2,14005 - 1,62862*OBESIDAD\ CENTRAL=0(NO) + 13,1467*HIPERTENSION=0(NO) - 14,2572*TRAT.\ ANTIH=0(NO) + 0,0688735*EDAD + 0,0937055*PCR$$

- 2) Para pacientes SIN Obesidad central (=0), SIN Hipertensión (=0), CON un Tratamiento Antihipertensivo (=1) y con cierta edad y PCR,

$$z = b_0 + b_1x_1d_1 + b_4x_2d_2 + b_5x_3d_1 + b_7x_4 + b_8x_5$$

$$Z = -16,3972 - 1,62862*OBESIDAD\ CENTRAL=0(NO) + 13,1467*HIPERTENSION=0(NO) + 14,2572*TRAT.ANTIH=1(SI) + 0,0688735*EDAD + 0,0937055*PCR +$$

- 3) Para pacientes SIN Obesidad central (=0), CON Hipertensión (=1), SIN un Tratamiento Antihipertensivo (=0) y con cierta edad y PCR,

$$z = b_0 + b_2x_1d_2 + b_3x_2d_1 + b_5x_3d_1 + b_7x_4 + b_8x_5$$

$$Z = 11,0066 - 1,62862*OBESIDAD\ CENTRAL=0(NO) - 13,1467*HIPERTENSION=1(SI) - 14,2572*TRAT.\ ANTIH=0(NO) + 0,0937055*PCR + 0,0688735*EDAD$$

- 4) Para pacientes CON Obesidad central (=1), SIN Hipertensión (=0), SIN un Tratamiento Antihipertensivo (=0) y con cierta edad y PCR,

$$z = b_0 + b_1x_1d_1 + b_3x_2d_1 + b_5x_3d_1 + b_7x_4 + b_8x_5$$

$$Z = -3,76867+1,62862*OBESIDAD\ CENTRAL=1(SI) +13,1467*HIPERTENSION=0(NO) -14,2572*TRAT.ANTIH=0(NO) + 0,0937055*PCR + 0,0688735*EDAD$$

- 5) Para pacientes CON Obesidad central (=1), SIN Hipertensión (=0), CON un Tratamiento Antihipertensivo (=1) y con cierta edad y PCR,

$$z = b_0 + b_2x_1d_2 + b_3x_2d_1 + b_6x_3d_2 + b_7x_4 + b_8x_5$$

$$Z = -18,0259 + 1,62862*OBESIDAD\ CENTRAL=1(SI) +13,1467*HIPERTENSION=0(NO) + 14,2572*TRAT.ANTIH=1(SI) + 0,0937055*PCR + 0,0688735*EDAD$$

- 6) Para pacientes CON Obesidad central (=1), CON Hipertensión (=1), SIN un Tratamiento Antihipertensivo (=0) y con cierta edad y PCR,

$$z = b_0 + b_2x_1d_2 + b_4x_2d_2 + b_5x_3d_1 + b_7x_4 + b_8x_5$$

$$Z = 9,37798 + 1,62862*OBESIDAD\ CENTRAL=1(SI) -13,1467*HIPERTENSION=1(SI) - 14,2572*TRAT.ANTIH=0(NO) + 0,0937055*PCR + 0,0688735*EDAD$$

- 7) Para pacientes SIN Obesidad central (=0), CON Hipertensión (=1), CON un Tratamiento Antihipertensivo (=1) y con cierta edad y PCR,

$$z = b_0 + b_1x_1d_1 + b_4x_2d_2 + b_6x_3d_2 + b_7x_4 + b_8x_5$$

$$Z = -3,25059 - 1,62862*OBESIDAD\ CENTRAL=0(NO) -13,1467*HIPERTENSION=1(SI) + 14,2572*TRAT.ANTIH=1(SI) + 0,0937055*PCR + 0,0688735*EDAD$$

8) Para pacientes CON Obesidad central (=1), CON Hipertensión (=1), CON un Tratamiento Antihipertensivo (=1) y con cierta edad y PCR,

$$z = b_0 + b_2x_1d_2 + b_4x_2d_2 + b_6x_3d_2 + b_7x_4 + b_8x$$

$$Z = -4,87921 + 1,62862*OBESIDAD\ CENTRAL=1(SI) -13,1467*HIPERTENSION=1(SI) + 14,2572*TRAT.ANTIH=1(SI) + 0,0937055*PCR + 0,0688735*EDAD$$

Ahora sólo basta calcular las probabilidades para cada caso, lo cual se hará a continuación para un caso particular y además más adelante a través de la herramienta interactiva desarrollada adjunta a este informe de título.

**Por ejemplo para el primer modelo posible en este caso, se tiene que:**

En la codificación dummy la OBESIDAD CENTRAL = 0 (: NO) toma el valor 1 en  $X_1d_1$ , la Hipertensión = 0 (: NO), también toma el valor 1 en este caso y el trat. Antihipertensivo ídem, por lo tanto al reemplazar estos valores “1” en la ecuación z que se obtuvo para este caso, queda lo siguiente:

$$z = -2,14005 - 1,62862*(1) + 13,1467*(1) - 14,2572*(1) + 0,0937055*PCR + 0,0688735*EDAD$$

Finalmente dadas las condiciones antes señaladas en este caso particular para cierto paciente y suponiendo que tiene una PCR de 2mg/L y una edad de 35, la probabilidad de que tenga una condición de artritis POSITIVA es:

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

$$z = -2,14005 - 1,62862*(1) + 13,1467*(1) - 14,2572*(1) + 0,0937055*(2) + 0,0688735*(35)$$

$$z = -2.2811865$$

$$P(Y = Condición\_AR = 1 / X_i) = \frac{1}{1 + e^{-(-2.2811865)}} = 0.0926 \text{ ó } 9\%.$$

Así también se puede calcular la probabilidad de que esa misma persona tenga una condición de artritis NEGATIVA, sólo calculando el complemento de esta última probabilidad obtenida  $(1 - P(Y = Condición\_AR = 1 / X_i)) = 0.9074$  ó 90.7%

En general se puede calcular cualquier probabilidad que se desee estimar para este modelo de regresión logística, más prácticamente a través de la herramienta Excel.

### 10.8.5.2 Modelo de predicción para el tipo de Riesgo Cardiovascular Global de los pacientes

El modelo final que se dejó en este punto es el que se muestra a continuación, el cual fue elegido y desarrollado con los mismos criterios usados en el punto anterior.

Tabla 47: Modelo logístico para el tipo de RCG de los pacientes.

Estimated Regression Model (Maximum Likelihood)

Parameter	Estimate	Standard Error	Estimated Odds Ratio
CONSTANT	14,5678	20,9214	
HIPERGLICEMIA	0,151891	0,0705788	1,16403
HIPERTRIGLICERIDE	-8,82327	13,762	0,000147267
HIPERTENSION=0	-9,51276	11,3498	0,0000739026
OBESIDAD=0	-10,9906	10,1382	0,0000168588

Analysis of Deviance

Source	Deviance	Df	P-Value
Model	69,4431	4	0,0000
Residual	22,3715	89	1,0000
Total (corr.)	91,8146	93	

Percentage of deviance explained by model = 75,6341  
 Adjusted percentage = 64,7426

Likelihood Ratio Tests

Factor	Chi-Square	Df	P-Value
HIPERGLICEMIA	37,7897	1	0,0000
HIPERTRIGLICERIDEMIA.	7,60711	1	0,0058
HIPERTENSION	39,6511	1	0,0000
OBESIDAD	39,494	1	0,0000

Consideremos ahora los datos relacionados con **el tipo de RCG** (variable respuesta en este caso) de los pacientes e ilustremos con estos datos, la siguiente regresión logística que contiene variables predictoras ó independientes cualitativas y cuantitativas. Usaremos como predictoras **la obesidad central, la hipertensión y la hipertrigliceridemia** (todas con dos categorías), y como variable numérica se tiene a la **Hiperglicemia**.

La obesidad central y la hipertensión son variables cualitativas por lo que para sus dos categorías se usaron variables dummy anteriormente, mientras que para el caso de la hipertrigliceridemia, se tiene que las variables dummy asociadas son las siguientes:

Tabla 48: Variables dummy para la hipertrigliceridemia.

X <sub>3</sub> :Hipertrigliceridemia	X <sub>3</sub> d <sub>1</sub>	X <sub>3</sub> d <sub>2</sub>
0 : NO	1	0
1 : SI	0	0

Por lo tanto el modelo general para el cálculo de la probabilidad de tener un tipo de riesgo cardiovascular global ALTO, incluyendo a estas tres variables categóricas, usando variables dummy y la otra variable numérica (Hiperglicemia), queda expresado de la siguiente forma:

$$P(Y = Tipo\_de\_RCG\_ALTO = 1 / X_i) = \frac{1}{1 + e^{-z}}$$

$$z = b_0 + b_1x_1d_1 + b_2x_1d_2 + b_3x_2d_1 + b_4x_2d_2 + b_5x_3d_1 + b_6x_3d_2 + b_7x_4$$

Así ahora se puede ir combinando todos los posibles modelos logísticos para cada variable que presente o no cierta condición. En total nuevamente son 8 modelos los que se pueden obtener, los cuales se muestran a continuación:

La fórmula para la probabilidad es la misma, solo varían los coeficientes  $b_i$  en cada modelo lineal (**z**) a utilizar.

1) Para pacientes SIN Obesidad central (=0), SIN Hipertensión (=0), SIN Hipertrigliceridemia (=0) y con cierto valor de Hiperglicemia.

$$Z = 14,5678 - 10,9906*OBESIDAD\ CENTRAL=0(NO) - 9,51276*HIPERTENSION=0(NO) - 8,82327*HIPERTRIGLICERIDEMIA=0(NO) + 0,151891*HIPERGLICEMIA$$

2) Para pacientes SIN Obesidad central (=0), SIN Hipertensión (=0), CON Hipertrigliceridemia (=1) y con cierto valor de Hiperglicemia.

$$Z = 5,74449 - 10,9906*OBESIDAD\ CENTRAL=0(NO) - 9,51276*HIPERTENSION=0(NO) + 8,82327*HIPERTRIGLICERIDEMIA=1(SI) + 0,151891*HIPERGLICEMIA$$

3) Para pacientes SIN Obesidad central (=0), CON Hipertensión (=1), SIN Hipertrigliceridemia (=0) y con cierto valor de Hiperglicemia.

$$Z = 5,05499 - 10,9906*OBESIDAD\ CENTRAL=0(NO) + 9,51276*HIPERTENSION=1(SI) - 8,82327*HIPERTRIGLICERIDEMIA=0(NO) + 0,151891*HIPERGLICEMIA$$

4) Para pacientes CON Obesidad central (=1), SIN Hipertensión (=0), SIN Hipertrigliceridemia (=0) y con cierto valor de Hiperglicemia.

$$Z = 3,57711 + 10,9906*OBESIDAD\ CENTRAL=1(SI) - 9,51276*HIPERTENSION=0(NO) - 8,82327*HIPERTRIGLICERIDEMIA=0(NO) + 0,151891*HIPERGLICEMIA$$

5) Para pacientes CON Obesidad central (=1), SIN Hipertensión (=0), CON Hipertrigliceridemia (=1) y con cierto valor de Hiperglicemia.

$$Z = -5,24615 + 10,9906*OBESIDAD\ CENTRAL=1(SI) - 9,51276*HIPERTENSION=0(NO) + 8,82327*HIPERTRIGLICERIDEMIA=1(SI) + 0,151891*HIPERGLICEMIA$$

6) Para pacientes CON Obesidad central (=1), CON Hipertensión (=1), SIN Hipertrigliceridemia (=0) y con cierto valor de Hiperglicemia.

$$Z = -5,7 + 10,9906*OBESIDAD\ CENTRAL=1(SI) + 9,51276*HIPERTENSION=1(SI) - 8,82327*HIPERTRIGLICERIDEMIA=0(NO) + 0,151891*HIPERGLICEMIA$$

7) Para pacientes SIN Obesidad central (=0), CON Hipertensión (=1), CON Hipertrigliceridemia (=1) y con cierto valor de Hiperglicemia.

$$Z = -3,5 - 10,9906*OBESIDAD\ CENTRAL=0(NO) + 9,51276*HIPERTENSION=1(SI) + 8,82327*HIPERTRIGLICERIDEMIA=1(SI) + 0,151891*HIPERGLICEMIA$$

8) Para pacientes CON Obesidad central (=1), CON Hipertensión (=1), CON Hipertrigliceridemia (=1) y con cierto valor de Hiperglicemia.

$$Z = -14,7589 + 10,9906*OBESIDAD\ CENTRAL=1(SI) + 9,51276*HIPERTENSION=1(SI) + 8,82327*HIPERTRIGLICERIDEMIA=1(SI) + 0,151891*HIPERGLICEMIA$$

Ahora sólo basta calcular las probabilidades para cada caso al igual que en el punto anterior, lo cual se hará a continuación para un caso particular y luego más adelante a través de la herramienta interactiva desarrollada adjunta a este informe de título.

**Por ejemplo para el primer modelo posible en este caso, se tiene que:**

En la codificación dummy la OBESIDAD CENTRAL = 0 (: NO) toma el valor 1 en  $X_1d_1$ , la Hipertensión = 0 (: NO), también toma el valor 1 en este caso y la hipertrigliceridemia ídem, por lo tanto al reemplazar estos valores "1" en la ecuación z que se obtuvo para este caso, queda lo siguiente:

$$Z = 14,5678 - 10,9906*(1) - 9,51276*(1) - 8,82327*(1) + 0,151891*HIPERGLICEMIA$$

Finalmente dadas las condiciones antes señaladas en este caso particular para cierto paciente y suponiendo que tiene una Hiperglicemia de 80, la probabilidad de que tenga un riesgo cardiovascular global ALTO es:

$$P(Y = \text{Tipo\_de\_RCG\_ALTO} = 1 / X_i) = \frac{1}{1 + e^{-z}}$$

$$Z = 14,5678 - 10,9906*(1) - 9,51276*(1) - 8,82327*(1) + 0,151891*(80)$$

$$Z = -2.60755$$

$$P(Y = \text{TIPO\_DE\_RCG\_ALTO} = 1 / X_i) = \frac{1}{1 + e^{-(-2.60755)}} = 0.0686 \text{ ó } 6.86\%$$

También se puede calcular la probabilidad de que esa misma persona tenga un Tipo de RCG BAJO, sólo calculando el complemento de esta última probabilidad obtenida ( $1 - P(Y = \text{TIPO\_DE\_RCG\_ALTO} = 1 / X_i) = 0.931345$  ó 93.13%).

En general también en este modelo se puede calcular cualquier probabilidad que se desee estimar para este modelo de regresión logística, más prácticamente a través de la herramienta Excel.

## 10.8.6 Herramienta Interactiva

Ahora para hacer más interesante este informe de Título II, se ha desarrollado una herramienta interactiva, a través de Microsoft Excel 2007, con el fin de predecir todas las combinaciones de modelos de regresión logística posibles para un determinado paciente, tanto para la condición de artritis como para la condición de riesgo cardiovascular global que presenten los pacientes.

Estas aplicaciones se encuentran en el CD adjunto a este informe de título II y tiene como nombre “Modelos de predicción logísticos.xlsx”.

### 10.8.6.1 Portada de la Herramienta



Figura 15: Portada de herramienta interactiva para un modelos de predicción logística.

La portada de esta herramienta interactiva muestra el título y el índice que contiene a los dos modelos antes señalados, además de un modelo de regresión múltiple extra para calcular el valor real del Riesgo cardiovascular Global, el cual corrobora la fuerte influencia de ciertas variables en ambos modelos.

## 10.8.6.2 Herramienta interactiva de predicción para el ejemplo de la condición de artritis de los pacientes

**MODELO DE PREDICCIÓN PARA LA CONDICIÓN DE ARTRITIS REUMATOIDE**

**DATOS DEL MODELO**

N = 97	VARIABLE DEPENDIENTE	VARIABLES INDEPENDIENTES				
PACIENTE	ARTRITIS REUMATOIDE	VARIABLES			BIOMARCADORES	
Nº	Y=CONDICIÓN DE AR	X1=OBESIDAD CENTRAL	X2= HIPERTENSIÓN	X3=TRATAMIENTO ANTIHIPERTENSIVO	X4=PCR	X5=ROS
1	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
2	1:PACIENTE CON AR	0:NO	1:SI	1:SI	0: < 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
3	1:PACIENTE CON AR	1:SI	0:NO	0:NO	1: > ó = 3,19 mg/L	0: < 2,12 IF/ug proteína
4	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	0: < 2,12 IF/ug proteína
5	1:PACIENTE CON AR	0:NO	1:SI	1:SI	0: < 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
6	1:PACIENTE CON AR	0:NO	0:NO	0:NO	0: < 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
7	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	0: < 2,12 IF/ug proteína
8	1:PACIENTE CON AR	1:SI	0:NO	0:NO	0: < 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
9	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
10	1:PACIENTE CON AR	0:NO	0:NO	0:NO	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
11	1:PACIENTE CON AR	1:SI	1:SI	1:SI	0: < 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
12	1:PACIENTE CON AR	1:SI	0:NO	0:NO	0: < 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
13	1:PACIENTE CON AR	1:SI	0:NO	0:NO	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
14	1:PACIENTE CON AR	1:SI	0:NO	0:NO	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
15	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
16	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
17	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
18	1:PACIENTE CON AR	0:NO	0:NO	0:NO	1: > ó = 3,19 mg/L	1: > ó = 2,12 IF/ug proteína
19	1:PACIENTE CON AR	1:SI	1:SI	1:SI	1: > ó = 3,19 mg/L	0: < 2,12 IF/ug proteína

CALCULAR PREDICCIÓN

VOLVER

Figura 16: Imagen 1 de herramienta interactiva para un modelo de predicción logístico para la condición AR de los pacientes.

**MODELO DE PREDICCIÓN PARA LA CONDICIÓN DE ARTRITIS REUMATOIDE**

**COEFICIENTES ESTIMADOS**  
(FACTORES FIJOS)

B1	B2	B3	B4	B5
1,629	13,1467	14,2572	0,094	0,069

**ESTADÍSTICOS DESCRIPTIVOS**  
(VALORES POSIBLES PARA PREDECIR)

OBESIDAD CENTRAL	HIPERTENSIÓN	TRAT. ANTIHIPERTENSIVO
0	0	0
1	1	1

(INTERVALOS POSIBLES PARA PREDECIR)

	PCR	EDAD
MÍNIMO VALOR	0,19	30
MÁXIMO VALOR	87,2	76

**MODELOS LOGÍSTICOS ASOCIADOS**

Y = 0 =  $\frac{\text{EXP}(-Z)}{1 + \text{EXP}(-Z)}$ , PARA PREDECIR UNA CONDICIÓN DE ARTRITIS REUMATOIDE **NEGATIVA** DE UN PACIENTE

, Z = (+ ó -) B0 (+ ó -) 1,629\*OBESIDAD CENTRAL (+ ó -) 13,1467\*HIPERTENSION (+ ó -) 14,2572\*TRATAMIENTO ANTIHIPERTENSIVO (+ ó -) 0,094\*PCR (+ ó -) 0,069\*ROS = -2,281

Y = 1 =  $\frac{1}{1 + \text{EXP}(-Z)}$ , PARA PREDECIR UNA CONDICIÓN DE ARTRITIS REUMATOIDE **POSITIVA** DE UN PACIENTE

**ESTIMACIÓN DE PROBABILIDADES**

1\*) INGRESAR TIPO DE CONDICIÓN DE ARTRITIS A PREDECIR: **POSITIVA**

2\*) INGRESAR LOS VALORES DE:

X1= OBESIDAD CENTRAL	X2= HIPERTENSIÓN	X3=TRATAMIENTO ANTIHIPERTENSIVO	X4=PCR	X5=EDAD
0	0	0	2	35

3\*) P(Y=CONDICIÓN AR=POSITIVA / Xi) = **0,092695** ó **9,269 %**

**CONCLUSIÓN:** LA PROBABILIDAD DE QUE UN PACIENTE TENGA UNA CONDICIÓN DE ARTRITIS REUMATOIDE **POSITIVA** CONSIDERANDO QUE **NO** TIENE OBESIDAD CENTRAL, QUE **NO** TIENE HIPERTENSIÓN, QUE **NO** TIENE UN TRATAMIENTO ANTIHIPERTENSIVO, CON UNA PCR DE **2** Y UNA EDAD DE **35** ES DE **9,269**

Figura 17: Imagen 2 de herramienta interactiva para un modelo de predicción logístico para la condición AR de los pacientes.

### 10.8.6.4 Herramienta interactiva de predicción para el ejemplo del Tipo de riesgo cardiovascular Global de los pacientes

**MODELO DE PREDICCIÓN PARA EL TIPO DE RIESGO CARDIOVASCULAR GLOBAL**

**DATOS DEL MODELO**

N = 101	VARIABLE DEPENDIENTE	VARIABLES INDEPENDIENTES				
PACIENTE	RIESGO GLOBAL	VARIABLES				
Nº	Y=TIPO DE RCG	X1:OBESIDAD CENTRAL	X2 = HIPERTENSIÓN	X3=ANTECEDENTES FAMILIARES	X4=HIPERTRIGLICERIDEMIA	X5=HIPERGLICEMIA
1	1:PACIENTE CON AR	1: SI	1: SI	1: SI	0: NO	98
2	1:PACIENTE CON AR	0: NO	1: SI	1: SI	1: SI	78
3	1:PACIENTE CON AR	1: SI	0: NO	1: SI	1: SI	215
4	1:PACIENTE CON AR	1: SI	1: SI	0: NO	1: SI	95
5	1:PACIENTE CON AR	0: NO	1: SI	0: NO	0: NO	82
6	1:PACIENTE CON AR	0: NO	0: NO	0: NO	0: NO	98
7	1:PACIENTE CON AR	1: SI	1: SI	1: SI	1: SI	92
8	1:PACIENTE CON AR	1: SI	0: NO	0: NO	0: NO	91
9	1:PACIENTE CON AR	1: SI	1: SI	0: NO	0: NO	158
10	1:PACIENTE CON AR	0: NO	0: NO	1: SI	0: NO	87
11	1:PACIENTE CON AR	0: NO	0: NO	0: NO	0: NO	89
12	1:PACIENTE CON AR	1: SI	1: SI	1: SI	1: SI	85
13	1:PACIENTE CON AR	1: SI	0: NO	1: SI	0: NO	89
14	1:PACIENTE CON AR	1: SI	0: NO	0: NO	0: NO	73
15	1:PACIENTE CON AR	1: SI	0: NO	0: NO	0: NO	84
16	1:PACIENTE CON AR	1: SI	1: SI	0: NO	0: NO	111
17	1:PACIENTE CON AR	1: SI	1: SI	0: NO	0: NO	185
18	1:PACIENTE CON AR	1: SI	1: SI	1: SI	0: NO	214
19	1:PACIENTE CON AR	0: NO	0: NO	0: NO	0: NO	95

Figura 18: Imagen 1 de herramienta interactiva para un modelo de predicción logístico para el tipo de RCG de los pacientes.

**MODELO DE PREDICCIÓN PARA LA CONDICIÓN DE ARTRITIS REUMATOIDE**

**COEFICIENTES ESTIMADOS**  
(FACTORES FIJOS)

B1	B2	B3	B4
10,9906	9,5127	8,82327	0,1519

**ESTADÍSTICOS DESCRIPTIVOS**  
(VALORES POSIBLES PARA PREDECIR)

OBESIDAD CENTRAL	HIPERTENSIÓN	HIPERTRIGLICERIDEMIA
0	0	0
1	1	1

(INTERVALOS POSIBLES PARA PREDECIR)

HIPERGLICEMIA	
MÍNIMO VALOR	71
MÁXIMO VALOR	215

**MODELOS LOGÍSTICOS ASOCIADOS**

Y = 0 =  $\frac{EXP(-Z)}{1 + EXP(-Z)}$ , PARA PREDECIR UNA CONDICIÓN DE ARTRITIS REUMATOIDE **NEGATIVA** DE UN PACIENTE

Y = 1 =  $\frac{1}{1 + EXP(-Z)}$ , PARA PREDECIR UNA CONDICIÓN DE ARTRITIS REUMATOIDE **POSITIVA** DE UN PACIENTE

Z = (+ ó -) B0 (+ ó -) 10,9906\*OBESIDAD CENTRAL (+ ó -) 9,5127\*HIPERTENSION (+ ó -) 8,82327\*HIPERTRIGLICERIDEMIA (+ ó -) 0,1519\*HIPERGLICEMIA = -2,607

**ESTIMACIÓN DE PROBABILIDADES**

1\*) INGRESAR TIPO DE RIESGO CARDIOVASCULAR GLOBAL A PREDECIR: **ALTO**

2\*) INGRESAR LOS VALORES DE:

X1 = OBESIDAD CENTRAL	X2 = HIPERTENSIÓN	X3 = HIPERTRIGLICERIDEMIA	X4 = HIPERGLICEMIA
0	0	0	80

3\*) P(Y=CONDICIÓN AR= **ALTO** / X1) = 0,068658      6      6,866 %

**CONCLUSIÓN:** LA PROBABILIDAD DE QUE UN PACIENTE TENGA UN TIPO DE RIESGO CARDIOVASCULAR GLOBAL **ALTO** CONSIDERANDO QUE **NO** TIENE OBESIDAD CENTRAL, QUE **NO** TIENE HIPERTENSIÓN, **NO** TIENE ANTECEDENTES FAMILIARES Y CON UNA PCR DE **80** ES DE **6,866** ES DE **6,866** %

Figura 19: Imagen 2 de herramienta interactiva para un modelo de predicción logístico para el tipo de RCG de los pacientes.

En general en cada una de las imágenes se puede ver que las probabilidades generadas por la herramienta coinciden con las calculadas anteriormente en los ejemplos señalados.

Cada una de estas imágenes además muestra:

- Los datos que se utilizaron en la construcción del modelo: *Con ellos se ha calculado tanto los coeficientes estimados como las estadísticas descriptivas.*
- Los coeficientes estimados (*B*) del modelo: *Estos factores son fijos en esta herramienta y son necesarios para el cálculo de  $z$ , que es parte de los modelos logísticos asociados.*
- Los estadísticos descriptivos de cada variable: *Son útiles como intervalos para introducir el valor de cada biomarcador en la predicción final, por lo que se recomienda hacer la predicción dentro de los límites de cada variable.*
- Los modelos logísticos asociados a los datos: Cada modelo logístico tiene dos posibles modelos a utilizar, uno que calculará la probabilidad de una categoría de la variable respuesta, en este caso, un tipo de RCG BAJO (codificada como 0) y otro que calculará la probabilidad de la otra categoría de la variable respuesta, o sea, un tipo de RCG ALTO (codificada como 1).
- También se han incorporado mensajes de error para el caso en que los valores a ingresar por algún usuario estén mal realizados, en la sección “*estimación de probabilidades*”.

Es claro que cada uno de ellos es indispensable para la aplicación de la estimación de probabilidad del modelo logístico final que se construye.

Finalmente, se muestra la conclusión de cada predicción realizada.

## 10.8.6.7 Desarrollo de herramienta interactiva

1°) Para el desarrollo de la herramienta interactiva en la hoja 2, se parte por ingresar la probabilidad de **Artritis Reumatoide** que se desea predecir para cada caso en la celda **X18**, la que puede ser, una predicción de CONDICIÓN AR POSITIVA ó NEGATIVA.

Para el desarrollo de la herramienta interactiva en la hoja 3, se parte por ingresar la probabilidad de Tipo de RCG que se desea predecir para cada caso en la celda **X18**, la que puede ser, una predicción de Riesgo Cardiovascular Global ALTO ó BAJO.

2°) Luego sólo basta ingresar los valores numéricos especificados para cada variable en las celdas **AC20 y AD20** para la PCR y la EDAD, respectivamente, además de la condición de obesidad central, hipertensión y tratamiento antihipertensivo en las celdas **V20, X20, AA20, para el caso del modelo de predicción asociado a la condición AR de los pacientes** y la condición de obesidad central, hipertensión, antecedentes familiares y Hiperglicemia en las celdas **V20, X20, AA20, AC20, respectivamente para el caso del modelo de predicción asociado al Tipo de RCG de los pacientes.**

3°) Por último en la celda **V22** se obtiene el valor de la probabilidad calculada para los modelos de predicción y en **Y22** el porcentaje de dichas probabilidades.

Más aún en el **recuadro final** se muestra la interpretación completa de las predicciones realizadas.

# 11. Conclusiones

**OBJETIVO B1):** Analizar la relación de parámetros antropométricos, clínicos, bioquímicos e inmunológicos y biomarcadores de disfunción endotelial e inflamación celular de acuerdo a la condición AR de los pacientes y Tipo de Riesgo cardiovascular Global.

**OBJETIVO B2):** Determinar y contrastar perfiles para cada una de las relaciones anteriores.

**OBJETIVO B3):** Identificar variables y biomarcadores modificables para disminuir un evento cardiovascular a futuro en los pacientes.

## **CONCLUSIÓN PARA OBJETIVO B1, B2 Y B3:**

En el caso de las variables que se pueden modificar para mejorar la condición de riesgo de los pacientes, son la hipertensión, un tratamiento antihipertensivo, que está totalmente relacionado a la variable anterior, la obesidad central y la Proteína C reactiva, que son las variables que más determinan una condición de artritis de otra y también una condición de tipo de riesgo cardiovascular global de las otras.

En el caso de los biomarcadores que más determinan una condición de artritis de otra y también una condición de tipo de riesgo cardiovascular global de las otras, son las especies reactivas de oxígeno y la Proteína C reactiva, que en este caso está considerada como biomarcador y corrobora que anteriormente como variable también es determinante en las condiciones para este grupo de pacientes.

También se puede ver de los perfiles que la edad es una variable determinante en que una persona tenga o no la condición de artritis reumatoide.

Por otra parte los antecedentes familiares, la diabetes, el colesterol total/HDL, la hiperglicemia, la hipertrigliceridemia, la hiperinsulinemia, el HOMA-IR y el IMC son variables determinantes en que una persona tenga un riesgo bajo, moderado o alto; y no así es el resto de variables que no son parte de estos perfiles.

**OBJETIVO B4) Comparar el riesgo cardiovascular global entre una población de pacientes con artritis reumatoide y sujetos controles, considerando los parámetros o variables en estudio.**

Podemos afirmar que efectivamente hay diferencias estadísticamente significativas de riesgo cardiovascular global entre pacientes con artritis reumatoide y sujetos controles, pero además se debe considerar que hay otras tres variables ( bloques) que influyen en la comparación de estos dos grupos de pacientes, por lo que se hace muy recomendable separar previamente para efecto de estas comparaciones entre las categorías de la variable hipertensión (si/no), obesidad central (si/no) y PCRus (si/no).

**OBJETIVO B5) Comparar los biomarcadores de inflamación celular (TNF- $\alpha$ , y proteína C reactiva ultrasensible (PCRus)) y disfunción endotelial (especies reactivas de oxígeno (ROS)) entre una población de pacientes con artritis reumatoide y sujetos controles, considerando los parámetros o variables en estudio.**

Podemos afirmar que efectivamente hay diferencias estadísticamente significativas de TNFa entre pacientes con artritis reumatoide y sujetos controles.

Podemos afirmar que efectivamente hay diferencias estadísticamente significativas de PCR entre pacientes con artritis reumatoide y sujetos controles con un 95% de confianza, pero además se debe considerar que hay otra variable (bloque) que influye en la comparación de estos dos grupos de pacientes, por lo que se hace muy recomendable separar previamente para efecto de estas comparaciones entre las categorías de la variable colesterol HDL bajo (SI/NO).

Podemos afirmar que efectivamente hay diferencias estadísticamente significativas de ROS entre pacientes con artritis reumatoide y sujetos controles.

**OBJETIVO B6) Crear un modelo de predicción de regresión logística interactivo, a través de Microsoft Excel, para el Riesgo Cardiovascular Global de pacientes con artritis reumatoide y sujetos controles, considerando biomarcadores.**

Con los modelo logísticos de predicción podremos afirmar si es que la condición y las especificaciones de valores en las distintas variables de un paciente influyen o no en la probabilidad de que éste tenga primero una condición o no de artritis reumatoide positiva y luego un riesgo cardiovascular global que sea preocupante para considerar a futuro.

Profundizando más aun el tema, específicamente, en la herramienta interactiva se puede apreciar la utilidad que tiene realizar un modelo de predicción logístico ya que se puede determinar claramente la resolución de diferentes ejemplos ilustrativos tal como los que se propusieron.

En general se puede apreciar que los modelos finales que se dejaron poseen variables muy significativas ya que al variar en sus valores hacen que las probabilidades cambien significativamente.

Para el caso del primer modelo de predicción relacionado a la condición de artritis de los pacientes, tener o no hipertensión y seguir o no un tratamiento antihipertensivo condicionan demasiado a si un paciente tiene o no una alta probabilidad de tener artritis reumatoide. En cuanto a la obesidad central también es determinante en que cierto paciente tengo una alta probabilidad o no de tener artritis pero en menor escala a las anteriores y por último las variables numéricas incluidas en este modelo si bien son significativas igualmente, un cambio en ellas no afecta tanto como las otras en que una persona tenga una alta o baja probabilidad de tener una condición de artritis reumatoide a futuro.

Para el caso del segundo modelo de predicción relacionado al Tipo de Riesgo Cardiovascular Global de los pacientes tanto la condición de tener o no obesidad central, hipertensión o hipertrigliceridemia tienen igualmente una fuerte influencia en la probabilidad de tener un riesgo global ALTO, ya que sólo basta que una de ellas este presente para que la probabilidad de tener este riesgo sea grande, esto es porque existe una gran correlación entre las variables predictoras y dependiente.

En cuanto a la variable hiperglicemia también es significativa pero no tanto como el resto de las variables, por lo que produce menos cambios en las probabilidades calculadas.

## **OBJETIVO GENERAL**

**Analizar el riesgo cardiovascular global y proponer un modelo de predicción en pacientes con artritis reumatoide y sujetos controles considerando biomarcadores de inflamación celular y disfunción endotelial.**

Este estudio servirá para caracterizar a una muestra representativa de la población chilena con AR, y generará información de utilidad para mejorar los protocolos establecidos en la Guía Clínica de Artritis Reumatoide del Minsal. Además, la mejor estimación del Riesgo Cardiovascular Global permitirá a los clínicos abordar al paciente AR con intervenciones más agresivas y seguimientos de los factores de riesgo que permitan disminuir las comorbilidades y la principal causa de muerte en estos pacientes.

Con los resultados de este proyecto, principalmente en base a la validación estadística de los biomarcadores, permitirá proponer una nueva herramienta diagnóstica que se podría incorporar al Laboratorio como prestación de servicios. Es factible además, generar un software para incorporar todos los datos de los pacientes, incluidos los biomarcadores y con este nuevo estimador de riesgo se podrá facilitar la interpretación combinada de muchos exámenes clínicos y factores de riesgos de los individuos que aún no desarrollan patologías (salud preventiva) y de aquellos que ya la presentan (intervención terapéutica).

**Así se puede concluir finalmente que las variables Hipertensión y Obesidad Central son los factores de RCG con mayor preocupación entre todas las variables para las personas que formaron parte de este estudio. Se recomienda tomar medidas más agresivas ante estos factores ya que condicionan demasiado al grupo de pacientes. Además se debe considerar que los biomarcadores que se estudiaron no se deben dejar de lado por ningún motivo, pues estos si determinan una condición de riesgo preocupante de otra, particularmente el ROS y PCR, ya que se comprobó estadísticamente entre los dos grupos de pacientes que estos biomarcadores determinan que un paciente esté o no con un riesgo cardiovascular global elevado y a la vez preocupante, es decir, si presentan diferencias significativas de RCG.**

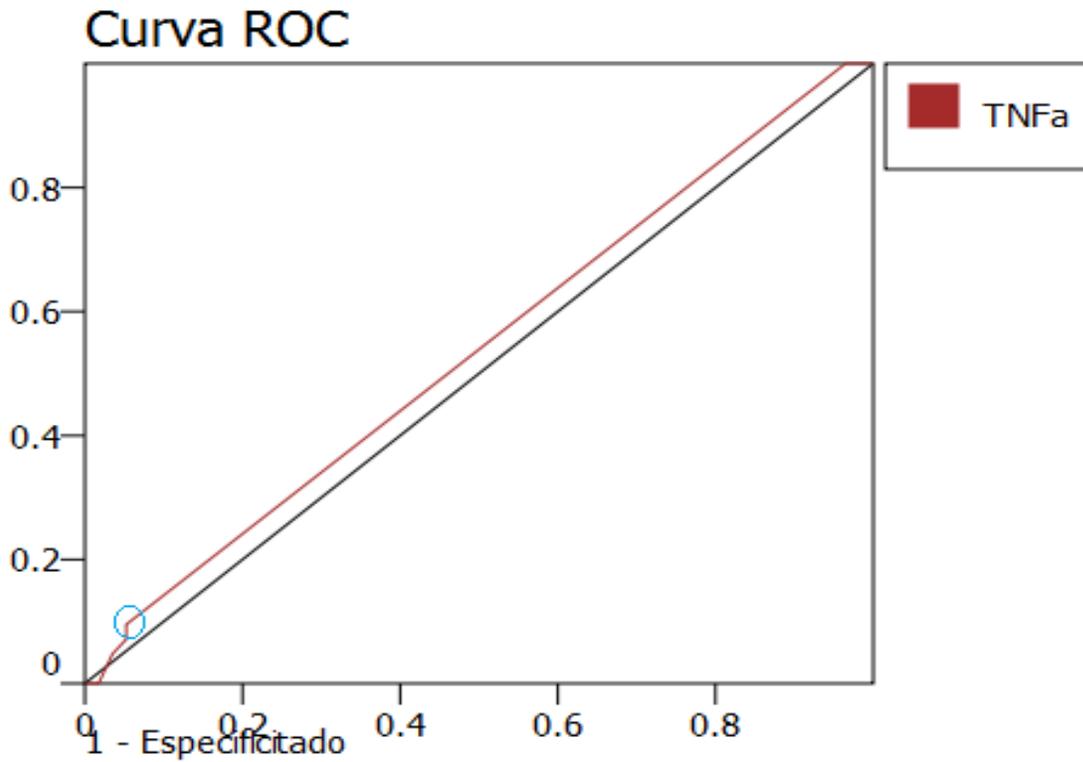
## 12. Bibliografía

1. Ministerio de Salud. Guía Clínica Artritis Reumatoide. Serie Guías Clínicas MINSAL N° 52, 2007.
2. Ministerio de Salud. Implementación del enfoque de riesgo en el Programa de Salud Cardiovascular. Serie Guías Clínicas MINSAL, 2009.
3. Proyecto de investigación, Formulario DIUC, “NUEVOS BIOMARCADORES DE DISFUNCION ENDOTELIAL PARA ESTRATIFICAR EL RIESGO CARDIOVASCULAR GLOBAL EN PACIENTES CON ARTRITIS REUMATOIDE”.
4. Rolph Anderson, Joseph Hair: Métodos Multivariados.
5. Probabilidad y estadística: aplicaciones y métodos. Canavos, George C [1988].
6. Diseño y análisis de experimentos. Montgomery, Douglas C [1991].
7. Cid, Luis (2011): Apuntes del curso “Diseño de experimentos” Ing. Estadística UBB.
8. Vargas, Gilda (2009): Apuntes del curso “Control estadístico de Procesos” Ing. Estadística UBB.
9. Neter, et al., 1996
10. Agresti, 2002

# 13. Anexos

## ANEXO 1

Curva ROC para el biomarcador TNFa.

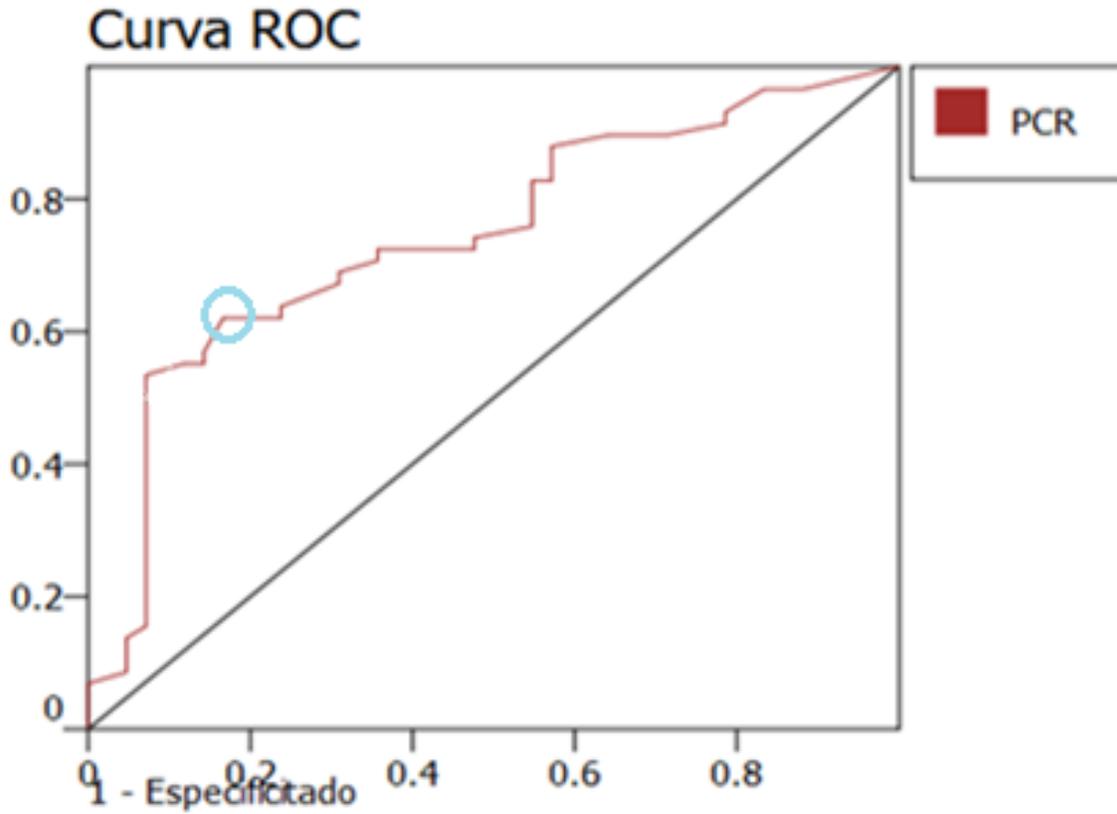


Área	Error Estándar	Sig. Asintótica	Asymp. 95% Intervalo de Confianza	
			Límite Inferior	Límite Superior
,54	,06	,53	,44	,63

Positivo si es mayor o igual a	Sensibilidad	1 - Especificidad
1,70	1,00	1,00
3,35	1,00	,96
3,60	1,00	,96
4,65	,10	,05
7,75	,07	,05
7,95	,07	,05
21,65	,05	,04
28,40	,05	,04
30,05	,05	,04
95,65	,00	,02
188,30	,00	,00

## ANEXO 2

Curva ROC para el biomarcador PROTEÍNA C REACTIVA.



Área	Error Estándar	Sig. Asintótica	Asymp. 95% Intervalo de Confianza	
			Límite Inferior	Límite Superior
,74	,05	,00	,66	,83

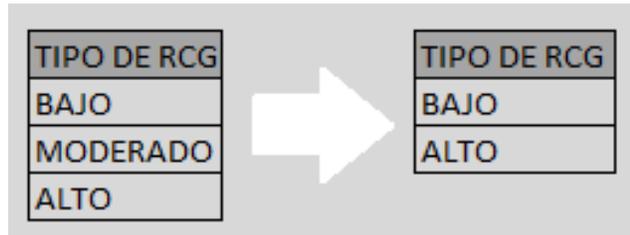
Positivo si es mayor o igual a	Sensibilidad	1 - Especificidad
- ,81	1,00	1,00
,37	,97	,88
,37	,97	,88
,49	,97	,83
,55	,93	,79
,56	,93	,79
,60	,91	,79
,69	,90	,71
,74	,90	,64
,87	,88	,57
,97	,84	,57
,97	,84	,57
1,00	,83	,57
1,02	,83	,57
1,03	,83	,57
1,08	,83	,55
1,10	,83	,55
1,11	,83	,55

1,11	,83	,55
1,12	,83	,55
1,12	,83	,55
1,30	,81	,55
1,32	,81	,55
1,33	,79	,55
1,34	,78	,55
1,35	,78	,55
1,40	,76	,55
1,42	,76	,55
1,57	,74	,48
1,58	,74	,48
1,71	,72	,48
1,75	,72	,48
1,76	,72	,45
1,83	,72	,43
1,93	,72	,40
1,97	,72	,40
2,00	,72	,38
2,04	,72	,36
2,09	,72	,36
2,13	,71	,36
2,18	,69	,31
2,30	,67	,31
2,73	,64	,24
2,75	,62	,24
2,84	,62	,21
3,10	,62	,21
3,16	,62	,21
3,19	,62	,17
3,48	,57	,14
3,50	,57	,14
3,50	,57	,14
3,54	,55	,14
3,77	,55	,14
3,83	,55	,14
3,88	,55	,12
3,96	,53	,07
4,35	,53	,07
4,46	,53	,07
4,47	,53	,07
4,69	,53	,07
4,70	,50	,07
4,84	,48	,07
4,84	,48	,07
4,90	,47	,07
4,91	,47	,07
4,91	,47	,07
4,98	,45	,07
5,03	,45	,07
5,10	,45	,07
5,22	,43	,07
5,22	,43	,07
5,30	,43	,07
5,45	,40	,07
5,76	,38	,07
5,87	,36	,07
6,10	,36	,07
6,17	,36	,07
6,21	,36	,07
6,42	,36	,07
6,59	,36	,07
6,67	,34	,07
6,68	,33	,07
6,83	,33	,07
6,91	,33	,07
7,58	,28	,07
7,87	,26	,07
9,39	,19	,07
10,07	,16	,07
10,35	,14	,05
11,06	,14	,05
11,32	,12	,05
12,23	,12	,05
12,84	,12	,05
17,11	,10	,05
18,72	,09	,05
18,77	,09	,05
18,93	,09	,05
25,44	,07	,00

## ANEXO 3

Para facilitar el modelo de regresión de predicción y la interpretación de éste, se ha recodificado la variable tipo de RCG, que originalmente tenía tres categorías, es decir,

Donde la “nueva” codificación contiene en el tipo de RCG ALTO a un RCG MODERADO Y ALTO.



Además la categoría 3 (RCG ALTO) original, contenía solo a seis observaciones, por lo que se determinó en conjunto con los investigadores patrocinantes de este proyecto de título que no era una buena categoría de pacientes para realizar una predicción con ella.

Para comprobar que estas dos categorías se pueden juntar sin ningún problema, se desarrollaron tablas ANOVA de un factor para cada biomarcador. Las hipótesis planteadas son las siguientes:

H0: La media del biomarcador para pacientes con riesgo bajo es igual a la media de pacientes con riesgo moderado y es igual a la media de pacientes con riesgo alto

H1: La media del biomarcador para pacientes con riesgo bajo es distinta a la media de pacientes con riesgo moderado y es distinta a la media de pacientes con riesgo alto

Por lo tanto para ver si se aceptan o no estas hipótesis, a continuación se presentan las salidas del software SPSS, que corresponden a las tablas ANOVA y test de comparación múltiple respectivamente, para cada uno de los biomarcadores en estudio.

Acá se puede apreciar de las tablas ANOVA que no existen pruebas suficientes para rechazar las hipótesis H0, ya que en todos los casos el valor- $p > 0.05$  con un 95% de confianza, es decir, no existen diferencias significativas entre pacientes con un tipo de RCG bajo, moderado y alto (1,2 y 3 respectivamente) en ninguno de los tres biomarcadores.

Para corroborar esta decisión se realizaron además test de comparación múltiple, que indican si es que hay diferencias significativas entre cada par de las categorías de la variable Tipo de RCG para cada uno de los biomarcadores.

**ANOVA**

TNFa

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	139,042	2	69,521	,167	,847
Intra-grupos	38761,731	93	416,793		
Total	38900,772	95			

**Comparaciones múltiples**

Variable dependiente: TNFa

HSD de Tukey

(I) Riesgo CV	(J) Riesgo CV	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
1	2	2,02602	5,27321	,922	-10,5338	14,5858
	3	5,31053	9,56044	,844	-17,4606	28,0817
2	1	-2,02602	5,27321	,922	-14,5858	10,5338
	3	3,28451	8,67962	,924	-17,3887	23,9577
3	1	-5,31053	9,56044	,844	-28,0817	17,4606
	2	-3,28451	8,67962	,924	-23,9577	17,3887

**ANOVA**

PCR

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	25,025	2	12,512	,102	,904
Intra-grupos	11586,993	94	123,266		
Total	11612,018	96			

**Comparaciones múltiples**

Variable dependiente: PCR  
HSD de Tukey

(I) Riesgo CV	(J) Riesgo CV	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
1	2	,26334	2,86351	,995	-6,5558	7,0825
	3	-1,85577	5,19923	,932	-14,2372	10,5257
2	1	-,26334	2,86351	,995	-7,0825	6,5558
	3	-2,11911	4,71766	,895	-13,3538	9,1155
3	1	1,85577	5,19923	,932	-10,5257	14,2372
	2	2,11911	4,71766	,895	-9,1155	13,3538

**ANOVA**

ROS

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	18,942	2	9,471	,523	,595
Intra-grupos	1612,202	89	18,115		
Total	1631,144	91			

**Comparaciones múltiples**

Variable dependiente: ROS  
HSD de Tukey

(I) Riesgo CV	(J) Riesgo CV	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
1	2	-1,15934	1,17938	,589	-3,9705	1,6518
	3	-,43729	2,03747	,975	-5,2937	4,4191
2	1	1,15934	1,17938	,589	-1,6518	3,9705
	3	,72205	1,81049	,916	-3,5934	5,0374
3	1	,43729	2,03747	,975	-4,4191	5,2937
	2	-,72205	1,81049	,916	-5,0374	3,5934

## ANEXO 4

La toma de decisiones clínicas es un proceso extremadamente complejo en el que deberá finalmente ser valorada la utilidad para el manejo del paciente de cualquier prueba diagnóstica. En este contexto, es imprescindible conocer detalladamente la exactitud de las distintas pruebas diagnósticas, es decir, su capacidad para clasificar correctamente a los pacientes en categorías o estados en relación con la enfermedad (típicamente dos: estar o no estar enfermo, respuesta positiva o negativa a la terapia...).

### Sensibilidad y Especificidad

Generalmente, la exactitud diagnóstica se expresa como sensibilidad y especificidad diagnósticas. Cuando se utiliza una prueba dicotómica (una cuyos resultados se puedan interpretar directamente como positivos o negativos), la **sensibilidad** es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como positivo respecto a la condición que estudia la prueba, razón por la que también es denominada fracción de verdaderos positivos (FVP). La **especificidad** es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como negativo. Es igual al resultado de restar a uno la fracción de falsos positivos (FFP).

Cuando los datos de una muestra de pacientes se clasifican en una tabla de contingencia por el resultado de la prueba y su estado respecto a la enfermedad, es fácil estimar a partir de ella la sensibilidad y la especificidad de la prueba (**tabla 1**).

		Verdadero Diagnóstico	
		Enfermo	Sano
Resultado de la Prueba	Prueba Positiva	Verdadero Positivo (VP)	Falso Positivo (FP)
	Prueba Negativa	Falso Negativo (FN)	Verdadero Negativo (VN)
		VP + FN	VN + FP
Sensibilidad	= $VP/(VP + FN)$ = FVP (fracción de verdaderos positivos)		
Especificidad	= $VN/(VN + FP)$ = FVN (fracción de verdaderos negativos) = 1 - FFP (fracción de falsos positivos)		