



Universidad Autónoma de Guerrero

Facultad de Matemáticas

Maestría en Matemáticas Aplicadas

**Construcción y validación de
índices de síndrome metabólico
para cada sexo en adultos
mexicanos: Un análisis factorial
confirmatorio**

Eduardo Pérez Castro

DIRECTORES DE TESIS:

MC. Cruz Vargas de León

Dr. Flaviano Godínez Jaimes

Enero 2022.

Dedicatoria

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología, por el apoyo económico recibido para la realización de mis estudios de Maestría.

A la Maestría en Matemáticas Aplicadas, por haberme brindado la oportunidad de realizar mis estudios de Maestría.

Abreviaturas

SM Síndrome Metabólico

OA Obesidad abdominal

RI Resistencia a la insulina

DM Diabetes Mellitus tipo 2

ECV Enfermedad Cardiovascular

NCEP Programa Nacional de Educación para el Colesterol

AHA Asociación Estadounidense del Corazón

NHLBI Instituto Nacional del Corazón, los Pulmones y la Sangre

FID Federación Internacion de Diabetes

NHANES Encuesta Nacional de Examen de Salud y Nutrición

CC Circunferencia de Cintura

HDL-c Colesterol de lipoproteina de alta densidad

PAS Presión arterial sistólica

Tri Triglicéridos

Glu Glucosa

hs-CRP Proteína C-reactiva de alta sensibilidad

IMC Índice de Masa Corporal

HOMA-RI Modelo de homeostasis de resistencia a la insulina

PAL Producto de acumulación de lípidos

SPISE Estimador de sensibilidad a la insulina de un solo punto

IAV Índice de adiposidad visceral

TyG Índice de triglicéridos y glucosa

TG/HDL Razón-triglicéridos/colesterol HDL

IRC Índice de redondez corporal

GC Grasa Corporal

TAV Tejido Adiposo Visceral

QUICKI Índice cuantitativo de comprobación de la sensibilidad a la insulina	CCC Coeficiente de Correlación de Concor-dancia
METS-IR Score Metabólico para la Resis-tencia a la Insulina	ROC Curva Característica Operativa del Receptor
METS-VF Score Metabólico para la Grasa Visceral	AUC Área Bajo la Curva
AFC Análisis Factoral Confirmatorio	ENSANUT Encuesta Nacional de Salud y Nutrición
NFI Índice de ajuste Normalizado	INEGI Instituto Nacional de Estadística y Geografía
CFI Índice de Ajuste Comparativo	
AIC Criterio de información de Akaike	INSP Instituto Nacional de Salud Pública
GFI Índice de bondad de ajuste comparativo	UNAM Universidad Nacional Autónoma de México
AGFI Índice ajustado de bondad de ajuste	
RMSEA Raíz del error cuadrático medio de aproximación	ETBC Estudio transversal de base comuni-taria
CCI Coeficiente de Correlación Intraclase	

Resumen

El síndrome metabólico (SM) es un conjunto de anormalidades metabólicas consideradas como factores de riesgo para desarrollar enfermedad cardiovascular y diabetes. Generalmente el diagnóstico del SM es de tipo binario, Sí o No. Este se basa en puntos de cortes para las variables que definen el SM: glucosa, presión arterial, triglicéridos, colesterol HDL y adiposidad central. Existen índices que proporcionan una medición continua del SM para la población de los Estados Unidos, pero no son válidos para la población mexicana debido a las diferentes características fisiológicas y antropométricas. En este trabajo se proponen a partir de datos de la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2018 índices del SM en la población adulta de México tanto en hombres como mujeres usando diferentes medidas de adiposidad, posteriormente se evalúan mediante los coeficientes de correlación intraclase y de concordancia. Se utilizaron curvas características del receptor (ROC) para determinar la capacidad de los índices para identificar SM.

Índice general

Índice de figuras	VII
Índice de Cuadros	VIII
1. Introducción	1
1.1. Síndrome Metabólico	1
1.2. Planteamiento del problema	3
1.3. Objetivos	3
2. Marco teórico	5
2.1. Criterios diagnósticos del Síndrome Metabólico	5
2.1.1. Índices antropométricos	7
2.2. Marcadores subrogados de resistencia a la insulina y obesidad	10
2.3. Análisis factorial confirmatorio	14
2.3.1. Estimación de parámetros	17
2.3.2. Estadísticos de bondad de ajuste	18
2.4. Medidas de concordancia	22
2.4.1. Coeficiente de correlación intraclase	23
2.4.2. Coeficiente de correlación de concordancia	26
2.4.3. Gráficos de Bland y Altman	28
2.5. Curvas ROC	30
2.5.1. Definiciones previas	30
2.5.2. Sensibilidad y especificidad	31
2.5.3. Pruebas continuas	33

2.5.4.	Curva Característica Operativa del Receptor	35
2.5.5.	Cálculo del area bajo la curva	38
2.5.6.	Contraste para la aleatoriedad de la prueba	41
2.5.7.	Comparación de pruebas	42
3.	Metodología	44
3.1.	Bases de datos	44
3.1.1.	ENSANUT 2018	44
3.1.2.	Estudio transversal de base comunitaria	44
3.2.	Construcción de los índices	45
3.2.1.	Análisis factorial confirmatorio	45
3.2.2.	Validación de los índices propuestos	47
4.	Resultados	48
4.1.	Derivación de los índices propuestos y validación interna en ENSANUT	48
4.1.1.	Análisis exploratorio	48
4.1.2.	Estimación de los modelos AFC	50
4.1.3.	Ecuaciones para SM en adultos mexicanos	52
4.2.	Validación de los índices propuestos en la base ETBC	53
5.	Discusión y Conclusiones	64
5.1.	Discusión	64
5.2.	Conclusiones	67
	Referencias	67

Índice de figuras

1figure.caption.8	
2.1. Gráfico dispersión de los pares de datos alrededor de la línea de 45° (izquierda) y gráfico de Bland y Altman (derecha) para las medidas de los evaluadores 1 y 4.	30
2.2. Clasificación de un individuo como sano o enfermo en función de un límite fijo c	34
2.3. Curva ROC de una prueba diagnóstica continua.	36
3.1. Modelos AFC para SM.	46
4.1. Diagrama de flujo para la obtención de la base de datos para la formulación de MetSx.	49
4.2. Gráficos de Bland Altman para los índices en mujeres.	58
4.3. Gráficos de Bland Altman para los índices en hombres.	59
4.4. AUCs de los índices en mujeres.	62
4.5. AUCs de los índices en hombres.	63

Índice de Cuadros

2.1. Valores étnicos específicos para la circunferencia de la cintura según la Federación Internacional de Diabetes (21).	7
2.2. Criterios diagnósticos de SM.	8
2.3. Índices para la RI y sus rangos de referencia (56–61).	13
2.4. Interpretación del RMSEA (71).	21
2.5. Estadísticos de bondad de ajuste y criterios de referencia (73).	22
2.6. Modelo ANOVA de efectos aleatorios de una sola vía (Ecuación 2.20.)	24
2.7. Ejemplo de Shrout y Fleiss (1979).	29
2.8. Clasificación de los resultados de las pruebas por estado de enfermedad. . . .	32
2.9. Clasificación de los resultados de las pruebas continuas por estado de enfermedad.	35
2.10. Interpretación de Swets para valores del AUC.	37
4.1. Características de los adultos de 20 a 59 años de ENSANUT 2018 sobre componentes del SM y otros parámetros bioquímicos ($n = 6,567$).	50
4.2. Índices de bondad de ajuste para los modelos del AFC.	51
4.3. Estimación del análisis factorial confirmatorio con variables transformadas y estandarizadas.	52
4.4. MetSx por sexo para adultos mexicanos.	53
4.5. Estadísticas descriptivas de la base de datos CBHS para la validación de MetSx.	55
4.6. Prevalencia de SM según los diferentes criterios.	56
4.7. Correlaciones (IC del 95 %) de parámetros bioquímicos y marcadores subrogados con MetSx.	57
4.8. Coeficientes de concordancia y correlación intraclase entre índices.	60
4.9. AUC de MetSx para el diagnóstico binario de SM.	61

1. Coeficientes de concordancia y correlación intraclassa entre MetSx-IRC y MetSx- IMC.	83
--	----

Introducción

1.1. Síndrome Metabólico

El síndrome metabólico (SM) es una condición patológica caracterizada por obesidad abdominal (OA), resistencia a la insulina (RI), hipertensión e hiperlipidemia (1). El SM se reconoció por primera vez en 1923. Kylin, un médico sueco, describió un síndrome que involucraba hipertensión, hiperglucemia e hiperuricemia (2). Posteriormente, en 1965 Avogaro y Crepaldi (3) describieron un síndrome que consistía en hipertensión, hiperglucemia y obesidad. Según Reaven, el mecanismo subyacente a este síndrome era la RI (4). El SM está asociado con un riesgo para desarrollar Diabetes Mellitus Tipo 2 (DM) y Enfermedad Cardiovascular (ECV) (5).

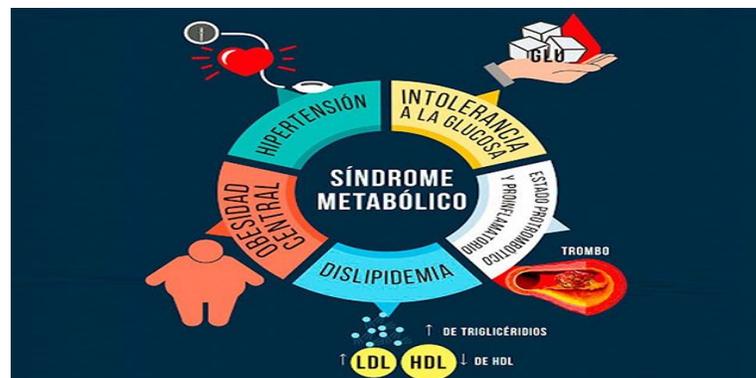


Figura 1.1: Síndrome Metabólico ¹.

El mundo se enfrenta a una creciente epidemia de SM. En 2017 se estimó que aproximadamente una cuarta parte de la población adulta mundial (mil millones) estaba afectada por SM (1), ese mismo año el SM afectó al 20% de la población de América del Norte, al 25% de la población europea y aproximadamente al 15% de la población china (1). En México, aproximadamente 6 de cada 10 adultos mexicanos tiene SM (57.9% en hombres y 63.2% en mujeres)(6).

SM se diagnostica tradicionalmente con una clasificación binaria (presente/ausente) basada en los criterios propuestos por el Programa Nacional de Educación sobre el Colesterol, Programa de Tratamiento para Adultos III (ATP-III); la Asociación Estadounidense del Corazón/Instituto Nacional del Corazón, los Pulmones y la Sangre (AHA/NHLBI); y la Federación Internacional de Diabetes (FID). Cada criterio requiere que una persona tenga anomalías en al menos tres componentes, pero los criterios difieren en que tiene niveles de corte ligeramente diferentes para designar una anomalía (7). Gurka y colaboradores en los años 2012 y 2014 propusieron índices de SM específicos por sexo y etnia para adolescentes (entre 12 y 19 años) y adultos (entre 20 y 64 años) estadounidenses (8,9). Estos índices fueron derivados a partir de datos de la Encuesta Nacional de Examen de Salud y Nutrición (NHANES). Estos índices utilizaron los cinco componentes tradicionales del SM: circunferencia de la cintura (CC), colesterol de lipoproteínas de alta densidad (HDL-c), presión arterial sistólica (PAS), triglicéridos (Tri) y glucosa (Glu), además demostraron que están altamente correlacionados con otros marcadores subrogados de SM, incluida la proteína C reactiva altamente sensible (hs-CRP), el ácido úrico y el modelo de homeostasis de resistencia a la insulina (HOMA-RI) (8,9). Cuatro años más tarde, los mismos autores proponen otro índice de SM para adultos, el cual utiliza el índice de masa corporal (IMC) en lugar de la CC como medida de adiposidad. Este segundo índice proporcionó un poder similar para predecir enfermedades futuras como el índice basado en la CC y un alto grado de concordancia con el anterior (10).

Las ecuaciones de SM que se presentan corresponden a los hispano estadounidenses.

¹Publicado por Calevit, Calevit. Síndrome Metabólico y Covid-19 [Internet]. [Citado el 23 de noviembre de 2021]. Disponible en: <https://www.calevit.cl/sindrome-metabolico-y-covid-19/>

SM-CC-M	$-7.7641 + 0.0162CC - 0.0157HDL + 0.0084PAS + 0.8872 \ln(\text{Tri}) + 0.0206\text{Glu}$
SM-CC-H	$-5.5541 + 0.0135CC - 0.0278HDL + 0.0054PAS + 0.8340 \ln(\text{Tri}) + 0.0105\text{Glu}$
SM-IMC-M	$-7.1844 + 0.0333IMC - 0.0166HDL + 0.0085PAS + 0.8625 \ln(\text{Tri}) + 0.0221\text{Glu}$
SM-IMC-H	$-4.8198 + 0.0355IMC - 0.0303HDL + 0.0051PAS + 0.7835 \ln(\text{Tri}) + 0.0104\text{Glu}$

1.2. Planteamiento del problema

Los componentes más frecuentes del SM en la población mexicana son OA, HDL-c bajo y niveles altos de Tri (6). La OA y las dislipidemias por separado o combinadas son los componentes del SM más prevalentes en la población mexicana, lo que expone a esta población a un alto riesgo de DM (11, 12) y ECV (13). Además, la hipertrigliceridemia en la población mexicana es más alta que en otros países en desarrollo y desarrollados, como India, Nigeria, China, Japón y Estados Unidos (14). Debido a que las características fisiológicas y antropométricas de la población mexicana son diferentes a las de la población estadounidense, es necesario tener índices específicos por sexo para medir la severidad del SM en adultos mexicanos.

1.3. Objetivos

Objetivo general

Proponer índices de SM específicos para adultos mexicanos (MetSx) por sexo, utilizando CC, IMC o IRC como medida de adiposidad.

Objetivos específicos

- Proponer índices de SM específicos para adultos mexicanos.
- Evaluar la concordancia entre los índices propuestos.
- Evaluar la asociación entre los índices propuestos y varios marcadores subrogados de resistencia a la insulina y obesidad como evaluación del modelo homeostático para

la resistencia a la insulina (HOMA-IR), producto de acumulación de lípidos (PAL), estimador de sensibilidad a la insulina de un solo punto (SPISE), índice de adiposidad visceral (IAV), índice de triglicéridos y glucosa (TyG) y razón triglicéridos/colesterol-HDL (TG/ HDL-c), entre otros.

Marco teórico

2.1. Criterios diagnósticos del Síndrome Metabólico

El primer criterio de diagnóstico propuesto fue de la Organización Mundial de la Salud (OMS) y como requisito principal para el diagnóstico de SM, era que una persona presentara RI (15) y dos factores de riesgo adicionales: obesidad, hipertensión, dislipidemia y albuminuria.

En 2001, el ATP-III del Programa Nacional de Educación sobre el Colesterol (NCEP) publicó sus propios criterios para definir el SM (16). Según la definición de NCEP ATP-III, el SM está presente si se cumplen tres o más de los siguientes cinco criterios: CC mayor de 102 cm (hombres) o 88 cm (mujeres), PAS mayor a 130 *mmHg*, Tri superior a 150 mg/dl, HDL-c inferior a 40 *mg/dl* (hombres) o 50 *mg/dl* (mujeres) y Glu superior a 100 *mg/dl* en ayunas. La definición de NCEP ATP-III es uno de los criterios de SM ampliamente utilizados. Incorpora las características clave de RI, OA, dislipidemia aterogénica e hipertensión (17). Utiliza mediciones y resultados de laboratorio que están fácilmente disponibles para los médicos, lo que facilita su aplicación clínica y epidemiológica. También es simple y fácil de recordar. Es importante destacar que no requiere que se cumpla ningún criterio específico; sólo que se cumplan al menos tres de los cinco criterios. Por lo tanto, la definición no se basa en ninguna noción preconcebida de la causa subyacente del SM, ya sea la RI o la obesidad.

La FID tuvo como objetivo crear una definición sencilla y clínicamente útil para identificar

a las personas en cualquier país del mundo con alto riesgo de ECV y DM que permitieran estudios epidemiológicos comparativos. Esto resultó en la definición de la FID de SM en 2005. La obesidad central, según las mediciones de la CC específicas de la etnia/región era un requisito para el síndrome. Además, la definición requería la presencia de dos de los siguientes cuatro elementos: Tri \geq 150 mg/dL, HDL-c $<$ 40 mg/dL en hombres o $<$ 50 mg/dL en mujeres, PAS \geq 130 mmHg, Glu en ayunas $>$ 100 mg/dl incluidos los diabéticos y aquellos con un diagnóstico o tratamiento previo de cualquiera de estas afecciones (DM o hipertensión arterial) (18).

En 2005, la Asociación Estadounidense del Corazón (AHA)/Instituto Nacional del Corazón, los Pulmones y la Sangre (NHLBI) también sugirió criterios para el diagnóstico del SM. Su definición se basó en los criterios ATP-III y requería tres de cualquiera de los cinco criterios siguientes: CC elevada (\geq 102 cm en hombres y \geq 88 cm en mujeres), Tri \geq 150 mg/dL y HDL-c $<$ 40 mg/dL en hombres y $<$ 50 mg/dL en mujeres, PAS \geq 130 mmHg y Glu en ayunas elevada $>$ 100 mg/dL. Como sugirió la IDF, las CC específicas de la etnia/regiones (Cuadro 2.1) se tuvieron en cuenta al usar esta definición. Además, la alteración de la glucosa en ayunas se definió como $>$ 100 mg/dL, lo que también coincidía con las directrices de la FID (19).

El año 2009, representantes de la FID, AHA/NHLBI y de las guías del ATP-III discutieron resolver las diferencia entre las definiciones del SM, lo que permitió unificar criterios. Esta unificación de criterios fue publicada bajo el título de Harmonizing the Metabolic Syndrome, en la revista Circulation en su edición de diciembre del año 2009, donde se consideró al perímetro abdominal como uno más de los componentes diagnósticos del SM, no siendo prioridad su presencia para el diagnóstico. El SM debía ser definido como la presencia de tres componentes descritos por FID y AHA/ NHLBI, la cual considera la población y el país específico para la definición del corte de perímetro abdominal.

Una de las principales diferencias en los criterios es la OA. La FID considera que debiera ser el principal requisito para diagnosticar SM. Para el resto de los criterios es un componente de los cinco propuestos, pero no excluyente para diagnosticar SM. La mayoría de las investigaciones nivel internacional han sido realizados con los criterios del ATP III. Sin embargo, considerar al perímetro abdominal como criterio principal y excluyente en el diagnóstico de síndrome refuerza el hecho que la OA es la causa del desarrollo de RI y de los demás compo-

nentes SM (20). En el Cuadro 2.2 se presentan los puntos de cortes para diagnosticar SM de acuerdo a los criterios mencionados anteriormente.

Cuadro 2.1: Valores étnicos específicos para la circunferencia de la cintura según la Federación Internacional de Diabetes (21).

País/grupo étnico	Valores de la CC	
	Hombres	Mujeres
Europeos*	≥ 94 cm	≥ 80 cm
Asiáticos del Sur**	≥ 90 cm	≥ 80 cm
Chinos	≥ 90 cm	≥ 80 cm
Japoneses	≥ 85 cm	≥ 90 cm
Sudamericanos y centroamericanos	Use las recomendaciones del sur de Asia hasta que haya datos más específicos disponibles	
Africanos subsaharianos	Usar datos europeos hasta que haya datos más específicos disponibles	
Poblaciones del Mediterráneo oriental y Oriente Medio (árabe)	Usar datos europeos hasta que haya datos más específicos disponibles	

*En los EE. UU., los valores de ATP III (102 cm masculino; 88 cm femenino) se utilizan con fines clínicos, ** Basado en poblaciones de chinos, malayos e indios asiáticos.

2.1.1. Índices antropométricos

La grasa visceral juega un papel fundamental en la patogénesis del SM (22). Dado que la grasa visceral se concentra principalmente en el abdomen, durante mucho tiempo, las personas simplemente han examinado y manejado la obesidad visceral mediante índices antropométricos de obesidad, como la CC y el IMC.

Índice de masa corporal

Uno de los métodos más simples para evaluar la composición corporal es el IMC (23). El IMC es la razón entre el peso en kilogramos y la altura en metros al cuadrado.

$$\text{IMC} = \frac{\text{Peso(kg)}}{\text{Estatura(m)}^2}$$

Cuadro 2.2: Criterios diagnósticos de SM.

	ATP III	AHA/NHLBI	FID (Requerido)	Criterio FID armonizado
	CC	CC		CC
Obesidad	≥ 102 cm (Hombres) ≥ 88 cm (Mujeres)	≥ 102 cm (Hombres) ≥ 88 cm (Mujeres)	≥ 90 cm (Hombres) ≥ 80 cm (Mujeres)	≥ 90 cm (Hombres) ≥ 80 cm (Mujeres)
Glucosa	≥ 110 mg/dL ^a o diagnóstico previo de DM	≥ 100 mg/dL or diagnóstico previo de DM	≥ 100 mg/dL o diagnóstico previo de DM	≥ 100 mg/dL o tratamiento farmacológico para el control de la glucosa
Triglicéidos	≥ 150 mg/dL	≥ 150 mg/dL o tratamiento farmacológico	≥ 150 mg/dL tratamiento farmacológico	≥ 150 mg/dL o tratamiento farmacológico para controlar
PAS	≥ 130 mm Hg o diagnóstico previo de hipertensión	≥ 130 mm Hg ^b o diagnóstico previo de hipertensión	≥ 130 mm Hg o previous diagnosis of hypertension	≥ 130 mm Hg o diagnóstico previo de hipertensión
HDL-c	< 40 mg/dL (Hombres) < 50 mg/dL (Mujeres)	< 40 mg/dL (men) < 50 mg/dL (Mujeres)	< 40 mg/dL (Hombres) < 50 mg/dL (Mujeres)	< 40 mg/dL (Hombres) < 50 mg/dL (Mujeres)
Diagnostico	Cualquiera de los 3 componentes	Cualquiera de los 3 componentes	CC+2 de los 3 componentes	Cualquiera de los 3 componentes

^amg/dL: miligramos por decilitro

^bmm Hg: milímetros de mercurio

El IMC se utiliza para evaluar el estado nutricional de diferentes tipos de población, ampliamente utilizado por su buena correlación con la masa corporal ($r \approx 0.8$) y baja correlación con la altura. El IMC se clasifica en ($< 18.5 \text{ kg/m}^2$), peso normal ($18.5 - 24.9 \text{ kg/m}^2$), sobrepeso ($25.0 - 29.9 \text{ kg/m}^2$), obesidad grado I ($30.0 - 34.9 \text{ kg/m}^2$), obesidad grave o grado II ($35.0 - 39.9 \text{ kg/m}^2$) y obesidad mórbida o grado III ($> 40.0 \text{ kg/m}^2$) (OMS, 2017). El IMC se asocia regularmente con el riesgo de ECV (24,25), incluida la cardiopatía isquémica (26) y el de DM (27).

Las limitaciones del IMC son su baja precisión en la evaluación del tejido adiposo y el hecho de que no distingue el dimorfismo sexual y las diferencias étnicas en la adiposidad, la distribución del tejido adiposo y la composición corporal relacionada con la edad (28).

Circunferencia de cintura

La CC es una medida simple de la circunferencia alrededor de la cintura entre la costilla más baja y la cresta ilíaca.

Numerosos estudios han informado que las mediciones de CC son similares, si no más precisas, para predecir el riesgo de mortalidad (24), el riesgo de ECV (29,30) y resultados de salud (31) en comparación con el IMC. Ross y colaboradores en 2008 (32) consideraron los aspectos prácticos para medir la CC. A pesar de que la CC parece estar ampliamente recomendada para uso práctico y clínico como un concepto relativamente fácil, aún no existe un protocolo estándar para medir la CC.

La desventaja de la CC es que las personas más altas tienen circunferencias más grandes. Además, los valores de la CC difieren entre los grupos étnicos (33). La principal desventaja de la CC es su falta de capacidad para diferenciar la acumulación de grasa subcutánea de la visceral (34).

Índice de redondez corporal

Un índice desarrollado recientemente por Thomas y colaboradores en el 2013 (35) es el índice de redondez corporal (IRC), que combina la altura y la CC para predecir el porcentaje de grasa corporal (GC) y evaluar el estado de salud del paciente (36). El IRC se derivó para cuantificar la forma del cuerpo independientemente de la estatura, como predictor del porcentaje de GC y del tejido adiposo visceral (TAV).

Para desarrollar este índice, los autores modelaron la forma del cuerpo humano como una elipse o una forma ovalada, que captura la circunferencia del cuerpo en relación con la altura (redondez del cuerpo) (36). Esta nueva medición se deriva geoméricamente de la forma corporal y proporciona predicciones consistentes del porcentaje de GC y TAV. Los valores de IRC oscilan entre 1 y 16, y los individuos más redondos tienen valores más grandes.

Se ha reportado que el IRC tiene una capacidad superior para predecir SM y riesgo cardio metabólico en comparación con otros índices antropométricos (37–40). El IRC ha mostrado un potencial razonable para servir como un índice alternativo para la RI y los factores inflamatorios y una equivalencia con los índices de obesidad tradicionales (41). Además, la capacidad del IRC para predecir la dislipidemia e hipertensión fue comparable a los indicadores clásicos de obesidad (42–44). El IRC ha mostrado tener una capacidad predictiva de hiperuricemia mayor que el IMC y un poder predictivo similar que la CC en mujeres (45).

Una ventaja del IRC en relación al IMC, es la mejora en las predicciones del TAV en los pacientes, que es la más peligrosa para la salud. El IRC está relacionado positivamente con la mortalidad y otros marcadores de salud (como el cardiovascular y el desarrollo de DM) (35). El IRC es dado por la ecuación

$$\text{IRC} = 364.2 - 365.5 \sqrt{1 - \left(\frac{(\text{CC}(\text{m})/2\pi)^2}{(0.5 \text{ Talla (m)})^2} \right)}$$

2.2. Marcadores subrogados de resistencia a la insulina y obesidad

La insulina es un regulador clave de la homeostasis de la glucosa. La RI conduce a una intolerancia a la Glu y juega un papel fisiopatológico importante en el desarrollo de la DM. Además, la RI conduce a muchas de las anomalías metabólicas asociadas con el SM. Es probable que los pacientes con RI tengan niveles de Glu plasmática en ayunas alterados, lo que a su vez aumenta la prevalencia de LDL más pequeñas, densas y aterogénicas (46). La pinza euglucémica hiperinsulinémica, es el estándar de oro para estimar sensibilidad a la insulina, pero es una técnica muy compleja e invasiva. Varios índices subrogados se han desarrollado para simplificar y mejorar la determinación de la RI (47).

La adiposidad visceral se asocia con trastornos cardiovasculares y metabólicos (48) que

incluyen RI (49), DM (50), dislipidemia (51), enfermedad del hígado graso no alcohólico (52), hipertensión (53) y varios tipos de cánceres (54). La evaluación de la adiposidad visceral se recomienda en muchas situaciones y se ha convertido en una práctica clínica habitual. Sin embargo, las mediciones más precisas de la adiposidad visceral (resonancia magnética, tomografía computarizada) rara vez están disponibles para la mayoría de los médicos, ya que las técnicas de imagen involucradas son costosas, pueden implicar exposición a radiación o requerir entrenamiento intensivo (55). Por tanto, las técnicas de imagen son sustituidas con frecuencia por índices subrogados de adiposidad visceral. La CC y el índice cintura-cadera (ICC) son los índices subrogados de adiposidad visceral más utilizados.

En este apartado se presentan algunos de los marcadores subrogados de RI y obesidad.

Modelo de homeostasis de resistencia a la insulina

La evaluación del modelo de homeostasis (HOMA) fue propuesto en 1985 por Matthews y colaboradores (56). Se ha utilizado ampliamente como una estimación de la RI en diferentes tipos de estudios transversales, longitudinales y prospectivos. Es una de las medidas subrogadas de RI más utilizada y se basa en medidas sencillas de insulina¹ y Glu² en ayunas.

$$\text{HOMA} - \text{RI} = \frac{\text{Ins}(\mu\text{UI/L}) \times \text{Glu}(\text{mg/dL})}{18 \times 22.5(\text{mg/mldL})} \quad (2.1)$$

Estimador de sensibilidad a la insulina de un solo punto

El estimador de sensibilidad a la insulina de punto único (SPISE) propuesto por Paulmichl y colaboradores en el año 2016 (57) es un biomarcador de sensibilidad a la insulina basado en el IMC, Tri y HDL-c.

$$\text{SPISE} = \frac{600 \times \text{HDL}(\text{mg/dL})^{0.185}}{\text{Tri}(\text{mg/dL})^{0.2} \times \text{IMC}^{1.338}} \quad (2.2)$$

Índice de comprobación cuantitativa de la sensibilidad a la insulina

Este índice propuesto por Katz en el año 2000 se basa en el mismo principio que HOMA y se define como el inverso de la suma de los logaritmos de la insulina en ayunas y la Glu en ayunas (58).

¹ μUI : Microunidades internacionales por mililitro

² mg/dL : miligramos por decilitro

$$\text{QUICKI} = \{\ln[\text{Ins}(\mu\text{UI/L})] + \ln[\text{Glu}(\text{mg/dL})]\}^{-1} \quad (2.3)$$

Score Metabólico para la Resistencia a la Insulina

En 2018, Bello-Chavolla y colaboradores (59) desarrollaron un índice para determinar la RI. El cálculo se realiza mediante mediciones en ayuno de Glu, HDL-c, Tri y el IMC. METS-IR es una puntuación para evaluar el riesgo cardiometabólico en sujetos sanos y en riesgo.

$$\text{METS} - \text{IR} = \frac{\ln(2 \times \text{Glu}(\text{mg/dL}) + \text{Tri}(\text{mg/dL})) \times \text{IMC} (\text{kg/m}^2)}{\ln(\text{HDL}(\text{mg/dL}))} \quad (2.4)$$

Índice Triglicéridos/HDL-colesterol

La razón triglicéridos/colesterol HDL (TG/HDL) se considera un parámetro subrogado de RI (60).

$$\text{TG/HDL} = \frac{\text{Tri}(\text{mg/dL})}{\text{HDL}(\text{mg/dL})} \quad (2.5)$$

Índice Triglicéridos-Glucosa

Guerrero y colaboradores (61) propusieron un nuevo índice sencillo y económico para evaluar la RI.

$$\text{TyG} = \frac{\ln(\text{Tri}(\text{mg/dL}) \times \text{Glu}(\text{mg/dL}))}{2} \quad (2.6)$$

El Cuadro 2.3 presenta los rangos de valores de los diferentes índices para RI. El índice SPISE, varía de acuerdo a la población de estudio, en la población europea el punto de corte para la RI es de 6.61 (57), en un estudio realizado en el norte de la India el punto de corte fue de 5.82.

Cuadro 2.3: Índices para la RI y sus rangos de referencia (56–61).

Índice	Rangos de referencia
	Normal: ≤ 1.6
HOMA-IR	RI temprana: 1.7-2.4 RI: ≥ 2.5
QUICKY	Normal: > 0.339 RI: ≤ 0.339
SPISE	RI: >6.61
METS-IR	RI: >50
TG/HDL	RI: >2
TyG	RI: >8

Producto de acumulación de lípidos

El producto de acumulación de lípidos (PAL) fue introducido por primera vez por Kahn (los datos del estudio provienen de la tercera Encuesta Nacional de Examen de Salud y Nutrición, NHANES III) en el 2005 (62). PAL combina las medidas de la CC y los niveles de Tri en ayunas, lo que refleja los cambios anatómicos y fisiológicos asociados con la sobreacumulación de lípidos. PAL estuvo estrechamente asociado con ECV, DM y SM y superó al IMC para identificar estas enfermedades³.

$$\begin{aligned} \text{PAL}_{\text{Mujeres}} &= [\text{CC}(\text{cm}) - 58] \times \text{Tri}(\text{mmol/L}) \\ \text{PAL}_{\text{Hombres}} &= [\text{CC}(\text{cm}) - 65] \times \text{Tri}(\text{mmol/L}) \end{aligned} \tag{2.7}$$

Índice de adosidad visceral

El Índice de Adiposidad Visceral (IAV), utiliza parámetros simples tanto antropométricos (IMC y CC) como funcionales (Tri y HDL-c). Este índice, mostró una fuerte asociación independiente con eventos cardiovasculares y cerebrovasculares (63) y un mejor poder predictivo para DM, que sus componentes individuales (CC, IMC, Tri y HDL) (63).

³*mmol/L*: milimoles por cada litro

$$\begin{aligned}
IAV_{Mujeres} &= \left(\frac{CC(\text{cm})}{(1.89 \times \text{IMC} (\text{kg}/\text{m}^2)) + 36.58} \right) \times \left(\frac{\text{Tri}(\text{mmol}/\text{L})}{0.81} \right) \times \left(\frac{1.52}{\text{HDL}(\text{mmol}/\text{L})} \right) \\
IAV_{Hombres} &= \left(\frac{CC(\text{cm})}{(1.88 \times \text{IMC}) + 39.68} \right) \times \left(\frac{\text{Tri}(\text{mmol}/\text{L})}{1.03} \right) \times \left(\frac{1.31}{\text{HDL}(\text{mmol}/\text{L})} \right)
\end{aligned} \tag{2.8}$$

Score metabólico para la grasa visceral

El estándar de oro para evaluar el tejido adiposo visceral (TAV) es la resonancia magnética (RM), pero es cara y puede no ser adecuada en la práctica clínica general y en estudios de cohortes grandes (64). El Score Metabólico para la Grasa Visceral (METS-VF) se desarrolló en 2019 en México y se encontró positivamente asociado con el riesgo de hipertensión y DM. Este índice toma en cuenta la RI y combina los niveles de Glu, Tri, HDL-c, IMC, razón cintura-altura (RCA), la edad y el sexo. En sujetos con obesidad metabólicamente sana, un valor de METS-VF ≥ 7.20 tenían un mayor riesgo de desarrollar DM.

$$\begin{aligned}
\text{METS - VF} &= 4.466 + 0.011(\ln(\text{METS} - \text{IR}))^3 \\
&\quad + 3.239(\ln(\text{RCA}))^3 + 0.319(\text{Sexo}) + 0.594 \ln(\text{Edad})
\end{aligned} \tag{2.9}$$

La ecuación para determinar el VAT es:

$$\text{VAT} = e^{4.466+0.011[(\ln(\text{METS}-\text{IR}))^3]+3.239[(\ln(\text{RCA}))^3]+0.319(\text{Sexo})+0.594(\ln(\text{Edad}))} \tag{2.10}$$

donde Sexo = 1(0) si es hombre (mujer) y $\text{RCA} = \frac{\text{CC}}{\text{Altura}}$ denota la razón cintura-altura.

2.3. Análisis factorial confirmatorio

Un Análisis Factorial consiste en determinar el número y la naturaleza de las variables latentes (conocidas también como factores o constructos) que pueden ser generadas por un conjunto de variables medidas. El Análisis Factorial Exploratorio (AFE) consiste en determinar el número de factores y las variables observadas que lo forman para explicar de forma correcta las relaciones existentes, dado un conjunto de variables medidas. El objetivo del Análisis Factorial Confirmatorio (AFC) es el de confirmar una hipótesis sobre los Factores o Variables Latentes a partir de un conjunto de variables medibles.

A diferencia del AFE, en el AFC el modelo es construido de antemano, el investigador es el que establece el número de factores, qué factores están correlacionados y qué variables observadas miden cada factor (65).

Las variables latentes son las representaciones de los conceptos en los modelos de medición, y para su construcción es necesaria una definición teórica que explique el significado de un concepto, y así poder identificar las variables latentes para representarlo, definir los indicadores de las variables latentes, y especificar la relación entre las medidas y las variables latentes (66).

El modelo matemático del AFC está dado por:

$$\mathbf{X} = \mathbf{\Lambda}\xi + \delta \quad (2.11)$$

donde

- \mathbf{X} es un vector $q \times 1$, q es el número de variables observadas.
- ξ es un vector $s \times 1$, el modelo contiene s factores comunes o variables latentes.
- $\mathbf{\Lambda}$ es una matriz $q \times s$ que contiene las cargas de los factores de las variables latentes.
- δ es el vector $q \times 1$ que contiene los errores o factores específicos de las variables medidas.

Una observación importante es que el número de variables medidas debe ser mayor que el número de variables latentes ($q > s$). Este principio parte de la idea que se desea reducir el número de variables o encontrar una dimensión menor.

De la Ecuación 2.11 se puede notar la relación existente entre los factores y las variables observadas. Se asume que

$$E[\mathbf{X}] = \mathbf{0}, \quad E[\xi] = \mathbf{0}, \quad E[\delta] = \mathbf{0} \quad (2.12)$$

la varianza de \mathbf{X} , Σ es:

$$\Sigma = E[\mathbf{X}\mathbf{X}'] = E[(\mathbf{\Lambda}\xi + \delta)(\mathbf{\Lambda}\xi + \delta)'] \quad (2.13)$$

Ya que la transpuesta de una suma de matrices es la suma de las matrices transpuestas y que la transpuesta de un producto es el producto de las matrices transpuestas en orden inverso, de la Ecuación 2.13 se tiene que:

$$\Sigma = E[\mathbf{X}\mathbf{X}'] = E[(\mathbf{\Lambda}\xi + \delta)(\xi'\mathbf{\Lambda}' + \delta')] \quad (2.14)$$

El argumento de la esperanza se puede escribir como:

$$(\mathbf{\Lambda}\xi + \delta) (\xi' \mathbf{\Lambda}' + \delta') = \mathbf{\Lambda}\xi\xi' \mathbf{\Lambda}' + \mathbf{\Lambda}\xi\delta' + \delta\xi' \mathbf{\Lambda}' + \delta\delta' \quad (2.15)$$

Al aplicar la esperanza a la Ecuación 2.15 se tiene:

$$\begin{aligned} E[\mathbf{\Lambda}\xi\xi' \mathbf{\Lambda}'] + E[\mathbf{\Lambda}\xi\delta'] + E[\delta\xi' \mathbf{\Lambda}'] + E[\delta\delta'] \\ = \mathbf{\Lambda}E[\xi\xi'] \mathbf{\Lambda}' + \mathbf{\Lambda}E[\xi\delta'] + E[\delta\xi'] \mathbf{\Lambda}' + E[\delta\delta'] \end{aligned} \quad (2.16)$$

$\mathbf{\Lambda}$ es una matriz de escalares ya que son los parámetros del modelo AFC, en otras palabras, es el efecto que tienen las variables latentes sobre las variables medidas y los parámetros estimados son variables aleatorias. Por otro lado, se tiene que la esperanza de un vector aleatorio multiplicado por su transpuesto es la varianza del vector (cuando el vector aleatorio tiene media cero), y la esperanza de la multiplicación de dos vectores aleatorios, uno de ellos transpuesto, es la covarianza de los vectores (cuando los vectores aleatorios tienen media cero). Además, supongamos que:

$$E[\xi\xi'] = \Phi, \quad E[\delta\delta'] = \Theta \quad (2.17)$$

Asumiendo que los factores comunes y específicos no están correlacionados, es decir:

$$E[\xi\delta'] = E[\xi'\delta] = \mathbf{0} \quad (2.18)$$

Por las Ecuaciones (2.17) y (2.18) la matriz de varianzas-covarianzas de las variables medidas es igual:

$$\Sigma = \mathbf{\Lambda}\Phi\mathbf{\Lambda}' + \Theta \quad (2.19)$$

La Ecuación (2.19) es importante ya que a partir de ella se determinaran los estimadores del modelo y se sabrá si el modelo es identificable o no. Como observación se tiene:

1. La matriz Σ que es de orden q tiene $\frac{q(q+1)}{2}$ elementos no redundantes, ya que es una matriz de varianzas-covarianzas, lo que implica que es una matriz simétrica con $\frac{q(q+1)}{2}$ elementos posiblemente distintos.
2. La matriz de varianzas-covarianzas de los factores comunes Φ , tiene $\frac{s(s+1)}{2}$ elementos no redundantes.

3. Para la matriz de varianzas-covarianzas Θ de los factores específicos se tienen $\frac{q(q+1)}{2}$ elementos no redundantes. Pero considerando los supuestos del modelo de factores, en donde los errores son independientes, esta matriz se reduce a una matriz diagonal.

Por lo tanto, los $\left(\frac{q(q+1)}{2}\right)$ elementos de la matriz Σ se expresan en función de $\left[qs + \frac{s(s+1)}{2} + \frac{q(q+1)}{2}\right]$ parámetros de las matrices Σ, Φ y Θ . De esta manera los parámetros que se deben estimar estarán asociados por la Ecuación 2.19.

2.3.1. Estimación de parámetros

La estimación de los parámetros del modelo AFC consiste en que la matriz de varianzas-covarianzas Σ que se define en la Ecuación 2.19 debe ser lo más cercano (o igual) a la matriz S (la matriz de varianzas-covarianzas muestrales). La matriz Σ está expresada en términos de los parámetros a estimar, por lo tanto, lo que se busca es una expresión en la cual los argumentos sean Σ y S y se pueda minimizar esta diferencia a través de esa expresión. De esta función se obtendrá un conjunto de parámetros estimados que serán los elementos de la matriz $\hat{\Sigma}$.

La Ecuación 2.19 expresa el modelo que se desea estimar:

$$\Sigma = \Lambda\Phi\Lambda' + \Theta$$

A continuación se describen los métodos de estimación de máxima verosimilitud y estimación por la teoría de la distribución elíptica. Cabe mencionar que existen otros métodos para estimar los parámetros entre los cuales están: *mínimos cuadrados no ponderados (ULS, por sus siglas en inglés)* y *mínimos cuadrados generalizados*.

Estimación por máxima verosimilitud

Para usar este método de estimación las variables medidas deben cumplir el supuesto de normalidad. La función que debe ser minimizada para estimar los parámetros por el método de máxima verosimilitud (ML por sus siglas en inglés)

$$F_{ML}(S, \Sigma) = \text{tr}(S\Sigma^{-1}) + [\log|\Sigma| - \log|S|] - q$$

donde el segundo término consiste en restar los logaritmos naturales de los determinantes de las matrices de varianzas-covarianzas poblacional y muestral, y por último q que es el número

de variables medidas en el modelo. Cuando las matrices Σ y S sean muy parecidas entonces el producto de una de ellas y la inversa de la otra se aproximará a la matriz identidad; siendo su traza próxima a q . Por otra parte, los logaritmos de los determinantes se aproximarán a cero.

Estimación por la teoría de la distribución elíptica

La estimación basada en la teoría de la distribución elíptica tiene su fundamento en la distribución (Elliptical Distribution Theory, por sus siglas en inglés), se basa en la distribución de probabilidad de este nombre. La distribución normal multivariada es un caso particular de esta familia con parámetros de curtosis igual a cero. En este caso, la función a minimizar tiene la forma:

$$F_{EDT}(S; \Sigma) = \frac{1}{2}(k+1)^{-1} \text{tr} [(S - \Sigma) W^{-1}]^2 - \delta \text{tr} [(S - \Sigma) W^{-1}]^2$$

donde k y δ funciones de curtosis y de ponderación, y W es algún estimador consistente de Σ .

2.3.2. Estadísticos de bondad de ajuste

Existen dos conjuntos de estadísticos conocidos como: los índices comparativos de ajuste o de incremento y los índices de ajuste absoluto.

El primer conjunto de índices comparan el modelo teórico propuesto por el investigador contra el modelo de independencia, es decir, el modelo en el cual no existe relaciones entre las variables (las variables son no correlacionadas). Algunos de estos índices son: NFI, CFI, AIC, entre otros.

Los índices NFI y CFI deben estar lo más cercano a 1, o dicho de otra forma, si el investigador está comparando distintos modelos estimados, el mejor de ellos es aquel que su índice sea mayor al otro. Estos índices se encuentran en el intervalo $(0, 1)$. A continuación se da la expresión de ellos.

Índice NFI

El índice NFI (Normed Fit Index) compara el valor del estadístico χ^2 del modelo teórico con el modelo independiente:

$$NFI = \frac{\chi_{\text{indep}}^2 - \chi_{\text{teorico}}^2}{\chi_{\text{indep}}^2}$$

Este índice entre más cercano a 1, el ajuste del modelo es bueno. Algunos autores señalan que este índice tiende a subestimar el ajuste del modelo si las muestras son pequeñas, de tal manera que se presentan algunos índices modificados que solucionan estas situaciones.

Índice CFI

El índice CFI (Comparative Fit Index) corrige por el número de grados de libertad del siguiente modo:

$$CFI = \left[\frac{(\chi_{\text{indep}}^2 - gl_{\text{indep}}) - (\chi_{\text{teorico}}^2 - gl_{\text{teorico}})}{(\chi_{\text{indep}}^2 - gl_{\text{indep}})} \right]$$

AIC

Este índice pertenece a un nuevo grupo llamado índices de grado de parsimonia ya que toman en cuenta no sólo la bondad de ajuste del estadístico sino también el número de parámetros a estimar. El *AIC* por sus siglas en inglés, Akaike Information Criterion, se define por:

$$AIC = \chi_{\text{ind}}^2 - 2gl_{\text{ind}}$$

Su valor debe ser lo suficientemente pequeño, sin embargo, como el índice no varía en un intervalo $(0, 1)$, se entiende por suficientemente pequeño cuando se compara con otros modelos teóricos, es decir, este número nos ayuda a comparar modelos entre sí. Por tanto, se recomienda que el AIC se acompañe del modelo de independencia, el cual es la base de cualquier modelo teórico. Entonces, cuanto mayor sea la diferencia entre el modelo teórico y el de independencia, el modelo teórico es mejor.

Los índices de ajuste absoluto miden directamente el ajuste del modelo, es decir, miden en qué grado el modelo estimado predice la matriz de varianza covarianza observada. Estos índices normalmente se encuentran entre 0 y 1.

Índice GFI

El índice GFI, por sus siglas en inglés Goodness of Fit Index, es una razón entre los elementos ponderados de la matriz de covarianzas poblacional estimada y los elementos ponderados de la matriz de covarianzas muestral estimada directamente. Su expresión es la siguiente cuando el ML es utilizado:

$$\text{GFI} = 1 - \frac{\text{tr} \left[\left(\widehat{\Sigma}^{-1} \mathbf{S} - \mathbf{I} \right)^2 \right]}{\text{tr} \left[\left(\widehat{\Sigma}^{-1} \mathbf{S} \right)^2 \right]}$$

donde \mathbf{I} es la matriz identidad.

Índice AGFI

El índice AGFI por sus siglas en inglés Adjusted Goodness of Fit Index, es una corrección del anterior que se hace en función del número de variables (q) y los grados de libertad(df) del modelo.

$$\text{AGFI} = 1 - \left[\frac{q(q+1)}{2\text{df}} \right] [1 - \text{GFI}]$$

Índice RMR

Este índice está basado en los residuos, es decir, los promedios en las diferencias de las varianzas y covarianzas muestrales y las estimadas que se derivan del modelo. El índice está definido por:

$$\text{RMR} = \sqrt{\frac{\sum_{i=1}^q \sum_{j=1}^q (s_{ij} - \hat{\sigma}_{ij})^2}{q(q+1)/2}}$$

Dado que los residuos están afectados por la escala en que se miden las variables se recomienda usar el SRMR, el cual es la estandarización del RMR y toma valores en el intervalo (0, 1) y lo recomendable es que tome valores inferiores a 0.05.

RMSEA

RMSEA (Root Mean Square Error Approximation) por sus siglas en inglés, es un índice que estima la escasez en el modelo ajustado comparado a una perfecta saturación del modelo

(68); éste se puede interpretar como el error de aproximación medio por grado de libertad. El RMSEA es una medida no central que depende del tamaño de la muestra y los grados de libertad del modelo estimado (69, 70) y se calcula como la diferencia entre estos. Si la diferencia es menor a cero, el RMSEA se toma como cero. La expresión del RMSEA es:

$$RMSEA = \sqrt{\frac{\chi_{\text{teo}}^2 - \text{gl}_{\text{teo}}}{(N - 1) \text{gl}_{\text{teo}}}}$$

Si el indicador toma valores menores a 0.08 indica que el modelo tiene un buen ajuste, lo deseable es que el indicador se acerque a 0 (Cuadro 2.4).

Cuadro 2.4: Interpretación del RMSEA (71).

Valores de RMSEA	Conclusión
$0 < RMSEA < 0.05$	Muy buen ajuste
$0.05 < RMSEA < 0.08$	Buen ajuste
$0.08 < RMSEA < 0.1$	Raramente aceptable
$0.1 < RMSEA$	No aceptable

El investigador debe hacer una revisión de todos los indicadores ya que no hay una regla en que se mencione aquellos estadísticos en los que el investigador debe enfocarse para determinar si el modelo tiene un buen ajuste. Algunos autores mencionan que los indicadores que más se emplean son: GFI, SRMR, AGFI, CFI y RMSEA (72).

Resumen de los estadísticos de bondad de ajuste

El Cuadro 2.5 muestra un resumen de los estadísticos de bondad de ajuste, así como el valor de referencia para cuando se tiene un buen ajuste del modelo.

Cuadro 2.5: Estadísticos de bondad de ajuste y criterios de referencia (73).

Estadístico	Abreviatura	Referencia
Índice de ajuste normalizado	NFI	>0.90
Índice de ajuste comparativo	CFI	>0.90
Índice de bondad de ajuste	GFI	>0.90
Índice ajustado de bondad de ajuste	AGFI	>0.90
Raíz del residuo cuadrático medio	RMR	<0.08
Raíz cuadrada media residual estandarizada	SRMR	<0.08
Raíz del error cuadrático medio de aproximación	RMSEA	<0.08

2.4. Medidas de concordancia

Todo proceso de medición debe asegurar que las medidas generadas por distintos instrumentos o métodos concuerden entre si o concuerden con la verdadera medida o de referencia. Sin embargo, en la práctica es poco viable considerar una medida como verdadera pues, por un lado, se acepta que la medida considere un error tolerable por el investigador y, por el otro la medida verdadera no siempre se conoce o está disponible (74).

Cuando las medidas son continuas, se definen varios conceptos usados en diferentes áreas de estudio para referirse a la concordancia entre medidas. Conceptos como exactitud y precisión, repetibilidad y reproducibilidad, validez y confiabilidad, dependibilidad, generabilidad y acuerdo están estrechamente relacionados con la idea de cuantificar la “cercanía entre las medidas” considerando o no una medida de referencia (75).

Existen tres tipos de estudios para evaluar el acuerdo entre medidas, el primero es llamado estudio de comparación, el segundo el estudio de calibración y el tercero el estudio de conversión (76).

El estudio de comparación consiste en evaluar si el nuevo método genera medidas concordantes con el método de referencia, a menudo llamado Gold Standard, de modo que el nuevo método pueda reemplazar al antiguo. El estudio de calibración compara un método aproximado con uno que es conocido, preciso y exacto cuyo error de medición es marginal. El estudio de conversión compara métodos que miden la característica aproximada en diferentes

unidades de medida.

Para cualquier estudio de comparación, en la práctica, se han usado diferentes pruebas para evaluar el grado de acuerdo entre medidas. Barnhart y colaboradores (75) los resume en: métodos descriptivos, por ejemplo la dispersión de los pares de datos alrededor de la línea de 45 grados; índices no escalados, por ejemplo los límites de acuerdo propuesto por Bland y Altman (77) y los índices escalados, que toman valores entre -1 y 1 , por ejemplo el Coeficiente de Correlación Intra-clase (CCI) y el Coeficiente de Correlación de Concordancia (CCC) propuesto por Lin (78).

2.4.1. Coeficiente de correlación intraclase

Como enfoque alternativo, Fisher en 1958 desarrolló una metodología basada en el análisis de varianza (ANOVA), que dio lugar al CCI, que es un criterio de confiabilidad que proporciona la proporción de varianza atribuible a las diferencias entre métodos. El CCI es considerado como un índice de confiabilidad universal y ampliamente aceptado.

Suponga que cada uno de una muestra aleatoria de sujetos $(1, \dots, N)$ es calificado por un conjunto diferente de evaluadores $(1, \dots, R_i, i = 1, \dots, N)$. Primero se considera el caso de un número constante de calificaciones $(R_i = R, i = 1, \dots, N)$ para cada elemento. Cuando cada elemento es calificado por un conjunto diferente de evaluadores de R , seleccionados al azar de una población más grande de evaluadores (Ecuación 2.20), el efecto debido a los evaluadores, a la interacción entre los evaluadores y los elementos y al error aleatorio no se puede estimar por separado. Por lo tanto, solo el acuerdo absoluto es medible. Si $Y_{i,r}$ es la medida del evaluador r ($r = 1, \dots, R$) en el sujeto i ($i = 1, \dots, N$), se asume el siguiente modelo lineal,

$$Y_{i,r} = \mu + B_i + \epsilon_{i,r} \quad (i = 1, \dots, N; r = 1, \dots, R) \quad (2.20)$$

donde μ es la media poblacional de las mediciones, B_i es el efecto aleatorio que reflejan la variación entre los sujetos y $\epsilon_{i,r}$ es el error de medición que refleja la variación entre las mediciones dentro de un sujeto y los B_i son independientes de $\epsilon_{i,r}$

$$B_i \sim N(0, \sigma_B^2) \quad \text{y} \quad \epsilon_{i,r} \sim N(0, \sigma_\epsilon^2)$$

El CCI es dado por la expresión:

$$CCI = \frac{\text{cov}(Y_{i,r}, Y_{i,s})}{\sqrt{\text{var}(Y_{i,r}) \text{var}(Y_{i,s})}} = \frac{\sigma_B^2}{\sigma_\epsilon^2 + \sigma_B^2}$$

Cuadro 2.6: Modelo ANOVA de efectos aleatorios de una sola vía (Ecuación 2.20.)

Variabilidad	Suma de cuadrados	Grados de libertad	Cuadrados medios	E(MS)
Entre sujetos	SCE	$N - 1$	CME	$R\sigma_B^2 + \sigma_\epsilon^2$
Dentro de sujetos	SCD	$N(R - 1)$	CMD	σ_ϵ^2
Total	SCT	$NR - 1$		

Se puede ver en el Cuadro 2.6 que CMD es un estimador insesgado de σ_ϵ^2 y $(CMD - CME)/R$ es un estimador insesgado de σ_B^2 .

Un estimador del CCI es:

$$\widehat{CCI} = \frac{CME - CMD}{CME + (R - 1)CMD} \quad (2.21)$$

donde

$$SCE = R \sum_{i=1}^N (\bar{y}_{i,\cdot} - \bar{y}_{\cdot,\cdot})^2$$

$$SCD = \sum_{i=1}^N \sum_{r=1}^R (y_{i,r} - \bar{y}_{i,\cdot})^2$$

$$SCT = \sum_{i=1}^N \sum_{r=1}^R (y_{i,r} - \bar{y}_{\cdot,\cdot})^2$$

$y_{i,r}$ el valor observado de la variable aleatoria $Y_{i,r}$ ($i = 1, \dots, N; r = 1, \dots, R$), $\bar{y}_{i,\cdot}$ el valor medio sobre los evaluadores y $\bar{y}_{\cdot,\cdot}$ la media general, es decir:

$$\bar{y}_{i,\cdot} = \frac{1}{R} \sum_{r=1}^R y_{i,r}, \text{ y } \bar{y}_{\cdot,\cdot} = \frac{1}{NR} \sum_{i=1}^N \sum_{r=1}^R y_{i,r}$$

Esta estimación es consistente pero sesgada, ya que la esperanza de un cociente no es igual al cociente de las esperanzas.

Intervalo de confianza para CCI

Algunos de los métodos para obtener intervalos de confianza bilaterales para el CCI en el modelo de efectos aleatorios de una sola vía son los siguientes

Método basado en el error estándar de $\hat{\rho}$ (Asintótico simple)

El enfoque más simple para un solo ρ es aplicar el teorema del límite central junto con el teorema de Slutsky. Swiger y colaboradores (79) derivaron una aproximación a la varianza de muestra grande de \widehat{CCC} como

$$\widehat{V}(\widehat{CCC}) = \frac{2(NR - 1)(1 - \hat{\rho})^2(1 + (R - 1)\widehat{CCC})^2}{R^2(R - 1)N(N - 1)}$$

donde N es el número de sujetos en el estudio, p es el número de mediciones de cada sujeto (generalmente denominado tamaño de la clase). Por lo tanto, el intervalo de confianza de $(1 - \alpha)100\%$ para CCC puede calcularse como

$$\left(\widehat{CCC} - Z_{\alpha/2} \sqrt{\widehat{V}(\widehat{CCC})}, \widehat{\rho} + Z_{\alpha/2} \sqrt{\widehat{V}(\widehat{CCC})} \right)$$

Método basado en la transformación de Fisher

Fisher demostró que

$$z(\widehat{CCC}) = \frac{1}{2} \ln \frac{1 + (R - 1)\widehat{CCC}}{1 - \widehat{CCC}}$$

se distribuye asintóticamente normal

$$N \left(\frac{1}{2} \ln \frac{1 + (R - 1)CCC}{1 - \rho}, \frac{R}{2(R - 1)(N - 2)} \right)$$

Los límites de confianza aproximados de $(1 - \alpha)100\%$ para CCC están dados por

$$\left(I \left(z - Z_{\alpha/2} \sqrt{\widehat{V}(z)} \right), I \left(z + Z_{\alpha/2} \sqrt{\widehat{V}(z)} \right) \right)$$

donde

$$I(z) = \frac{e^{2z} - 1}{e^{2z} - 1 + R}$$

es el inverso de la transformación de Fisher.

Una ventaja de este método es que se puede ampliar fácilmente para construir intervalos de confianza para un CCI en el caso de diseños de muestreo más complejos.

Transformación Arctanh

Se ha sugerido que la transformación tangente hiperbólica inversa (a veces también denominada transformación arctanh) puede funcionar mejor que la transformación Z de Fisher. Los límites resultantes se basan en la distribución muestral de

$$z(\widehat{CCC}) = \frac{1}{2} \ln \frac{1 + \widehat{CCC}}{1 - \widehat{CCC}}$$

que puede tomarse como

$$N \left(\frac{1}{2} \ln \frac{1 + CCC}{1 - CCC}, \frac{2(NR - 1)(1 + (R - 1)CCC)^2}{R^2(R - 1)N(N - 1)(1 + CCC)^2} \right)$$

La inversión de la transformación arctanh está dada por

$$I(z) = \frac{e^{2z} - 1}{e^{2z} + 1}$$

los límites de confianza aproximados de $(1 - \alpha)100\%$ para ρ están dados por

$$\left(I \left(z - Z_{\alpha/2} \sqrt{V(\widehat{CCC})} \right), I \left(z + Z_{\alpha/2} \sqrt{V(\widehat{CCC})} \right) \right)$$

Método basado en las aproximaciones a la distribución F

También se puede construir un intervalo de confianza exacto para CCI basado en la distribución F. Este intervalo de confianza viene dado por

$$l = \frac{\left[R\widehat{CCC} + (1 - F_1 - \alpha/2)(1 - \widehat{CCC}) \right]}{\left[R - (R - 1)(1 - F_1 - \alpha/2)(1 - \widehat{CCC}) \right]}, u = \frac{\left[R\widehat{CCC} + (1 - F_{\alpha/2})(1 - \widehat{CCC}) \right]}{\left[R - (R - 1)(1 - F_{\alpha/2})(1 - \widehat{CCC}) \right]},$$

donde F_q es el q ésimo cuantil de la distribución F con grados de libertad de $N - 1$ y $N(R - 1)$.

2.4.2. Coeficiente de correlación de concordancia

Existe la necesidad de derivar un coeficiente que refleje el acuerdo entre dos evaluadores. El recurso al coeficiente de correlación de Pearson, la prueba t pareada, el análisis de mínimos cuadrados de la pendiente y el intersepto o al coeficiente de variación falla en algunos casos. Esto llevó a Lin en 1989 (78) a desarrollar el coeficiente de correlación de concordancia (CCC), un índice de reproducibilidad que mide la correlación entre dos lecturas que caen en la línea de 45° que pasa por el origen.

Suponga que la distribución conjunta de Y_1 y Y_2 es normal bivariada con media $(\mu_1, \mu_2)'$ y matriz de varianzas y covarianzas

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$$

El grado de concordancia entre Y_1 y Y_2 se puede caracterizar por el valor esperado de la diferencia al cuadrado

$$\begin{aligned} E(Y_1 - Y_2)^2 &= (\mu_1 - \mu_2)^2 + (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) \\ &= (\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2 + 2(1 - \rho)\sigma_1\sigma_2 \end{aligned}$$

donde $\rho = \text{corr}(Y_1, Y_2) = \sigma_{12}/\sigma_1\sigma_2$. Lin propuso aplicar una transformación para escalar el índice de acuerdo entre -1 y 1 , lo que lleva al coeficiente de correlación de concordancia (78)

$$CCC = 1 - \frac{E(Y_1 - Y_2)^2}{(\mu_1 - \mu_2)^2 + \sigma_1^2 + \sigma_2^2} = \frac{2\rho\sigma_1\sigma_2}{(\mu_1 - \mu_2)^2 + \sigma_1^2 + \sigma_2^2} = \rho C_b$$

donde $C_b = [(\nu + 1/\nu + u^2)/2]^{-1}$ con $\nu = \sigma_1/\sigma_2$ que representa el cambio de escala y $u = (\mu_1 - \mu_2)/\sqrt{\sigma_1\sigma_2}$ el cambio de ubicación en relación con la escala. C_b ($0 < C_b \leq 1$) es un factor de corrección de sesgo que mide qué tan lejos se desvía el mejor ajuste de la línea 45° (medida de precisión) (78). No se produce ninguna desviación cuando $C_b = 1$. El coeficiente de correlación de Pearson ρ mide cuánto se desvían las observaciones de la línea de mejor ajuste (medida de precisión).

El CCC posee las siguientes propiedades:

- $-1 \leq -|\rho| \leq CCC \leq |\rho| \leq 1$
- $CCC = 0$ si y solo si $\rho = 0$;
- $CCC = \rho$ si y solo si $\sigma_1 = \sigma_2$ y $\mu_1 = \mu_2$
- $CCC = \pm 1$ si y solo si cada par de medidas está en perfecto acuerdo o en perfecto acuerdo inverso.

Para una muestra de tamaño n , sean $(y_{i,1}, y_{i,2})$, si $\bar{y}_{.r}$ denota la media estimada y s_r^2 la varianza muestral de las mediciones realizadas por el evaluador r ($r = 1, 2$),

$$\bar{y}_{.r} = \frac{1}{n} \sum_{i=1}^n y_{i,r} \quad \text{y} \quad s_r^2 = \frac{1}{n} \sum_{i=1}^n (y_{i,r} - \bar{y}_{.r})^2$$

y $\hat{\rho}$ es el coeficiente de correlación muestral.

$$\hat{\rho} = \frac{\sum_{i=1}^n (y_{i,1} - \bar{y}_{.1})(y_{i,2} - \bar{y}_{.2}) / n}{s_1 s_2}$$

el CCC is estimado por

$$\widehat{CCC} = \frac{2\hat{\rho}s_1s_2}{(\bar{y}_{.1} - \bar{y}_{.2})^2 + s_1^2 + s_2^2}$$

Para lograr una mejor aproximación con la distribución normal, se propone transformar \widehat{CCC} mediante la transformación z de Fisher para obtener (78)

$$\hat{\lambda} = \tanh^{-1}(\widehat{CCC}) = \frac{1}{2} \ln \left(\frac{1 + \widehat{CCC}}{1 - \widehat{CCC}} \right)$$

Esta cantidad tiene una distribución asintóticamente normal con media

$$\lambda = \tanh^{-1}(CCC) = \frac{1}{2} \ln \left(\frac{1 + CCC}{1 - CCC} \right)$$

y

$$Var(CCC, u, n)^2 = \frac{1}{n-2} \left\{ \frac{(1-\rho^2)CCC^2}{(1-CCC^2)\rho^2} + \frac{2CCC^3(1-CCC)v^2}{\rho(1-CCC^2)^2} - \frac{CCC^4v^4}{2\rho^2(1-CCC^2)^2} \right\}$$

Esta varianza se estima utilizando

$$S_{\hat{\lambda}}^2 = \frac{1}{n-2} \left\{ \frac{(1-r^2)\widehat{CCC}^2}{(1-\widehat{CCC}^2)r^2} + \frac{2\widehat{CCC}^3(1-\widehat{CCC})v^2}{r(1-\widehat{CCC}^2)^2} - \frac{\widehat{CCC}^4v^4}{2r^2(1-\widehat{CCC}^2)^2} \right\}$$

donde

$$v = \frac{\bar{y}_{.1} - \bar{y}_{.2}}{\sqrt{s_1 s_2}}$$

2.4.3. Gráficos de Bland y Altman

Sea Y una cantidad asociada con cada sujeto de una cierta población en estudio. Sean Y_1 y Y_2 las cantidades correspondientes medidas por dos evaluadores o instrumentos distintos. En teoría, la concordancia perfecta entre los evaluadores ocurre cuando $Y_1 = Y_2$. Por lo tanto, dada una muestra, la concordancia entre los dos evaluadores o instrumentos se ve mejor trazando las observaciones emparejadas respecto a la línea de 45° ($Y_1 = Y_2$). Si los dos evaluadores están en perfecto acuerdo, todas las observaciones caerán en la línea de 45° . Por el contrario, el desacuerdo entre los dos evaluadores puede tomar diferentes formas: (i)

un sesgo constante ($Y_2 = a + Y_1, a \in \mathbb{R}$), (ii) un sesgo proporcional ($Y_2 = bY_1, b \in \mathbb{R}$), o (iii) ambos tipos de sesgos ($Y_2 = a + bY_1, a, b \in \mathbb{R}$) (80). Esto llevó a Bland y Altman (77) a sugerir graficar la diferencia de las medidas ($Y_2 - Y_1$) contra su media $(Y_1 + Y_2) / 2$, llamado “diagrama de Bland y Altman”. En caso de sesgo constante, $Y_2 - Y_1 = a$ ($a \neq 0$), las observaciones tenderán a situarse alrededor de una línea horizontal; en caso de sesgo proporcional, $Y_2 - Y_1 = (b - 1)Y_1$ ($b \neq 0$), los puntos se dispersarán alrededor de una línea creciente ($b > 1$) o decreciente ($b < 1$) que pasa por el origen; cuando ambos sesgos están presentes, la línea creciente o decreciente no pasará por el origen.

Los dos métodos descritos son esencialmente gráficos, aunque mediante análisis de regresión es posible estimar los factores de sesgo constante y proporcional.

Ejemplo: Shrout y Fleiss (80) consideraron el siguiente ejemplo hipotético, donde 4 evaluadores midieron 6 ítems en una escala de 10 puntos (Cuadro 2.7). Considere solo las medidas de los evaluadores 1 y 4.

Como se ve en el gráfico dispersión de los pares de datos alrededor de la línea de 45° y en la gráfica de Bland y Altman (Figura 2.1), la concordancia perfecta ocurre solo para un elemento (elemento 3). Además, parece que el evaluador 1 da, en general, valores más altos que el evaluador 4. La diferencia media ($\pm DE$) es 1.0 ± 1.67 ($IC95\% : [-2.28; 4.28]$).

Cuadro 2.7: Ejemplo de Shrout y Fleiss (1979).

Item	Evaluador				$Y_1 - Y_4$	$(Y_1 + Y_4) / 2$
	1	2	3	4		
1	9	2	5	8	1	8.5
2	6	1	3	2	4	4
3	8	4	6	8	0	8
4	7	1	2	6	1	6.5
5	10	5	6	9	1	9.5
6	6	2	4	7	-1	6.5

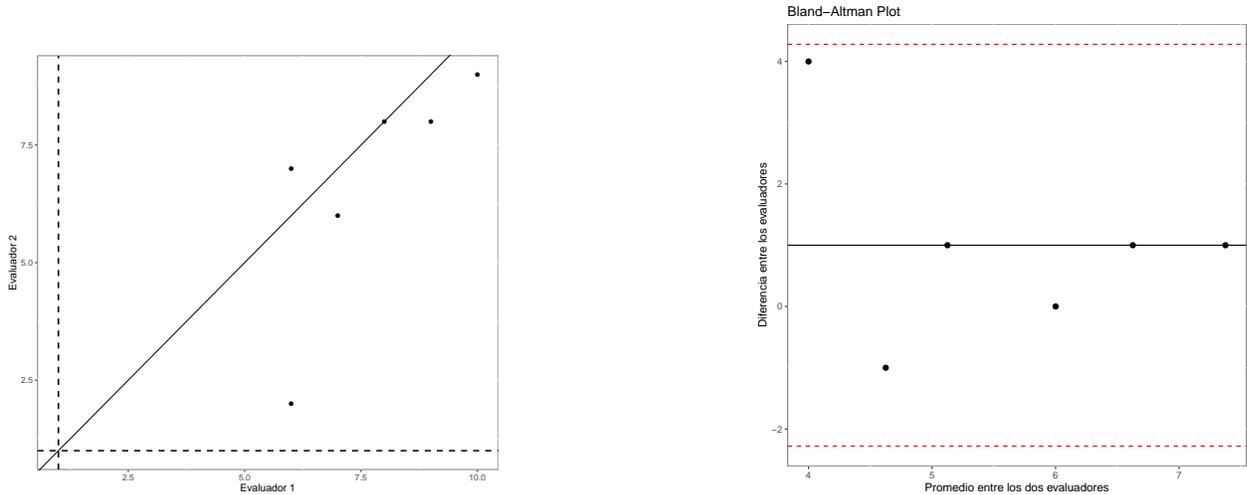


Figura 2.1: Gráfico dispersión de los pares de datos alrededor de la línea de 45° (izquierda) y gráfico de Bland y Altman (derecha) para las medidas de los evaluadores 1 y 4.

2.5. Curvas ROC

2.5.1. Definiciones previas

La prueba de diagnóstico más simple es aquella en la que el resultado de la prueba es dicotómica, es decir, es una prueba que tiene solo dos resultados posibles, positivo o negativo para la enfermedad (81).

Definiendo D como la variable aleatoria que modela el verdadero estado de la enfermedad, $D = 1$ indica que el paciente tiene la enfermedad y $D = 0$ el paciente no tiene dicha enfermedad.

$$D \sim Be(p)$$

Una prueba puede fallar en su tarea de detectar la enfermedad de dos maneras diferentes, clasificando incorrectamente a un paciente sano (un falso positivo, denotado por FP) o declarando que un paciente está sano cuando este enfermo (un falso negativo, indicado por FN). Por el contrario, la prueba puede clasificar correctamente a un paciente sano (un verdadero negativo, denotado por VN) o un paciente enfermo (un verdadero positivo, denotado por VP). La prevalencia de la enfermedad es la probabilidad de que un individuo tenga la enfermedad y normalmente se denota por π ; es decir, $\pi = Pr(D = 1)$.

2.5.2. Sensibilidad y especificidad

La validez de una prueba diagnóstica en la práctica clínica generalmente se mide por las probabilidades condicionales de clasificar correctamente a un paciente, la sensibilidad y especificidad las medidas de precisión más conocidas.

La sensibilidad (Se) es la probabilidad de clasificar correctamente a un paciente enfermo, es decir, la probabilidad de que un paciente enfermo obtenga un resultado positivo en la prueba. La sensibilidad es, por tanto, la capacidad de la prueba para detectar la enfermedad cuando está presente, es decir, qué tan sensible es la prueba a la presencia de la enfermedad, también se conoce como fracción de verdaderos positivos (FVP) o tasa verdadera positiva (TPV). La sensibilidad solo se puede calcular a partir de aquellos individuos que tienen la enfermedad. Esto significa que la sensibilidad de una prueba no nos dice nada acerca de si algunas personas sin la enfermedad también darían positivo y, de ser así, en qué proporción. A partir de la matriz de confusión dado en el Cuadro 2.8, se puede estimar la sensibilidad como la proporción de personas enfermas que dieron positivo, es decir:

$$Se = \Pr(Y+ | D = 1) = \Pr(VP) = VP/(VP + FN)$$

La especificidad (Sp) o fracción negativa verdadera (FNV) es la probabilidad de clasificar correctamente a un paciente sano, es decir, la probabilidad de que para un sujeto sano se obtenga un resultado negativo en la prueba. En otras palabras, la especificidad se puede definir como la capacidad de la prueba para excluir la enfermedad en pacientes libres de enfermedad. La especificidad solo se puede calcular a partir de aquellas personas que no tienen la enfermedad, por lo que la especificidad no nos dice nada sobre si algunas personas con la enfermedad también tendrían un resultado negativo y, de ser así, en qué proporción. En el Cuadro 2.8, se puede estimar la especificidad como:

$$Sp = \Pr(Y- | D = 0) = \Pr(VN) = VN/(VN + FP)$$

Cuadro 2.8: Clasificación de los resultados de las pruebas por estado de enfermedad.

Resultado de la prueba (Y)	Estado verdadero de la enfermedad (D)		
	Enfermo ($D = 1$)	Sano ($D = 0$)	Total
Positivo ($Y+$)	VP	FP	VP + FP
Negativo ($Y-$)	FN	VN	FN + VN
Total	VP + FN	FP + VN	$N = (VP + FP + FN + VN)$

Tanto la sensibilidad como la especificidad son probabilidades de éxito en el diagnóstico de enfermedades. A partir de estas decisiones correctas, las probabilidades de los correspondientes errores de clasificación se definen de la siguiente manera:

- 1-sensibilidad ($1 - Se$), también llamada fracción de falsos negativos (FFN), es la probabilidad de clasificar incorrectamente a un paciente enfermo:

$$1 - Se = \Pr(Y- | D = 1) = \Pr(FN) = FN / (VP + FN)$$

- 1-especificidad ($1 - Sp$), también llamada fracción de falsos positivos (FFP), es la probabilidad de clasificar incorrectamente a un sujeto sano:

$$1 - Sp = \Pr(Y+ | D = 0) = \Pr(FP) = FP / (VN + FP)$$

La suma de la probabilidad de un verdadero positivo (sensibilidad) y la probabilidad de un falso negativo es igual a uno:

$$\Pr(Y+ | D = 1) + \Pr(Y- | D = 1) = 1$$

De manera similar, la suma de la probabilidad de un verdadero negativo (especificidad) y la probabilidad de un falso positivo es igual a uno:

$$\Pr(Y- | D = 0) + \Pr(Y+ | D = 0) = 1$$

En el contexto de pruebas de hipótesis, la hipótesis nula ($D = 0$) versus la hipótesis alternativa ($D = 1$), la sensibilidad es el poder estadístico β y 1-especificidad es el nivel de significancia α .

El índice de Youden (Youden, 1950) (indicado en la literatura por J o IY) es otra medida de la eficacia de una prueba de diagnóstico definida como (82)

$$J = Se + Sp - 1 = Se - (1 - Sp)$$

para maximizar la eficacia de la prueba. J toma valores entre 0 y 1 y es la diferencia máxima entre la probabilidad de clasificar correctamente a los sujetos enfermos o la tasa de verdaderos positivos (sensibilidad) y la probabilidad de clasificar erróneamente a los individuos sanos o la tasa de falsos positivos (1-especificidad).

Algunas de las propiedades de este índice son

- Cuando $J = 0$ corresponde a una prueba completamente ineficaz para diagnosticar la enfermedad.
- Si $J = 1$ indica una perfecta efectividad diagnóstica.

2.5.3. Pruebas continuas

Una prueba de diagnóstico continua toma valores en un rango. En este tipo de pruebas surge la siguiente pregunta: ¿cómo clasificar a un individuo como sano o enfermo a partir de los valores de Y ? Cuando el resultado de la prueba de diagnóstico Y es ordinal o continuo, la regla de clasificación es más compleja que en el caso de una prueba binaria. En las pruebas diagnósticas continuas es necesario seleccionar un punto de corte o valor de discriminación c , un valor determinado de la prueba, que define el límite entre sano y enfermo. En general, sin pérdida de generalidad, asumiremos que valores más altos de la prueba diagnóstica están asociados con enfermedad. Bajo este supuesto, tiene la siguiente regla de clasificación (Figura 2.2): si el valor de la prueba de diagnóstico Y es igual o mayor que el punto de corte c , el paciente se clasifica como enfermo (prueba positiva resultado $Y+$), mientras que si el valor de la prueba es menor que c , el paciente se clasifica como sano (resultado negativo de la prueba $Y-$).

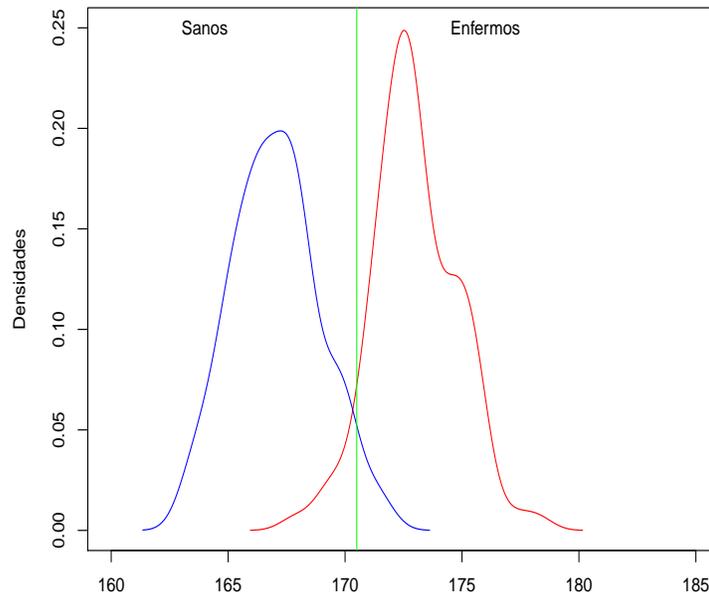


Figura 2.2: Clasificación de un individuo como sano o enfermo en función de un límite fijo c .

De manera análoga a una prueba de diagnóstico binaria, esta clasificación no está libre de errores y la prueba de diagnóstico continua puede fallar en su tarea de detectar enfermedades (Cuadro 2.9). Por lo tanto, antes de la aplicación rutinaria de una prueba de diagnóstico continua en la práctica, se deben cuantificar los errores de clasificación errónea.

La diferencia esencial ahora es que no tiene un solo par de valores de sensibilidad y especificidad que definen la precisión de la prueba, sino un conjunto de pares de valores correspondientes cada uno en un punto de corte diferente c .

$$Se(c) = FVP(c) = \Pr(Y+ | D = 1) = \Pr(Y \geq c | D = 1)$$

$$Sp(c) = FNV(c) = \Pr(Y- | D = 0) = \Pr(Y < c | D = 0)$$

Cuadro 2.9: Clasificación de los resultados de las pruebas continuas por estado de enfermedad.

Resultado de la prueba(Y)	Enfermo ($D = 1$)	Sano ($D = 0$)
Positivo ($Y \geq c$)	Verdadero positivo (TP(c))	Falso positivo (FP(c))
Negativa ($Y < c$)	Falso negativo (FN(c))	Verdadero negativo (VN(c))
Total	VP(c) + FN(c)	FP(c) + VN(c)

A partir de las medidas anteriores, puede definir las probabilidades correspondientes de clasificaciones incorrectas de diagnóstico.

$$1 - Se(c) = FFN(c) = \Pr(Y < c \mid D = 1)$$

$$1 - Sp(c) = FFP(c) = \Pr(Y \geq c \mid D = 0)$$

2.5.4. Curva Característica Operativa del Receptor

Una herramienta útil para evaluar la capacidad de diagnóstico de una prueba cuantitativa para todos los posibles valores de corte de una prueba de diagnóstico continua se llama curva de característica operativa del receptor (ROC) (83,84). La curva ROC se construye trazando las sensibilidades en el eje y y las correspondientes fracciones de falso positivo (1-sensibilidad) en el eje x :

$$ROC(.) = (1 - Sp(c), Se(c)), c \in (-\infty, \infty)$$

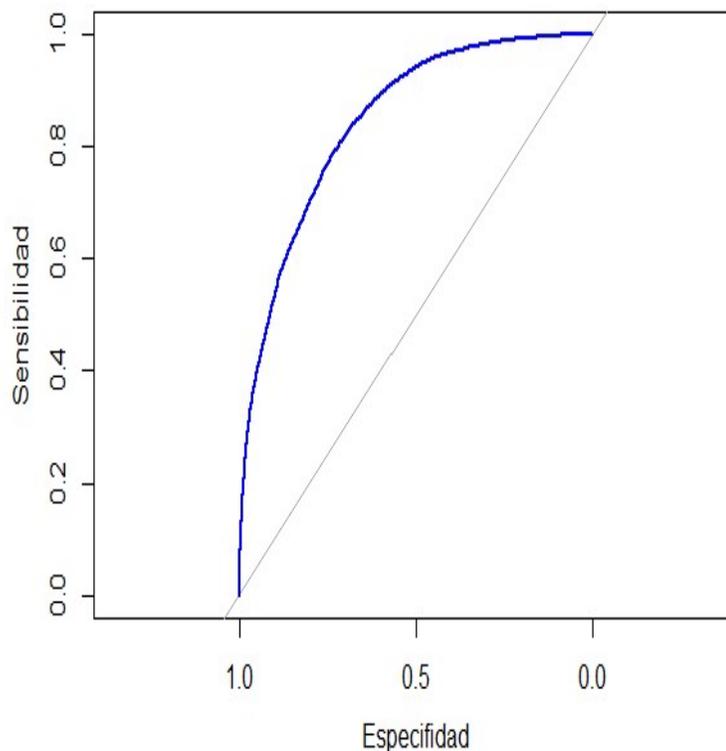


Figura 2.3: Curva ROC de una prueba diagnóstica continua.

La curva ROC se puede representar en términos de las CDF de la prueba de diagnóstico en ambas poblaciones (sanas y enfermas) de la siguiente manera:

$$ROC(t) = 1 - F_D(F_{\bar{D}}^{-1}(1 - t)), t \in (0, 1)$$

donde F_D and $F_{\bar{D}}$ son las funciones de distribución de la prueba de diagnóstico en las poblaciones (\bar{D}) enfermas y no enfermas, es decir,

$$F_D(y) = \Pr(Y \leq y \mid D = 1)$$

$$F_{\bar{D}}(y) = \Pr(Y \leq y \mid D = 0)$$

La curva ROC también tiene varias propiedades interesantes:

- La curva ROC es una función monótona creciente definida en el cuadrado unitario $[0; 1] \times [0; 1]$.
- La curva ROC une los puntos $(0, 0)$ y $(1, 1)$.
- La curva ROC es invariante a las transformaciones estrictamente crecientes del marcador Y .

Desde un punto de vista práctico, la curva ROC tiene una interpretación muy simple e interesante:

- Es una medida global de la capacidad discriminatoria (precisión) de una prueba diagnóstica continua (u ordinal), independiente del punto de corte y de la prevalencia de la enfermedad, que cumple las dos condiciones marcadas por Swets (1988) como características indispensables de cualquier indicador de la precisión del diagnóstico.
- Representa la capacidad discriminativa de una prueba de diagnóstico a lo largo de todos los puntos de corte posibles en la escala o rango del marcador de diagnóstico.
- Sirve de guía a la hora de seleccionar el punto de corte óptimo en las pruebas diagnósticas.
- Proporciona un método para comparar diferentes pruebas de diagnóstico.

El área bajo la curva es el estadístico por excelencia para medir la capacidad discriminante de una prueba. También para comparar pruebas entre sí y determinar cual es la más eficaz.

Los índices numéricos de las curvas ROC se utilizan a menudo para resumir la precisión de una prueba de diagnóstico.

Definición: Sea $ROC(t)$ la función asociada a la curva ROC. Se define el área bajo la curva bajo la curva ROC (area under curve, AUC) como:

$$AUC = \int_0^1 ROC(t)dt$$

Su rango de valores va desde 0.5, siendo este valor el correspondiente a una prueba sin capacidad discriminante, hasta 1, que es cuando los dos grupos están perfectamente diferenciados por la prueba. Cuanto mayor sea el AUC mejor será la prueba (84).

Cuadro 2.10: Interpretación de Swets para valores del AUC.

Baja exactitud	[0.5, 0.7)
Útiles para algunos propósitos	[0.7, 0.9)
Exactitud alta	[0.9, 1]

Bamber en 1975 (85) mostró que el AUC puede interpretarse como la probabilidad de que un individuo enfermo elegido al azar tenga un valor de prueba mayor que el de un individuo sano elegido al azar:

$$AUC = \Pr(Y_1 > Y_0)$$

donde Y_1 y Y_0 representan valores de la prueba de diagnóstico de los grupos sanos y enfermos, respectivamente.

2.5.5. Cálculo del area bajo la curva

De acuerdo a la definición dada de AUC, una primera forma de calcularla sería representando la curva ROC, o una estimación de ella, y obtener el porcentaje de área del cuadrado $[0, 1] \times [0, 1]$ que encierra bajo ella. Para profundizar más en el tópico de los métodos de estimación del AUC bajo distintos enfoques ver (86).

Considere dos muestras independientes de observaciones i.i.d. (independientes e idénticamente distribuidas), $\{Y_{0k_0}\}_{k_0=1}^{n_0}$ y $\{Y_{1k_1}\}_{k_1=1}^{n_1}$, tomado de las poblaciones sanas y enfermas, Y_0 y Y_1 , respectivamente, con tamaños de muestra n_0 y n_1 .

Modelo binormal

El primer método paramétrico se basa en el supuesto de que la distribución del marcador o prueba tanto en poblaciones sanas como enfermas sigue una distribución normal (87, 88).

Asumiendo que Y_i sigue una distribución normal con media μ_i y desviación estándar σ_i , para $i = 0, 1$, para cada punto de corte c , se deduce que

$$\begin{aligned} FFP(c) &= 1 - Sp(c) = P(Y_0 \geq c) = P\left(Z \geq \frac{c - \mu_0}{\sigma_0}\right) \\ Se(c) &= P(Y_1 \geq c) = P\left(Z \geq \frac{c - \mu_1}{\sigma_1}\right) \end{aligned} \quad (2.22)$$

donde Z denota una variable aleatoria que sigue la distribución normal estándar.

Bajo esta condición, existe una relación lineal entre los cuantiles de la distribución normal estándar definida por los valores de sensibilidad y 1-especificidad, es decir, se tiene que:

$$\Phi^{-1}(Se(c)) = a + b\Phi^{-1}(1 - Sp(c))$$

donde $a = \frac{\mu_1 - \mu_0}{\sigma_1}$ es el intercepto, $b = \frac{\sigma_0}{\sigma_1}$ es la pendiente, y Φ denota la CDF normal estándar. En consecuencia, se deduce que la curva ROC se puede expresar como:

$$ROC(t) = \Phi(a + b\Phi^{-1}(t))$$

donde $t = 1 - Sp(c)$ es la *FPF*.

En la literatura, se han propuesto varios métodos de inferencia para AUC, principalmente basados en el supuesto de binormalidad y la técnica de bootstrap.

Bajo el modelo binormal, el AUC se puede reescribir también en términos de a, b y Φ de la siguiente manera:

$$AUC = \Phi \left(\frac{a}{\sqrt{1 + b^2}} \right)$$

Métodos no paramétricos

Para resolver el problema de especificar una distribución adecuada, se han propuesto en la literatura diferentes métodos no paramétricos. Los métodos no paramétricos se caracterizan por no asumir ninguna distribución conocida para las pruebas diagnósticas tanto en poblaciones sanas como enfermas.

Método empírico

El método empírico es uno de los métodos no paramétricos más comunes y sencillos propuestos en la literatura (89,90). Estima la función de distribución de la prueba de diagnóstico mediante la función de distribución empírica de la muestra; es decir, el estimador empírico de la curva ROC simplemente aplica la definición de la curva ROC a los datos observados. Entonces, para cada valor de corte posible, c , la *FVP(Se)* empírica y la *FFP* empírica se estiman por $\widehat{Se}(c) =$ proporción de individuos enfermos en la muestra con $Y \geq c$ y $\widehat{FFP}(c) =$ proporción de individuos sanos con $Y \geq c$, es decir,

$$\begin{aligned} \widehat{Se}(c) &= \widehat{FVP}(c) = \frac{1}{n_1} \sum_{i=1}^{n_1} I[Y_{1i} \geq c] = \frac{\#[Y_{1i} \geq c]}{n_1} \\ 1 - \widehat{Sp}(c) &= \widehat{FFP}(c) = \frac{1}{n_0} \sum_{j=1}^{n_0} I[Y_{0j} \geq c] = \frac{\#[Y_{0j} \geq c]}{n_0} \end{aligned}$$

Por lo tanto, la curva ROC empírica es una gráfica de los pares $(\widehat{FPF}(c), \widehat{TPF}(c))$ para todo $-\infty < c < \infty$.

De manera equivalente, se puede escribir la curva ROC empírica de la siguiente manera:

$$\widehat{ROC}_{emp}(t) = 1 - \widehat{F}_1 \left(\widehat{F}_0^{-1}(1 - t) \right)$$

donde \widehat{F}_0 y \widehat{F}_1 son las funciones de distribución empírica obtenidas de las dos muestras de poblaciones sanas y enfermas, respectivamente:

$$\widehat{F}_0(c) = \frac{1}{n_0} \sum_{i=1}^{n_0} I [Y_{0i} \leq c]$$

$$\widehat{F}_1(c) = \frac{1}{n_1} \sum_{j=1}^{n_1} I [Y_{1j} \leq c]$$

El método empírico presenta las siguientes propiedades:

El estimador empírico del AUC viene dado por:

$$\widehat{AUC} = \int_0^1 \widehat{ROC}_{emp}(t) dt$$

Si no hay empates en los datos, la estimación empírica del AUC es equivalente a la prueba de Mann-Whitney:

$$\widehat{AUC} = \frac{1}{n_0 n_1} \sum_{i=1}^{n_0} \sum_{j=1}^{n_1} I [Y_{1j} \geq Y_{0i}]$$

Método de la función Kernel

El suavizado de kernel es un enfoque suave basado en la estimación de kernel (91) de las funciones de distribución F_1 y F_0 , proporcionando así una estimación de estas funciones. F_1 y F_0 se estiman, respectivamente

$$\widehat{F}_{1,h_1}(y) = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbb{K} \left(\frac{y - Y_{1i}}{h_1} \right)$$

$$\widehat{F}_{0,h_0}(y) = \frac{1}{n_0} \sum_{j=1}^{n_0} \mathbb{K} \left(\frac{y - Y_{0j}}{h_0} \right)$$

donde $\mathbb{K}(u) = \int_{-\infty}^u K(v) dv$, K denota una función Kernel (una densidad simétrica) y h_i , $i = 0, 1$ es el parámetro de suavizado para cada población.

La estimación suave del núcleo de la curva ROC se obtiene simplemente teniendo en cuenta las estimaciones anteriores, es decir, trazando los pares correspondientes de puntos de coordenadas, es decir

$$\widehat{ROC}(t) = 1 - \widehat{F}_{1,h_1} \left(\widehat{F}_{0,h_0}^{-1}(1 - t) \right), t \in (0, 1)$$

Otro método para estimar el AUC de la curva ROC es mediante el suavizamiento por splines.

2.5.6. Contraste para la aleatoriedad de la prueba

Para hacer inferencia sobre el AUC, es necesario estudiar el comportamiento asintótico de la estimación empírica del AUC. Por un lado, la construcción de intervalos de confianza se puede realizar utilizando los métodos no paramétricos propuestos por Hanley y McNeil (92); DeLong *et.al* (93). Por otro lado, es de gran interés probar si una prueba específica Y puede discriminar entre poblaciones sanas y enfermas, y esto se puede hacer con una prueba de hipótesis basada en AUC.

$$\begin{cases} H_0 : AUC = 0.5 \\ H_1 : AUC \neq 0,5 \end{cases}$$

lo que equivale a probar si existe algún grado de separación entre las dos distribuciones. Para tamaños de muestra grandes, la prueba rechazará H_0 si

$$\frac{\widehat{AUC} - 0.5}{\sqrt{\frac{n_0+n_1+1}{12n_0n_1}}} > z_{1-\alpha}$$

Mediante la equivalencia establecida entre el estimador empírico del AUC y el estadístico de Mann-Whitney, esta prueba de hipótesis también se puede resolver para tamaños de muestra pequeños mediante la distribución exacta del estadístico de Mann-Whitney.

También puede ser de interés determinar si tiene un AUC satisfactorio. En este caso, denotando por θ_0 el valor mínimo aceptable para AUC, la prueba de hipótesis es:

$$H_0 : AUC = \theta_0 \text{ versus } H_1 : AUC > \theta_0$$

Por tanto, dado un nivel de significación, α , H_0 será rechazada si

$$\frac{\widehat{AUC} - \theta_0}{\widehat{\text{var}}(\widehat{AUC})} > z_{1-\alpha}$$

donde \widehat{AUC} es el estimador empírico del AUC y $\widehat{\text{var}}(\widehat{AUC})$ es el estimador de $\text{var}(\widehat{AUC})$.

El estimador empírico del índice de Youden viene dado por:

$$\widehat{J} = \max_t \{ \widehat{ROC}(t) - t \}$$

y es equivalente al conocido estadístico de Kolmogorov-Smirnov para la comparación de funciones de distribución:

$$\widehat{J} = \sup_c \{ \widehat{FVP}(c) - \widehat{FFP}(c) \} = \sup_c \left\{ \left(1 - \widehat{F}_1(c) \right) - \left(1 - \widehat{F}_0(c) \right) \right\} = \sup_c \left\{ \widehat{F}_0(c) - \widehat{F}_1(c) \right\}$$

Por tanto, la siguiente prueba de hipótesis:

$$H_0 : J = 0 \quad \text{versus} \quad H_1 : J > 0$$

puede resolverse basándose en la prueba clásica de Kolmogorov-Smirnov para la igualdad de distribuciones.

2.5.7. Comparación de pruebas

Dadas dos pruebas es mejor, en el sentido discriminante, la que mayor área tenga y, si fuesen iguales, podría ser porque son la misma curva ROC o porque una nos ofrece una prueba más sensible y otra más específica. Surge la siguiente cuestión, ¿una prueba tiene mayor área que otra única y exclusivamente por tener mayor capacidad discriminante o es debido a la variabilidad de la muestra?, lo que se lleva a plantear el siguiente juego de hipótesis:

$$\begin{cases} H_0 : AUC_1 = AUC_2 \\ H_1 : AUC_1 \neq AUC_2 \end{cases}$$

Hanley y McNeil (92) proponen el siguiente estadístico para resolver este contraste:

$$z = \frac{(\widehat{AUC}_1 - \widehat{AUC}_2)}{SD(\widehat{AUC}_1 - \widehat{AUC}_2)}$$

donde SD es el error estándar calculado de la siguiente forma

$$\begin{aligned} SD(\widehat{AUC}_1 - \widehat{AUC}_2) &= \\ &= \sqrt{\hat{V}(\widehat{AUC}_1) + \hat{V}(\widehat{AUC}_2) - 2 \cdot r \cdot SD(\widehat{AUC}_1) \cdot SD(\widehat{AUC}_2)} \end{aligned}$$

y r es el coeficiente de correlación entre ambas áreas:

$$r = \frac{\text{cov}(\widehat{AUC}_1, \widehat{AUC}_2)}{SD(\widehat{AUC}_1) \cdot SD(\widehat{AUC}_2)}$$

Cuyo valor se obtiene a través de una tabla de correlaciones calculada por Hanley y McNeily (92) unos valores promedios de las correlaciones en el grupo con presencia del evento, $D = 1$ y con ausencia, $D = 0$ y, por otro lado de la área calculadas también en ámbos grupos por separado.

Dicho estadístico sigue una distribución $N(0, 1)$ bajo la hipótesis nula de modo que para un nivel de significación α se rechaza H_0 si $|z| > z_{\alpha/2}$, donde $z_{\alpha/2}$ es el cuantil de dicha distribución que cumple $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

Metodología

3.1. Bases de datos

3.1.1. ENSANUT 2018

Los datos se obtuvieron de la Encuesta Nacional de Salud y Nutrición de México 2018 (ENSANUT 2018). La metodología, los detalles del diseño y la determinación del tamaño de la muestra se describen en el documento elaborado conjuntamente por el Instituto Nacional de Salud Pública de México (INSP) y el Instituto Nacional de Estadística y Geografía (INEGI) (94,95) La base de datos incluye medidas antropométricas, así como la presión arterial. Las tomas de muestras biológicas de sangre periférica se realizó con un equipo calibrado (95). Para el presente estudio, los criterios de inclusión fueron hombres o mujeres mexicanos entre 20 y 59 años de edad, y los criterios de exclusión fueron mujeres embarazadas y participantes con diabetes conocida o desconocida (Glu plasmática en ayunas $> 125 \text{ mg/dL}$), ya que un escenario no sesgado de metabolismo se busca el trastorno, y es probable que todas estas condiciones alteren los niveles de lípidos e insulina (8-10).

3.1.2. Estudio transversal de base comunitaria

Se utilizaron datos de un estudio transversal de base comunitaria (ETBC) en una población compuesta por estudiantes, académicos y personal de apoyo de la Universidad Nacional

Autónoma de México (UNAM) para validar los índices propuestos y encontrar su correlación con diferentes marcadores subrogados de RI y obesidad. Los datos incluyen información sociodemográfica, mediciones antropométricas, medición de la presión arterial y parámetros bioquímicos determinados por los resultados de laboratorio de un ayuno de 8 horas (96).

3.2. Construcción de los índices

3.2.1. Análisis factorial confirmatorio

Se realizaron varios AFCs, que tienen en común cuatro componentes de SM en adultos (PAS, HDL-c, Tri y Glu), pero difieren en la inclusión de una variable que mide la adiposidad, CC (MetSx-CC), IMC (MetSx-IMC) o (MetSx-IRC). Siguiendo a Gurka *et al.* (8–10), se usó el logaritmo natural de Tri para acercarse a la distribución normal, y se utilizó el negativo de HDL-c (HDLN) para que una puntuación de carga factorial más alta fuera similar en interpretación a las otras medidas en el modelo. Posteriormente, se estandarizaron todas las variables (media = 0 y desviación estándar = 1) (CCZ, IMCZ, IRCZ, GluZ, lnTriZ, HDLNZ, PASZ) (Figura 3.1). Las cargas factoriales (λ) para cada una de las variables medidas indican la fuerza de la asociación entre las variables y el factor SM. La estimación de parámetros se realizó por ML. Las ecuaciones que definen el índice MetSx se obtuvieron por transformación inversa de los coeficientes que resultaron del AFC con las variables estandarizadas. Es decir, las ecuaciones MetSx utilizan las medidas originales de las variables para facilitar el uso clínico.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \lambda_{1,1} \\ \lambda_{2,1} \\ \lambda_{3,1} \\ \lambda_{4,1} \\ \lambda_{5,1} \end{bmatrix} [\xi_1] + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{bmatrix}$$

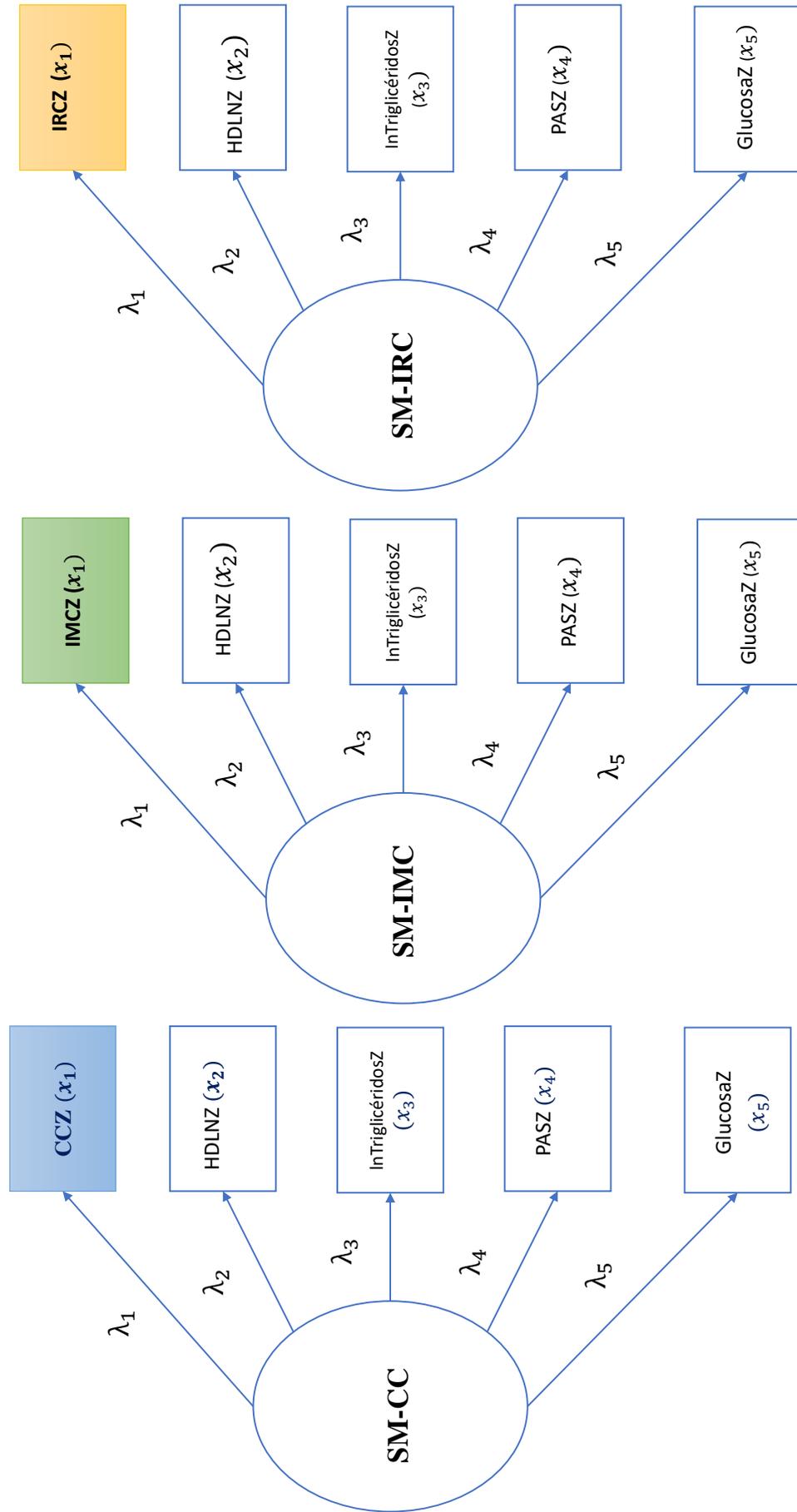


Figura 3.1: Modelos AFC para SM.

La calidad de los modelos AFC ajustados se evaluó utilizando diferentes estadísticas de bondad de ajuste, como el NFI (ajuste deficiente < 0.90), CFI (ajuste deficiente < 0.90), RMSEA (ajuste deficiente > 0.08), SRMR (ajuste deficiente > 0.08), GFI (ajuste deficiente < 0.90) e AGFI (deficiente ajuste < 0.90) (97). Además, se utilizó el AIC para la comparación del modelo, donde un valor pequeño de AIC indica un mejor ajuste.

3.2.2. Validación de los índices propuestos

Con la base de datos ETBC, la presencia de SM se declaró mediante los criterios ATP-III, AHA/HNLBI, FID y el criterio armonizado de la FID.

La concordancia entre los índices MetSx-CC y MetSx-IMC se evaluó mediante el CCI y el CCC (78, 98), donde los valores de CCI por debajo de 0.5 indican baja fiabilidad, los valores entre 0.5 y 0.75 indican moderada fiabilidad, los valores entre 0.75 y 0.9 indican una buena fiabilidad y los valores superiores a 0.90 indican una excelente fiabilidad. McBride (99), mediante el CCC, clasifica la fuerza del acuerdo como pobre (< 0.90), moderada (0.90–0.95), sustancial (0.95-0.99) y casi perfecta (> 0.99). Las curvas ROC se utilizaron para medir la capacidad discriminatoria de los índices sobre el rango de valores posibles en la detección de SM cuantificado por el AUC. Las AUC se compararon con el enfoque no paramétrico de DeLong *et al.* (81, 93).

Los marcadores subrogados conocidos y novedosos de RI y obesidad son HOMA-RI, PAL, SPISE, IAV, TyG, TG/ HDL, METS-RI, METS-VF y TAV se calcularon con las Ecuaciones 2.1-2.10.

Finalmente, se utilizó el coeficiente de correlación lineal de Pearson para medir la asociación entre los marcadores subrogados de RI y de obesidad y los índices MetSx propuestos por sexo con los datos de la población universitaria del ETBC, donde los valores de 0 a 0.3 (o de 0 a -0.3) son insignificantes; 0.3 a 0.5 (o -0.3 a -0.5) son débiles; 0.5 a 0.7 (o -0.5 y -0.7) son moderados; 0.7 a 0.9 (o -0.7 a -0.9) son fuertes; y las correlaciones > 0.9 (o < -0.9) son muy fuertes (100).

El análisis estadístico se realizó en el software R con un nivel de significancia de 0.05, mediante las funciones *cfa* del paquete *lavaan* (101), *CCC* del paquete *DescTools* (102), *icc* de *irr* (103) y *roc.test* del paquete *pROC* (104).

Resultados

4.1. Derivación de los índices propuestos y validación interna en ENSANUT

4.1.1. Análisis exploratorio

La base de datos ENSANUT 2018 tenía 13,100 observaciones. Se consideraron los registros de personas con edad menor a 59 años, por lo que 3,063 personas fueron eliminadas del estudio. Se excluyeron 2,698 registros de personas que no tenían reportado algunos de los valores de altura, peso ó CC, así como también aquellos que no tenían resultados de pruebas de laboratorio como Glu, Tri u otros parámetros bioquímicos. Además, se excluyeron del estudio 138 mujeres embarazadas y 634 personas con $\text{Glu} > 125 \text{ mg/dL}$ (Figura 4.1).

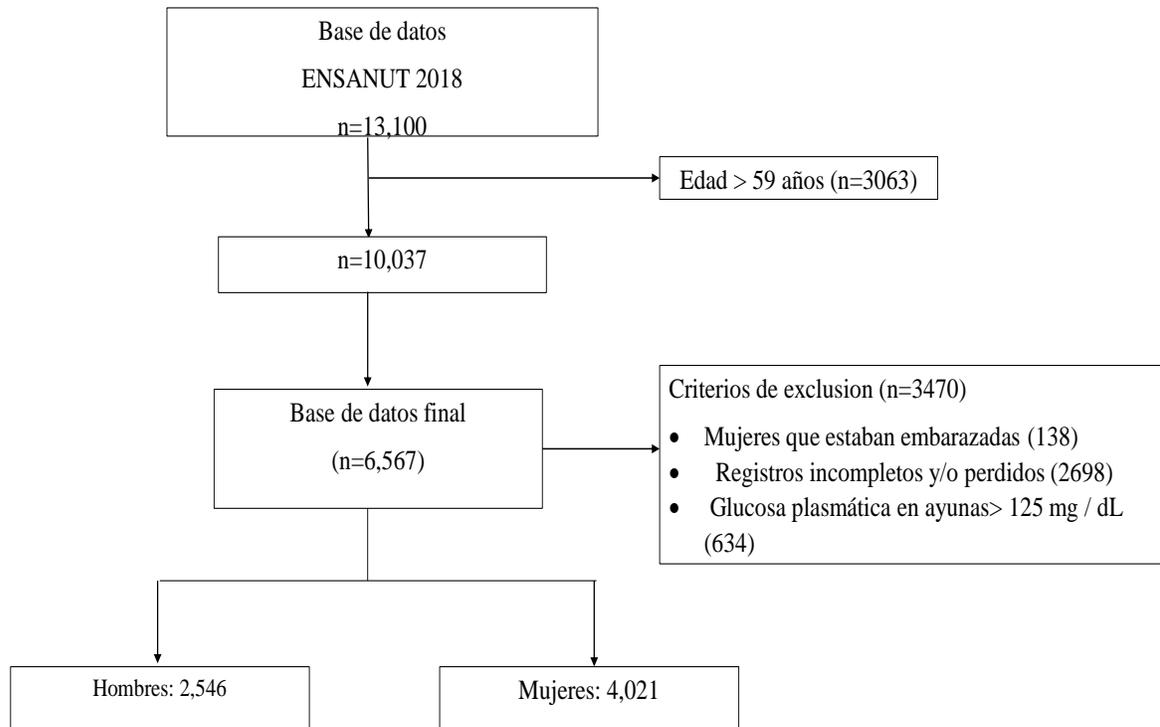


Figura 4.1: Diagrama de flujo para la obtención de la base de datos para la formulación de MetSx.

La muestra final estuvo compuesta por 6,567 adultos (2,546 hombres y 4,201 mujeres). La edad media fue de 37.90 (DE 11.41) años para los hombres y 38.15 (DE 10.77) años para las mujeres. Las estadísticas descriptivas de la base de datos ENSANUT 2018 se muestran en el Cuadro 4.1. El promedio de la CC tanto en hombres como en mujeres es superior a los valores de corte propuestos por la FID y el CA. La media del IMC en ambos sexos indica sobrepeso ($25.0 \leq \text{IMC} \leq 29.9$). El promedio de la PAS es mayor en hombres que en mujeres, la Glu media está por debajo de los valores de corte propuestos por los diferentes criterios y el promedio de los Tri en mujeres se acerca al valor de corte de los diferentes criterios (≥ 150 mg/dL).

Cuadro 4.1: Características de los adultos de 20 a 59 años de ENSANUT 2018 sobre componentes del SM y otros parámetros bioquímicos (n = 6, 567).

	Media (DE)		
	Total	Mujeres (n=4,201)	Hombres (n=2,546)
Antrpometría			
Edad (años)	38.05 (11.02)	38.15 (10.77)	37.90 (11.41)
CC (cm)	94.06 (13.72)	93.33 (13.49)	95.21 (14.00)
IMC (kg/m ²)	28.60 (5.75)	29.20 (5.93)	27.64 (5.31)
IRC (kg/m ²)	5.32 (2.29)	5.75 (2.21)	4.92 (2.21)
Presión arterial			
PA sistólica (mmHg)	131.59 (39.63)	127.84 (37.70)	137.50 (34.89)
PA diastólica (mmHg)	92.53 (48.45)	90.97 (48.54)	94.99 (48.21)
Parámetros bioquímicos			
Glucosa (mg/dL)	89.26 (11.68)	89.03 (11.58)	89.63 (11.64)
HbA1c (%)	5.19 (0.55)	5.23 (0.54)	5.15 (0.57)
Insulina (μ UI/ml)	12.25 (13.71)	13.40 (14.72)	10.44 (11.72)
Triglicéridos (mg/dL)	144.37 (53.51)	141.29 (52.79)	149.25 (54.29)
Colesterol (mg/dL)	179.23 (35.56)	178 (35.03)	179.92 (36.39)
Colesterol HDL (mg/dL)	45.63 (10.73)	46.34 (10.44)	44.52 (11.07)
Colesterol colesterol (mg/dL)	104.73 (30.21)	104.20 (29.40)	105.56 (31.43)
Creatinina (mg/dL)	0.71 (1.04)	0.62 (0.72)	0.85 (1.39)
Albumina (g/dL)	4.31 (0.88)	4.23 (0.89)	4.44 (0.85)
Marcadores subrogados			
HOMA-IR	2.80 (3.53)	3.05 (3.78)	2.41 (3.06)

4.1.2. Estimación de los modelos AFC

Los estadísticos de bondad de ajuste y el AIC de los modelos ajustados con las variables estandarizadas se presentan en el Cuadro 4.2. Para cada sexo, tanto MetSx-CC, MetSx-IMC como MetSx-IRC tuvieron un RMSEA menor que 0.08, lo que indica un buen ajuste. El resto de los índices indican una excelente bondad de ajuste (AGFI, GFI y CFI son todos superiores

a 0.90 y SRMR es inferior a 0.08). El valor del GFI en MetSx-CC para hombres y mujeres fue de 0.99, es decir, el 99 % de la variación de los datos observados en hombres y mujeres es explicado por el modelo, la interpretación es similar en los demás modelos. El valor del AGFI en MetSx-IMC para hombres fue de 0.98, lo que indica que el 98 % de la varianza de la población se explica por el modelo, la interpretación en MetSx-CC y MetSx-IRC tanto en hombres y mujeres es similar.

Cuadro 4.2: Índices de bondad de ajuste para los modelos del AFC.

	MetSx-CC		MetSx-IMC		MetSx-IRC		Ref
	Hombre	Mujer	Hombre	Mujer	Hombre	Mujer	
RMSEA	0.06	0.05	0.05	0.06	0.06	0.04	< 0.08
SRMR	0.03	0.02	0.02	0.03	0.03	0.02	< 0.08
AGFI	0.97	0.98	0.98	0.97	0.97	0.98	> 0.90
GFI	0.99	0.99	0.98	0.99	0.99	0.99	> 0.90
CFI	0.93	0.95	0.95	0.92	0.93	0.95	> 0.90
NFI	0.93	0.93	0.92	0.95	0.93	0.95	> 0.90
AIC	35630	56039	35448	56026	35630	56039	

En el Cuadro 4.3 se reportan las cargas factoriales estandarizadas para los índices MetSx-CC, MetSx-IMC y MetSx-IRC. Las cargas factoriales de lnTriz tienen valores grandes en comparación que los otros componentes, sin embargo, bajo MetSx-IRC en hombres fue la que presentó un valor mayor que el componente de obesidad. MetSx-IRC en hombres, las cargas factoriales de HDLNZ, SBPZ y GluZ fueron mayores que con MetSx-CC y MetSx-IMC. Las cargas factoriales de GluZ fueron mayores en mujeres que en hombres. Las cargas factoriales en mujeres fueron similares entre los tres MetSx. Las variables que forman parte de los modelos fueron estadísticamente significativas tanto en hombres y mujeres ($p < 0.01$).

Cuadro 4.3: Estimación del análisis factorial confirmatorio con variables transformadas y estandarizadas.

	CCZ		IMCZ		IRCZ	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
Medida de obesidad	1.000	1.000	1.000	1.000	1.000	1.000
HDLNZ	0.658	0.785	0.661	0.771	0.671	0.981
lnTriZ	0.946	0.898	0.895	0.860	0.980	1.277
PASZ	0.229	0.235	0.233	0.213	0.226	0.283
GluZ	0.691	0.503	0.697	0.485	0.696	0.597

4.1.3. Ecuaciones para SM en adultos mexicanos

Una vez obtenidas las estimaciones de los modelos AFC, las ecuaciones para SM usando la CC como medida de obesidad y siguiendo la metodología de Gurka *et.al* se determinan de la siguiente manera:

$$\text{MetSx-CC-M} = 1 \times \left(\frac{CC - 93.32}{13.49} \right) + 0.658 \times \left(\frac{HDL + 46.33}{10.44} \right) (-1) + 0.946 \times \left(\frac{\ln Tri - 4.87}{0.40} \right) \\ + 0.229 \times \left(\frac{PAS - 127.84}{37} \right) + 0.691 \times \left(\frac{Glu - 89.02}{11.57} \right)$$

$$\text{MetSx-CC-M} = (0.074 \times CC - 6.91) - 0.063 \times HDL - 2.920 + 2.365 \times \ln Tri - 11.51 \\ + 0.006 \times PAS - 0.791 + 0.060 \times Glu - 5.31$$

$$\text{MetSx-CC-M} = -27.441 + 0.074 \times CC - 0.063 \cdot HDL + 2.365 \cdot \ln Tri + 0.006 \cdot PAS + 0.060 \cdot Glu$$

$$\text{MetSx-CC-H} = 1 \times \left(\frac{CC - 95.20}{14.00} \right) + 0.785 \times \left(\frac{HDL + 44.51}{11.07} \right) (-1) + 0.898 \times \left(\frac{\ln Tri - 4.93}{0.40} \right) \\ + 0.235 \times \left(\frac{PAS - 137.50}{34.88} \right) + 0.503 \times \left(\frac{Glu - 89.63}{11.83} \right)$$

$$\text{MetSx-CC-H} = (0.071 \times CC - 6.80) - 0.070 \times HDL - 3.156 + 2.245 \times \ln Tri - 11.067 \\ + 0.006 \times PAS - 0.926 + 0.042 \times Glu - 3.810$$

$$\text{MetSx-CC-H} = -25.76 + 0.071 \cdot CC - 0.070 \cdot HDL + 2.245 \cdot \ln Tri + 0.006 \cdot PAS + 0.042 \cdot Glu$$

Analogamente se obtienen las ecuaciones para los índices que usan IMC o IRC como medida de obesidad.

El Cuadro 4.4 proporciona los índices para calcular el puntaje de severidad (MetSx) por sexo para la población adulta mexicana usando CC, IMC o IRC. Aunque el análisis se realizó utilizando variables transformadas, estos índices se dan en las escalas originales de las variable utilizadas.

Cuadro 4.4: MetSx por sexo para adultos mexicanos.

Usando CC	
Mujeres	$-27.441 + 0.074 \cdot CC - 0.063 \cdot HDL + 2.365 \cdot \ln Tri + 0.006 \cdot PAS + 0.060 \cdot Glu$
Hombres	$-25.760 + 0.071 \cdot CC - 0.070 \cdot HDL + 2.245 \cdot \ln Tri + 0.006 \cdot PAS + 0.042 \cdot Glu$
Usando IMC	
Mujeres	$-24.918 + 0.168 \cdot IMC - 0.063 \cdot HDL + 2.237 \cdot \ln Tri + 0.006 \cdot PAS + 0.060 \cdot Glu$
Hombres	$-23.400 + 0.188 \cdot IMC - 0.069 \cdot HDL + 2.150 \cdot \ln Tri + 0.006 \cdot PAS + 0.041 \cdot Glu$
Usando IRC	
Mujeres	$-27.563 + 0.451 \cdot IRC - 0.088 \cdot HDL + 3.192 \cdot \ln Tri + 0.008 \cdot PAS + 0.050 \cdot Glu$
Hombres	$-25.791 + 0.451 \cdot IRC - 0.064 \cdot HDL + 2.45 \cdot \ln Tri + 0.006 \cdot PAS + 0.060 \cdot Glu$

4.2. Validación de los índices propuestos en la base ETBC

Las características por sexo de los participantes del conjunto de datos de validación se presentan en el Cuadro 4.5. Los hombres tienen una CC mayor que las mujeres. Los parámetros bioquímicos son similares tanto en hombres como en mujeres. Los participantes tienen características diferentes (antropometría, presión arterial y parámetros bioquímicos) en com-

paración con los de ENSANUT 2018. Además, los biomarcadores asociados con la RI, como los niveles de insulina en sangre y HOMA-RI, fueron más altos en las mujeres. El IMC fue similar entre sexos, esto podría sugerir que la RI en la población mexicana tiene más causas fisiológicas a pesar de la obesidad y el sobrepeso. Se han encontrado múltiples situaciones que incluyen genética, estilo de vida y aculturación en el mayor riesgo de obesidad de las mujeres hispanas. Algunos trabajos han reportado que los promedios del IRC en hombres y mujeres son 4.64 (1.88) y 5.16 (2.24), respectivamente. En la base de datos de ETBC la media del IRC fue mayor en mujeres que hombres (5.75 y 4.92) respectivamente, lo cual concuerda con varios trabajos y están muy cercanos a los promedios de los valores reportados.

Cuadro 4.5: Estadísticas descriptivas de la base de datos CBHS para la validación de MetSx.

Parametros	Media (DE)		
	Total	Mujeres (n=561)	Hombres (n=301)
Antropometría			
Edad (años)	40.71 (10.26)	41.13 (10.05)	39.92 (10.62)
CC (cm)	93.59 (12.25)	90.34 (12.10)	95.55 (12.12)
IMC (kg/m ²)	26.85 (4.81)	26.69 (4.94)	26.96 (4.79)
IRC	5.03 (1.89)	5.06 (1.87)	4.71 (1.68)
Presión arterial			
PA Sistólica (mmHg)	111.87 (14.95)	110.12 (14.11)	115.12 (14.48)
PA Diastólica (mmHg)	74.95 (10.92)	72.41 (10.56)	75.75 (10.89)
Parámetros bioquímicos			
Glucosa (mg/dL)	99.01 (35.54)	99.21 (36.32)	95.53 (26.36)
HbA1c (%)	5.41 (1.24)	5.42 (1.37)	5.32 (1.07)
Insulina (μ UI/ml)	8.56 (6.42)	8.52 (6.02)	8.40 (6.48)
Triglicéridos (mg/dL)	169.41 (107.21)	169.10 (99.13)	167.63 (116.55)
Colesterol (mg/dL)	205.00 (37.93)	204.65 (36.50)	205.19 (40.63)
Colesterol HDL (mg/dL)	47.73 (12.35)	47.11 (11.64)	48.28 (13.00)
Colesterol LDL (mg/dL)	123.17 (33.51)	123.17 (33.51)	122.47 (33.74)
Creatinina (mg/dL)	0.81 (0.43)	0.83 (0.55)	0.78 (0.18)
Acido úrico(mg/dL)	5.40 (1.42)	5.38 (1.47)	5.37 (1.42)
Marcadores subrogados			
HOMA-RI	2.18 (2.07)	2.18 (2.24)	2.07 (1.88)
QUICKY	0.16 (0.01)	0.15 (0.02)	0.16 (0.02)
SPISE	5.85 (1.56)	5.87 (1.55)	5.85 (1.58)
METS-RI	41.16 (8.60)	41.10 (8.65)	41.20 (8.72)
TyG	4.76 (0.32)	4.78 (0.31)	4.74 (0.33)
TG/HDL	3.99 (3.21)	3.97 (2.93)	4.02 (3.61)
PAL	61.06 (47.31)	62.81 (45.35)	58.55 (47.22)
IAV	2.92 (2.20)	3.37(2.52)	2.36(2.12)
METS-VF	6.69 (0.57)	6.55(0.56)	6.79 (0.54)
VAT	985.43 (456.35)	803.83(400.30)	1011.75 (498.40)

La prevalencia de SM basada en los cuatro criterios de diagnóstico fue considerablemente mayor en mujeres (Cuadro 4.6) en la base de datos de la UNAM. De acuerdo con el criterio armonizado de la FID, el 44.38 % de las mujeres y el 28.90 % de los hombres tenían SM.

Cuadro 4.6: Prevalencia de SM según los diferentes criterios.

Criteria	Mujeres n (%)	Hombres n (%)
ATP III	94 (16.75)	38 (12.62)
AHA/NHLBI	202 (36.00)	46 (15.28)
FID	235 (41.88)	95 (27.24)
FID armonizado	249 (44.38)	87 (28.90)

En el Cuadro 4.6 se presenta el desempeño de los índices propuestos con diferentes marcadores subrogados de RI y obesidad. Los valores de MetSx se correlacionaron fuertemente con el índice TyG tanto en hombres como en mujeres. SPISE, METS-RI, TG/HDL, PAL y IAV se correlacionaron moderadamente con MetSx en mujeres, pero están fuertemente correlacionados con MetSx en hombres. MetSx tienen una correlación débil con el ácido úrico, METS-VF e IAV y una correlación nula con la creatinina.

En las Figuras 4.3 y 4.2 se presentan los diagramas de Bland Altman los cuales muestran el grado de concordancia con límites del 95 % (también llamados límites de acuerdo) para MetSx-CC, MetSx-IMC y MetSx-IRC tanto en hombres y mujeres después de calcular los tres índices utilizando las Ecuaciones presentadas en el Cuadro 4.4. MetSx-CC y MetSx-IMC presentan un alto grado de acuerdo en los dos sexos, sin embargo MetSx-CC y MetSx-IRC demuestran un grado de acuerdo pobre.

Cuadro 4.7: Correlaciones (IC del 95 %) de parámetros bioquímicos y marcadores subrogados con MetSx.

	Mujeres			Hombres		
	MetSx-CC	MetSx-IMC	MetSx-IRC	MetSx-CC	MetSx-IMC	MetSx-IRC
Parámetros bioquímicos						
Acido úrico	0.11 (0.03, 0.19)	0.11 (0.02, 0.18)	0.12 (0.03, 0.20)	-0.05 (-0.16, 0.05)	-0.07 (-0.18, 0.04)	-0.08 (-0.19, 0.03)
Creatinina	-0.01 (-0.09, 0.07)	-0.02 (-0.10, 0.06)	-0.01 (-0.09, 0.06)	-0.08 (-0.19, 0.03)	-0.08 (-0.19, 0.03)	-0.08 (-0.19, 0.03)
HbA1c (%)	0.72 (0.67, 0.75)	0.73 (0.68, 0.76)	0.72 (0.67, 0.75)	-0.14 (-0.25, -0.03)	-0.14 (-0.25, -0.03)	-0.14 (-0.25, -0.03)
Marcadores subrogados						
ln(HOMA-RI)	0.59 (0.54, 0.64)	0.59 (0.53, 0.64)	0.59 (0.54, 0.64)	0.01 (-0.10, 0.12)	0.01 (-0.10, 0.12)	0.01 (-0.09, 0.13)
QUICKY	-0.57 (-0.62, -0.51)	-0.56 (-0.62, -0.51)	-0.57 (-0.63, -0.52)	-0.23 (-0.33, -0.12)	-0.23 (-0.33, -0.12)	-0.23 (-0.34, -0.12)
SPISE	-0.57 (-0.62, -0.52)	-0.59 (-0.64, -0.54)	-0.57 (-0.62, -0.51)	-0.72 (-0.77, -0.66)	-0.78 (-0.82, -0.73)	-0.69 (-0.74, -0.62)
METS-IR	0.60 (0.54, 0.65)	0.62 (0.57, 0.67)	0.59 (0.54, 0.64)	0.72 (0.66, 0.77)	0.78 (0.74, 0.82)	0.68 (0.61, 0.73)
TyG	0.83 (0.81, 0.86)	0.83 (0.80, 0.85)	0.84 (0.82, 0.86)	0.87 (0.84, 0.89)	0.86 (0.82, 0.88)	0.91 (0.89, 0.92)
TG/HDL	0.56 (0.50, 0.61)	0.55 (0.48, 0.60)	0.56 (0.50, 0.62)	0.73 (0.67, 0.78)	0.72 (0.66, 0.77)	0.77 (0.72, 0.81)
PAL	0.62 (0.57, 0.67)	0.58 (0.52, 0.63)	0.61 (0.56, 0.66)	0.77 (0.73, 0.81)	0.75 (0.75, 0.80)	0.76 (0.71, 0.80)
VAI	0.56 (0.50, 0.61)	0.53 (0.47, 0.59)	0.56 (0.50, 0.62)	0.74 (0.69, 0.79)	0.72 (0.66, 0.77)	0.78 (0.73, 0.82)
METS-VF	0.36 (0.29, 0.43)	0.34 (0.26, 0.41)	0.35 (0.28, 0.42)	0.47 (0.38, 0.56)	0.48 (0.39, 0.57)	0.40 (0.31, 0.49)
VAT	0.39 (0.31, 0.45)	0.37 (0.29, 0.43)	0.38 (0.31, 0.45)	0.46 (0.37, 0.55)	0.48 (0.39, 0.56)	0.40 (0.30, 0.49)

Diagrama de Bland Altman

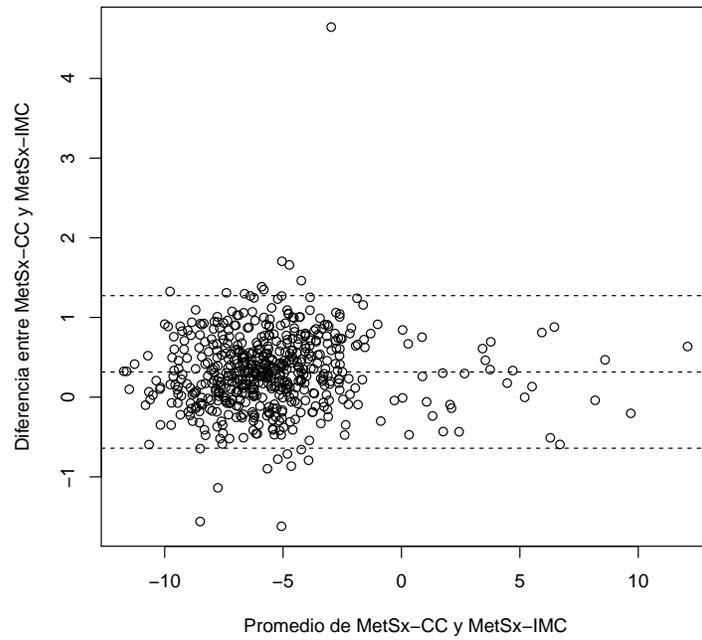


Diagrama de Bland Altman

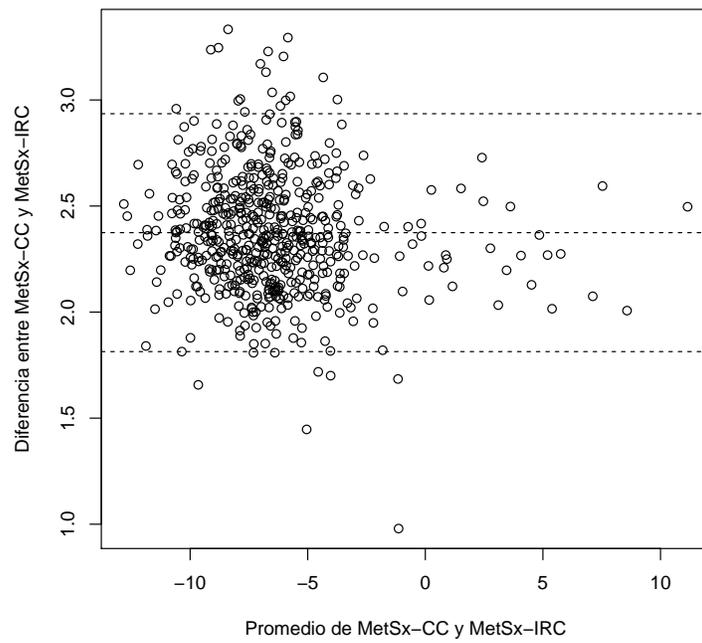


Figura 4.2: Gráficos de Bland Altman para los índices en mujeres.

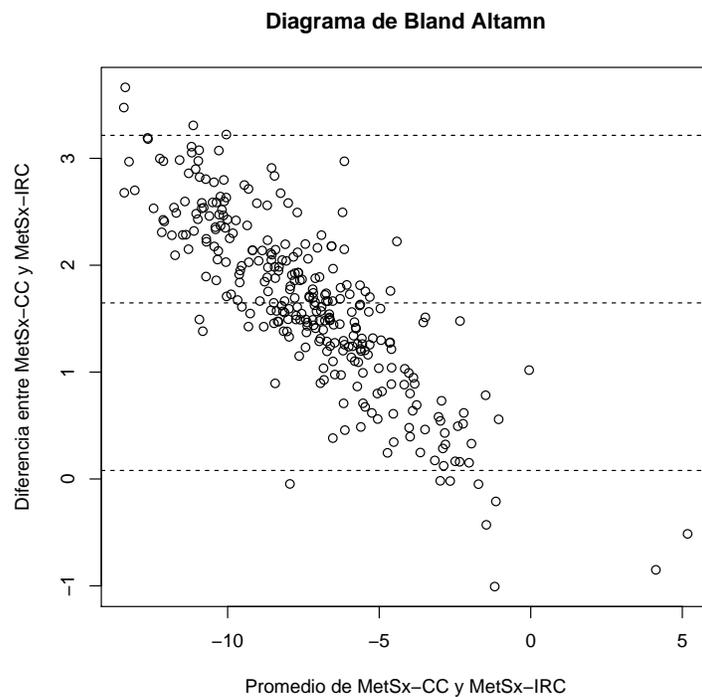
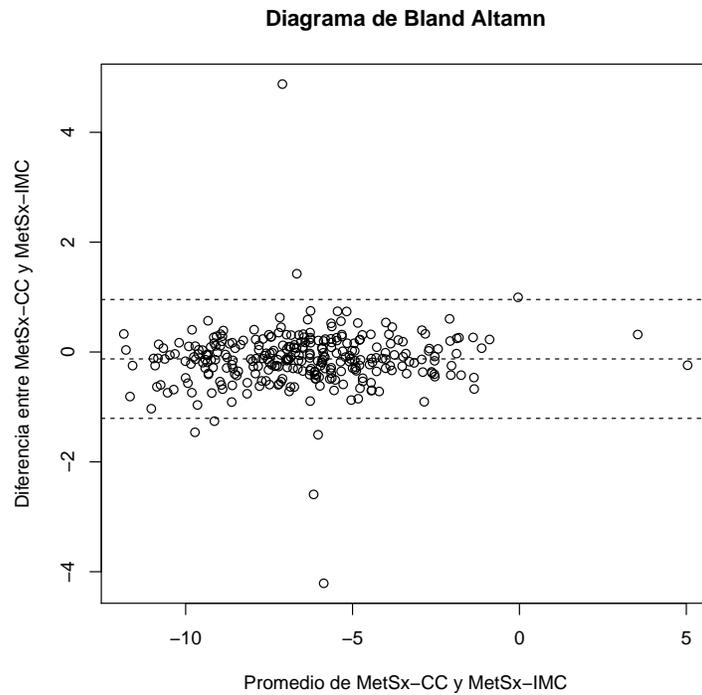


Figura 4.3: Gráficos de Bland Altman para los índices en hombres.

Los valores del CCI para MetSx-IMC y MetSx-IRC fueron, 0.982 y 0.975 para mujeres y hombres, lo cual indica que existe un grado de fiabilidad entre los dos MetSx. Por otro lado, los valores de CCC también muestran una excelente concordancia entre estos dos MetSx, son

0.982 y 0.975 para mujeres y hombres, respectivamente. Respecto a MetSx-CC y MetSx-IRC, los valores del CCI por sexo indican un grado de fiabilidad bueno y el CCC indica un grado de acuerdo pobre en hombres y mujeres (Cuadro 4.8).

Cuadro 4.8: Coeficientes de concordancia y correlación intraclase entre índices.

	MetSx-CC			
	CCI (95 % IC)		CCC (95 % IC)	
	Mujeres	Hombres	Mujeres	Hombres
MetSx-IMC	0.982 (0.979, 0.975)	0.975 (0.969, 0.980)	0.982 (0.979, 0.985)	0.975 (0.969, 0.980)
MetSx-IRC	0.744 (0.705, 0.779)	0.818 (0.777, 0.852)	0.772 (0.750, 0.792)	0.830 (0.804, 0.852)

El análisis ROC de las puntuaciones MetSx en la base de datos ETBC muestra una excelente capacidad para predecir una clasificación binaria de SM determinada por las definiciones operativas (Valor mínimo en hombres 0.875 y en mujeres 0.903). MetSx-CC tuvo el AUC más grande para ambos sexos en cualquier criterio. En las mujeres, el AUC fue menor en todos los casos para MetSx-IMC y en los hombres los AUCs mayores se obtuvieron con MetSx-CC bajo las cuatro definiciones (Cuadro 4.9).

En el Cuadro 4.9 se comparan las AUCs de los diferentes índices para predecir el SM tanto en hombres como en mujeres bajo los diferentes criterios. MetSx-IRC tuvo un AUC mayor que MetSx-IMC en mujeres bajo los diferentes criterios de SM, pero al hacer la comparación de las AUCs con el criterio de Delong, bajo ATP-III y el FID armonizado las AUCs fueron estadísticamente iguales ($p > 0.05$) y en los dos criterios restantes MetSx-IMC tuvo un AUC mayor ($p < 0.05$). Bajo los criterios ATP-III y FID las AUCs de los índices MetSx-CC y MetSx-IMC fueron estadísticamente iguales ($p > 0.05$), con los criterios AHA/NHLBI y FID armonizado, MetSx-CC presentó un AUC mayor que MetSx-IMC ($p < 0.05$). En hombres, las AUCs de los MetSx-IMC y MetSx-CC fueron estadísticamente iguales ($p > 0.05$). Bajo los criterios FID y FID armonizado MetSx-CC tuvo un AUC mayor que MetSx-IRC ($p < 0.01$) y con los dos criterios restantes los dos índices tuvieron AUCs estadísticamente iguales.

En las Figuras 4.4 y 4.5 se presentan las AUC de MetSx-CC para hombres y mujeres con

los 4 criterios, así como el valor del índice de Youden.

Cuadro 4.9: AUC de MetSx para el diagnóstico binario de SM.

Índice	Mujeres			Hombres		
	AUC	IC 95 %	p	AUC	IC 95 %	p
ATP III						
MetSx-CC	0.930	(0.908, 0.953)	Referencia	0.934	(0.902, 0.966)	Referencia
MetSx-IMC	0.925	(0.901, 0.948)	0.30	0.932	(0.900, 0.964)	0.58
MetSx-IRC	0.931	(0.908, 0.953)	0.77	0.923	(0.887, 0.958)	0.02
AHA/NHLBI						
MetSx-CC	0.935	(0.914, 0.956)	—	0.925	(0.885, 0.966)	—
MetSx-IMC	0.918	(0.895, 0.941)	< 0.01	0.923	(0.883, 0.963)	0.68
MetSx-IRC	0.929	(0.908, 0.951)	< 0.01	0.913	(0.868, 0.958)	< 0.01
FID						
MetSx-CC	0.916	(0.895, 0.938)	—	0.884	(0.844, 0.923)	—
MetSx-IMC	0.903	(0.880, 0.927)	< 0.01	0.875	(0.834, 0.916)	0.20
MetSx-IRC	0.911	(0.889, 0.934)	0.18	0.877	(0.836, 0.918)	0.15
FID armonizado						
MetSx-CC	0.939	(0.921, 0.958)	—	0.902	(0.866, 0.938)	0.20
MetSx-IMC	0.929	(0.909, 0.949)	0.01	0.894	(0.856, 0.931)	0.20
MetSx-IRC	0.937	(0.918, 0.956)	0.33	0.898	(0.860, 0.935)	0.14

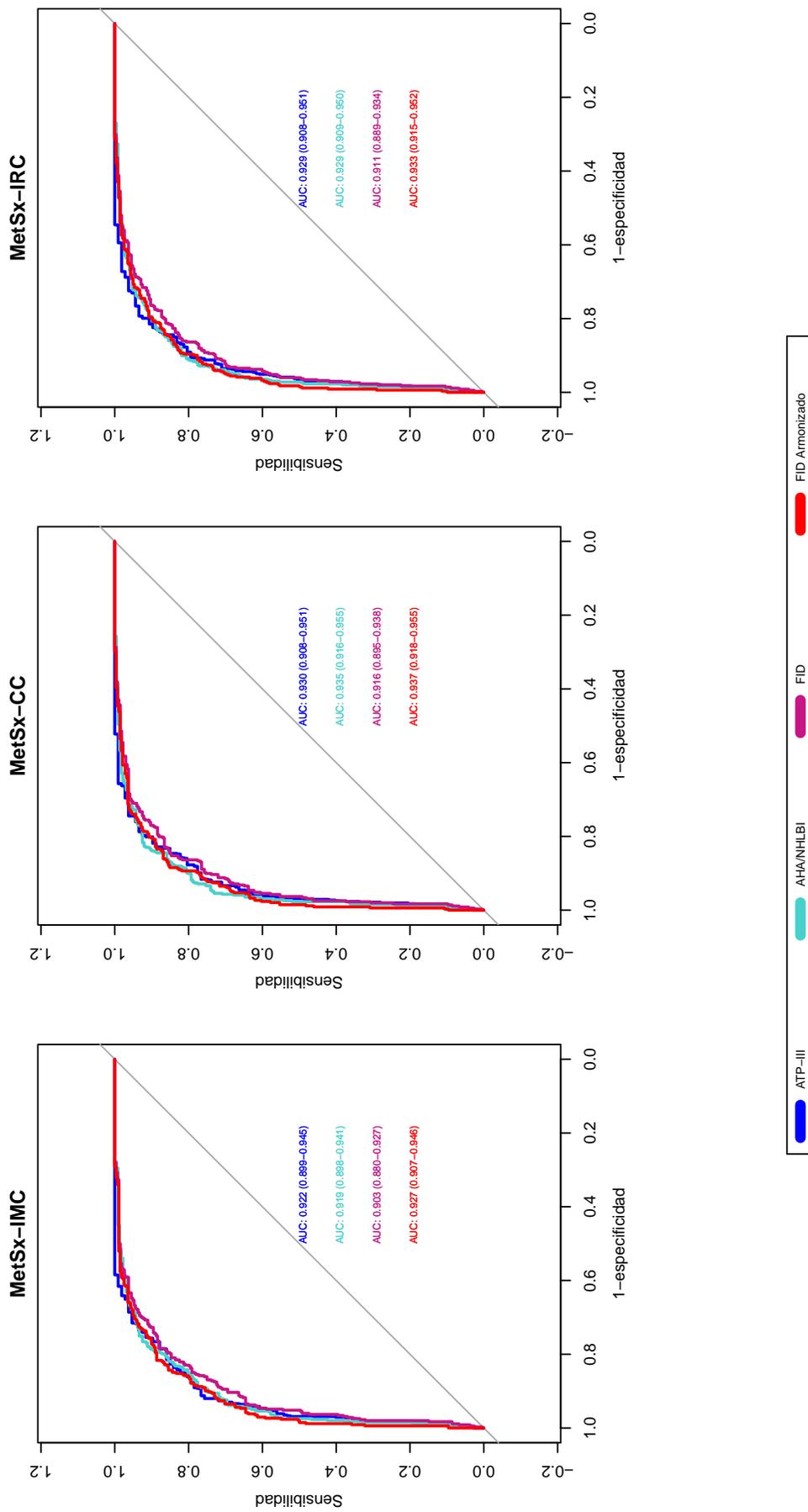


Figura 4.4: AUCs de los índices en mujeres.

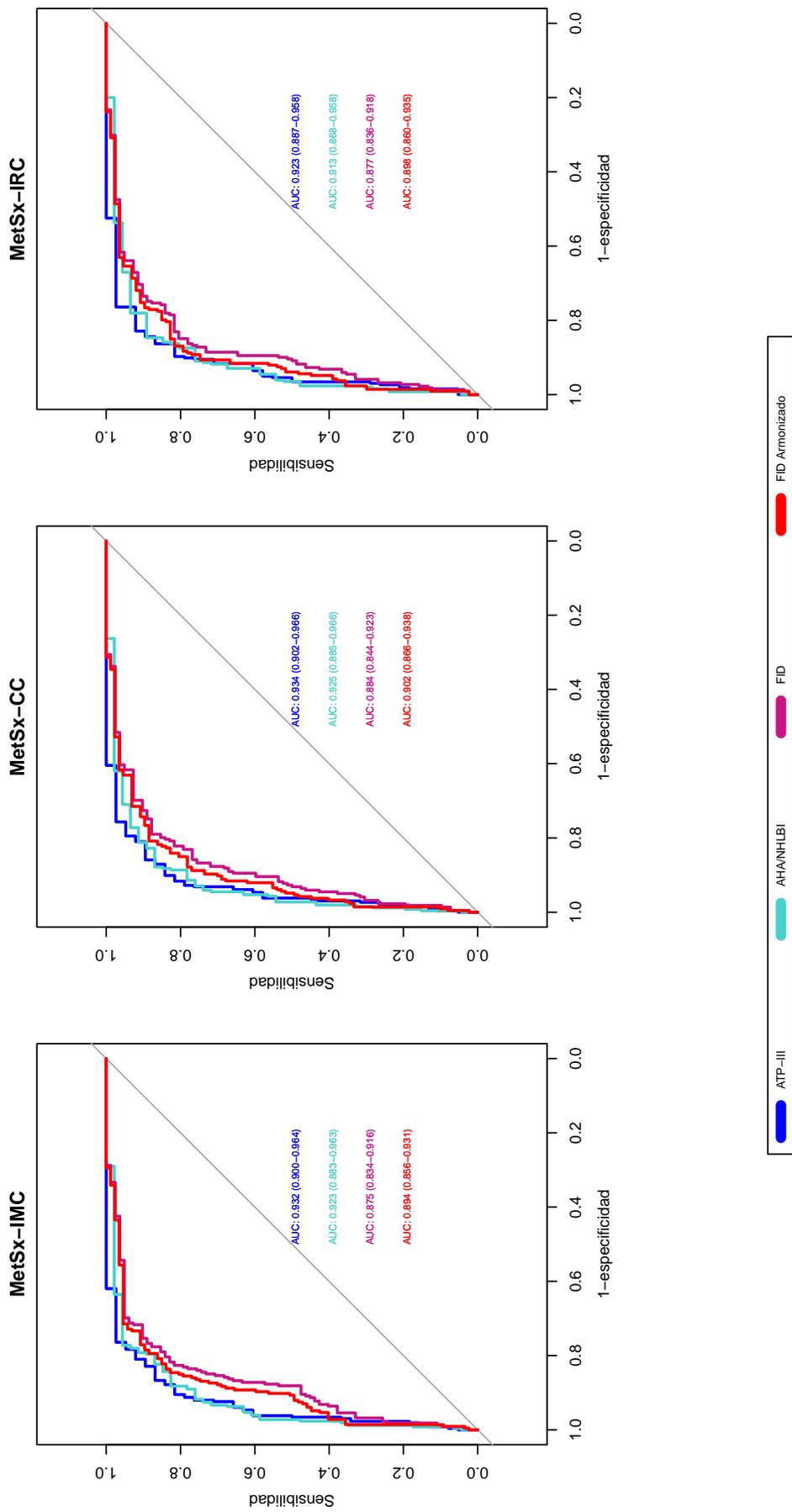


Figura 4.5: AUCs de los índices en hombres.

Discusión y Conclusiones

5.1. Discusión

Se ha reportado de una alta prevalencia de SM entre adultos mexicanos sanos, en comparación con los informes de otros países, incluidos los Estados Unidos y América Latina (104). Las estimaciones de la prevalencia de SM se han incrementado de manera sostenida en México, principalmente debido a varios componentes multifactoriales en los mexicanos, incluidos factores genéticos, un aumento en el consumo de dietas altas en carbohidratos, una alta prevalencia de sedentarismo y un uso limitado de políticas públicas dirigidas a reducir el impacto del SM en la población general (105). El componente más común del SM en los mexicanos es la AO y el colesterol HDL bajo (6). Entonces, es necesario desarrollar índices para cuantificar la severidad del SM en la población mexicana.

El presente estudio muestra diferencias de sexo entre adultos en los pesos factoriales estandarizados (PFE) del SM (Cuadro 4.2), lo que resulta en asociaciones moderadas o fuertes de los índices MetSx con marcadores subrogados de resistencia a la insulina y obesidad. En las puntuaciones MetSx para mujeres, el PFE de las medidas de obesidad (CC, IMC e IRC), Tri y Glu tienen el mayor peso relativo en las puntuaciones MetSx-CC MetSx-IMC y MetSx-IRC. Por otro lado, en las puntuaciones de los hombres, los componentes con mayor peso relativo son la medida de obesidad, Tri y HDL. Este dimorfismo sexual entre los componentes del SM se ha informado en las puntuaciones desarrolladas por Gurka *et al.* (8-10). En el grupo

étnico hispano de la población estadounidense, los componentes con mayor peso relativo en la puntuación son, en orden decreciente, Tri, HDL, obesidad abdominal, Glu y PAS en mujeres, y Tri, medida de obesidad, HDL, Glu, y PAS en hombres (10). Por el contrario, en la población mexicana, los componentes de SM con mayor peso relativo tanto en las puntuaciones de MetSx-CC como en las de MetSx-IMC y MetSx-IRC fueron la medida de obesidad y Tri, seguidos de Glu, HDL-c y PAS en mujeres y de HDL, Glu y PAS en hombres. Claramente, nuestros hallazgos son consistentes con los reportados en la población mexicana de que la AO y la dislipidemia son los componentes comunes del SM (6).

En la población universitaria, encontramos que los marcadores subrogados (conocidos y novedosos) de resistencia a la insulina y obesidad tienen una correlación de moderada a fuerte con los puntajes MetSx. El índice TyG, SPISE, METS-RI y la razón TG/HDL no tienen una asociación débil con las puntuaciones de MetSx, lo que muestra la asociación de resistencia a la insulina y SM. Diversos autores han identificado los índices PAL, IAV y TyG como los de mejor rendimiento para identificar casos de SM mediante el uso de criterios binarios de MetS (113–115). Gurka *et al.* informó una correlación moderada de puntuaciones con ln(HOMA-RI) (9, 116). En la población universitaria, las mujeres mostraron una asociación moderada de ln(HOMA-RI) con las puntuaciones MetSx, pero hubo una asociación nula en los hombres. PAL es un índice utilizado para evaluar la sobreacumulación de lípidos en adultos (62), encontramos una asociación moderada y fuerte en mujeres y hombres con ambas puntuaciones MetSx. Rotter *et al.* (117) informaron asociaciones entre PAL y SM y sus componentes en hombres de edad avanzada. Amato *et al.* (63) derivó y asoció el IAV con riesgo cardiometabólico. Encontramos una asociación de moderada a fuerte de IAV con las tres puntuaciones de los índices. Previamente, Wang y colaboradores (118) encontraron una asociación longitudinal de las puntuaciones de MetSx con el biomarcador de estrés oxidativo que conduce a ECV.

Utilizando el conjunto de datos de la población universitaria, se encontraron prevalencias estimadas de SM muy diferentes utilizando los cuatro criterios de identificación. La menor prevalencia estimada se obtiene con el criterio ATP-III, seguida de la obtenida con el criterio AHA/NHLBI, luego con el criterio FID, y la mayor prevalencia se obtiene con FID armonizado; los dos últimos fueron notablemente más altos (Cuadro 4.6). Con todos los criterios, la prevalencia estimada de SM es mayor para las mujeres que para los hombres. En las

mujeres, la prevalencia fue mayor según el criterio armonizado de la FID. Estas variaciones entre las prevalencias de los criterios de SM ya han sido reportadas por Espósito *et al.* (119), donde los criterios ATP-III identifican un número menor de casos de SM en comparación con los otros criterios. MetSx-IRC mostró un alto grado de fiabilidad entre MetSx-CC y un grado de fiabilidad bueno entre MetSx-IMC, aunque el grado de acuerdo entre MetSx-IRC y MetSx-IMC fue pobre. MetSx-IMC y MetSx-CC mostraron una excelente fiabilidad y grado de concordancia entre las dos puntuaciones por sexo. Un resultado similar fue encontrado por Gurka *et al.* (10) al validar la intercambiabilidad de las medidas de obesidad en la derivación de la puntuación. Las puntuaciones de MetSx mostraron una capacidad casi perfecta para discriminar una clasificación basada en los diferentes criterios de SM ($AUC > 0.89$). El AUC más alto se encuentra en los índices del FID armonizado en mujeres. En las mujeres, las AUCs de MetSx-IRC fueron más altas que las de MetSx-IMC con los diferentes criterios. MetSx-IRC y MetSx-CC tuvieron AUCs estadísticamente iguales bajo ATP-III y el FID armonizado. En hombres, MetSx-CC y MetSx-IMC tuvieron AUCs estadísticamente iguales bajo los cuatro criterios. MetSx-IRC y MetSx-CC tuvieron AUCs iguales con los criterios FID y FID armonizado.

Hasta el momento, no existen trabajos de investigación en México donde el IRC sea utilizado como medida de obesidad para predecir SM u otras enfermedades. Un estudio en adultos peruanos encontró que el IRC predecía mejor el SM que el IMC (120) al igual que otro estudio realizado en adultos polacos (38). En una revisión sistemática y meta-análisis se encontró que el IRC es superior en comparación a otros índices antropométricos para predecir SM (37). Existe literatura que apoya la asociación entre IRC y dislipidemia, y este índice puede estimar mejor los factores de riesgo cardiovascular (121–123). Chang y colaboradores (40) encontraron que el IRC era una medida superior en comparación con el IMC y CC para determinar la presencia de hipertrofia ventricular izquierda, además de que el IRC era un indicador superior de múltiples factores de riesgo cardio-metabólicos en adultos chinos, en comparación con otros índices (39). IRC y CC se han desempeñado mejor que el IMC en las mujeres para predecir glucemia y presión alta, así como también IRC ha demostrado una capacidad superior que el IMC para predecir hiperucemia (43, 120). En un estudio en China que incluía a 15,078 participantes, el IRC mostró un AUC para la clasificación de personas con prediabetes que el IRC (124). Recientemente en un estudio realizado en China

se encontró que el IRC tiende a funcionar de manera óptima en la predicción de diabetes entre la población hipertensa en comparación que la CC e IMC (125), es decir, los pacientes hipertensos con IRC elevado, independientemente del estado general de obesidad, tenían un mayor riesgo de diabetes.

Nuestro estudio tiene algunas limitaciones. Utilizamos la base de datos ENSANUT 2018, la cual se obtiene con un diseño muestral probabilístico complejo de todos los estados mexicanos. Aunque la metodología es robusta, es transversal. Estos puntajes deben usarse en el futuro para proporcionar una medida continua para rastrear los cambios en las anomalías relacionadas con el SM en cada sujeto a lo largo del tiempo. Otra limitación es que los puntajes solo se pueden aplicar a la población adulta mexicana. Queda pendiente por determinar de los diferentes valores de corte para los umbrales de riesgo asociados con diferentes enfermedades. La población universitaria del ETBC podría tener sesgo de selección.

5.2. Conclusiones

Propusimos una puntuación continua específica por sexo para medir la gravedad del SM en adultos mexicanos utilizando CC, IMC e IRC como medidas de adiposidad. Los índices propuestos asumen que el SM es un solo factor latente y que la suma de los componentes del SM no tiene el mismo peso. Es importante tener en cuenta que nuestro objetivo no fue explorar la prevalencia ni estudiar los factores del SM en la población mexicana. Ha existido un interés en evaluar la gravedad del SM mediante scores continuos en los últimos años (8–10, 126). Las puntuaciones MetSx-CC, MetSx-IMC y MetSx-IRC pueden ser herramientas importantes para el uso clínico, como la identificación de riesgos y la monitorización adecuada del paciente a lo largo del tiempo. Podrían usarse para cuantificar el grado de gravedad asociado con diabetes, ECV, enfermedad renal crónica, enfermedad de hígado graso disociado por disfunción metabólica (MAFLD), fibrosis avanzada y numerosas neoplasias malignas en poblaciones mexicanas.

Referencias

- [1] Saklayen MG. The Global Epidemic of the Metabolic Syndrome. *Curr Hypertens Rep* 2018;20:1–8. <https://doi.org/10.1007/s11906-018-0812-z>
- [2] Kylin E. Studien ueber das Hypertonie-Hyperglyca “mie-Hyperurika” miesyndrom. *Zentralblatt fuer Innere Medizin* 1923;44:105-27
- [3] Avogaro P, Crepaldi G. Essential hyperlipidemia, obesity and diabetes. *Diabetologia* 1965;1:137.
- [4] Reaven GM. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes*. 1988;37(12):1595–607. Available from: <https://pubmed.ncbi.nlm.nih.gov/3056758/>
- [5] Eberly LE, Prineas R, Cohen JD, Vazquez G, Zhi X, Neaton JD, et al. Metabolic syndrome: risk factor distribution and 18-year mortality in the multiple risk factor intervention trial. *Diabetes Care* 2006;29:123–30. <https://doi.org/10.2337/DIACARE.29.1.123>.
- [6] Campos-Nonato I, Aguilar-Salinas CA, Mendoza-Herrera K, Pedroza-Tobías A, Rojas R, Barquera S. Prevalence and trends of metabolic syndrome in Mexican adults: Data of the National Health and Nutrition Surveys 2006 and 2016. *Res Sq* 2016.
- [7] DeBoer MD, Gurka MJ. Clinical utility of metabolic syndrome severity scores: Considerations for practitioners. *Diabetes, Metab Syndr Obes Targets Ther* 2017;10:65–72. <https://doi.org/10.2147/DMSO.S101624>.
- [8] Gurka MJ, Ice CL, Sun SS, DeBoer MD. A confirmatory factor analysis of the me-

- tabolic syndrome in adolescents: an examination of sex and racial/ethnic differences. *Cardiovasc Diabetol* 2012;11:128. <https://doi.org/10.1186/1475-2840-11-128>.
- [9] Gurka MJ, Lilly CL, Oliver MN, DeBoer MD. An examination of sex and racial/ethnic differences in the metabolic syndrome among adults: A confirmatory factor analysis and a resulting continuous severity score. *Metabolism* 2014;63:218–25. <https://doi.org/10.1016/j.metabol.2013.10.006>.
- [10] Gurka MJ, Filipp SL, Musani SK, Sims M, DeBoer MD. Use of BMI as the marker of adiposity in a metabolic syndrome severity score: Derivation and validation in predicting long-term disease outcomes. *Metabolism* 2018;83:68–74. <https://doi.org/10.1016/j.metabol.2018.01.015>.
- [11] Warraich HJ, Rana JS. Dyslipidemia in diabetes mellitus and cardiovascular disease. *Cardiovasc Endocrinol* 2017;6:27–32. <https://doi.org/10.1097/XCE.0000000000000120>.
- [12] Freemantle N, Holmes J, Hockey A, Kumar S. How strong is the association between abdominal obesity and the incidence of type 2 diabetes? *Int J Clin Pract* 2008;62:1391–6. <https://doi.org/10.1111/j.1742-1241.2008.01805.x>.
- [13] Schneider HJ, Friedrich N, Klotsche J, Pieper L, Nauck M, John U, et al. The predictive value of different measures of obesity for incident cardiovascular events and mortality. *J Clin Endocrinol Metab* 2010;95:1777–85. <https://doi.org/10.1210/jc.2009-1584>.
- [14] Mendoza-Herrera K, Pedroza-Tobías A, Hernández-Alcaraz C, Ávila-Burgos L, Aguilar-Salinas CA, Barquera S. Attributable burden and expenditure of cardiovascular diseases and associated risk factors in Mexico and other selected megacountries. *Int J Environ Res Public Health* 2019;16:1–26. <https://doi.org/10.3390/ijerph16204041>.
- [15] Alberti KG., Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998. 15, 539–553

- [16] Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: A joint interim statement of the international diabetes federation task force on epidemiology and prevention; National heart, lung, and blood institute; American heart association; World heart federation; International . *Circulation* 2009;120:1640–5. <https://doi.org/10.1161/CIRCULATIONAHA.109.192644>.
- [17] Expert panel on detection evaluation and treatment of high blood cholesterol in adults. Executive summary of the third report (NCEP) -adult treatment panel III. *J Am Med Assoc* 2001;285:2486–97.
- [18] Ford ES. Prevalence of the Metabolic Syndrome Defined by the International Diabetes. *Diabetes Care* 2005;28:2745–9.
- [19] Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and management of the metabolic syndrome: An American Heart Association/National Heart, Lung, and Blood Institute scientific statement. *Circulation* 2005;112:2735–52. <https://doi.org/10.1161/CIRCULATIONAHA.105.169404>.
- [20] DeFronzo RA. Insulin resistance, lipotoxicity, type 2 diabetes and atherosclerosis: the missing links. The Claude Bernard Lecture. *Diabetologia*. 2009;53(7):1270-87.
- [21] The IDF consensus worldwide definition of the metabolic syndrome. Available from: [:http://www.idf.org/webdata/docs/IDF_Metasyndrome_definition.pdf](http://www.idf.org/webdata/docs/IDF_Metasyndrome_definition.pdf).
- [22] Tchernof A, Després JP. Pathophysiology of human visceral obesity: an update. *Physiol Rev*. 2013 Jan 1; 93(1):359–404. Available from: <https://pubmed.ncbi.nlm.nih.gov/23303913/>.
- [23] Nuttall FQ. Body Mass Index: Obesity, BMI, and Health: A Critical Review. *Nutr Today*. 2015 May 17; 50(3):117. Available from: </pmc/articles/PMC4890841/>
- [24] Visscher TLS, Seidell JC, Molarius A, van der Kuip D, Hofman A, Witteman JCM. A comparison of body mass index, waist-hip ratio and waist circumference as predictors of all-cause mortality among the elderly: the Rotterdam study. *Int J Obes*

Relat Metab Disord. 2001; 25(11):1730–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/11753597/>

- [25] Paniagua L, Lohsoonthorn V, Lertmaharit S, Jiamjarasrangsi W, Williams MA. Comparison of Waist Circumference, Body Mass Index, Percent Body Fat and Other Measure of Adiposity in Identifying Cardiovascular Disease Risks among Thai Adults. *Obes Res Clin Pract.* 2008 Sep 1; 2(3):215–23. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23599834>
- [26] Chen Z, Yang G, Zhou M, Smith M, Offer A, Ma J, et al. Body mass index and mortality from ischaemic heart disease in a lean population: 10 year prospective study of 220,000 adult men. *Int J Epidemiol.* 2006 Feb; 35(1):141–50. Available from: <https://pubmed.ncbi.nlm.nih.gov/16258057/>
- [27] National Institute for Health and Clinical Excellence [NICE]. Obesity guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children. (2006). Clinical guidelines 43.
- [28] Chen Z, Yang G, Zhou M, Smith M, Offer A, Ma J, et al. Body mass index and mortality from ischaemic heart disease in a lean population: 10 year prospective study of 220,000 adult men. *Int J Epidemiol.* 2006 Feb; 35(1):141–50. Available from: <https://pubmed.ncbi.nlm.nih.gov/16258057/>.
- [29] Lean MJ, Han TS. Waist worries. *Am J Clin Nutr.* 2002 Oct 1; 76(4):699–700. Available from: <https://academic.oup.com/ajcn/article/76/4/699/4677426>
- [30] Zhu S, Heymsfield SB, Toyoshima H, Wang ZM, Pietrobelli A, Heshka S. Race-ethnicity-specific waist circumference cutoffs for identifying cardiovascular disease risk factors. *Am J Clin Nutr.* 2005; 81(2):409–15. Available from: <https://pubmed.ncbi.nlm.nih.gov/15699228>
- [31] Woo J, Ho SC, Yu ALM, Sham A. Is waist circumference a useful measure in predicting health outcomes in the elderly? *Int J Obes* 2002 2610. 2002 Sep 27; 26(10):1349–55. Available from: <https://www.nature.com/articles/0802080>

- [32] Ross R, Berentzen T, Bradshaw AJ, Janssen I, Kahn HS, Katzmarzyk PT, et al. Does the relationship between waist circumference, morbidity and mortality depend on measurement protocol for waist circumference? *Obes Rev*. 2008 Jul; 9(4):312–25. Available from: <https://pubmed.ncbi.nlm.nih.gov/17956544/>
- [33] Yoo EG. Waist-to-height ratio as a screening tool for obesity and cardiometabolic risk. *Korean J Pediatr*. 2016 Nov 1; 59(11):425. Available from: </pmc/articles/PMC5118501/>
- [34] Onat A, Avci GŞ, Barlan MM, Uyarel H, Uzunlar B, Sansoy V. Measures of abdominal obesity assessed for visceral adiposity and relation to coronary risk. *Int J Obes Relat Metab Disord* [Internet]. 2004 Aug ;28(8):1018–25. Available from: <https://pubmed.ncbi.nlm.nih.gov/15197408/>
- [35] Thomas DM, Bredlau C, Bosy-Westphal A, Mueller M, Shen W, Gallagher D, et al. Relationships between body roundness with body fat and visceral adipose tissue emerging from a new geometrical model. *Obesity*. 2013 Nov 1; 21(11):2264–71. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/oby.20408>
- [36] Sardinha LB, Santos DA, Silva AM, Grøntved A, Andersen LB, Ekelund U. A Comparison between BMI, Waist Circumference, and Waist-To-Height Ratio for Identifying Cardio-Metabolic Risk in Children and Adolescents. *PLoS One*. 2016 Feb 1; 11(2). Available from: </pmc/articles/PMC4762486/>
- [37] Rico-Martín S, Calderón-García JF, Sánchez-Rey P, Franco-Antonio C, Martínez Alvarez M, Sánchez Muñoz-Torrero JF. Effectiveness of body roundness index in predicting metabolic syndrome: A systematic review and meta-analysis. *Obes Rev* 2020;21:1–23. Available from: <https://doi.org/10.1111/obr.13023>.
- [38] Suliga E, Ciesla E, Głuszek-Osuch M, Rogula T, Głuszek S, Koziel D. The Usefulness of Anthropometric Indices to Identify the Risk of Metabolic Syndrome. *Nutr* 2019, Vol 11, Page 2598 2019;11:2598. <https://doi.org/10.3390/NU11112598>.
- [39] Xu J, Zhang L, Wu Q, Zhou Y, Jin Z, Li Z, et al. Body roundness index is a superior indicator to associate with the cardio-metabolic risk: evidence from a

cross-sectional study with 17,000 Eastern-China adults. *BMC Cardiovasc Disord* 2021;21:1–12. <https://doi.org/10.1186/S12872-021-01905-X/TABLES/4>.

- [40] Chang Y, Guo X, Li T, Li S, Guo J, Sun Y. A Body Shape Index and Body Roundness Index: Two New Body Indices to Identify left Ventricular Hypertrophy among Rural Populations in Northeast China. *Heart Lung Circ* 2016;25:358–64. <https://doi.org/10.1016/J.HLC.2015.08.009>.
- [41] Li G, Wu H kun, Wu X wei, Cao Z, Tu Y chao, Ma Y, et al. The feasibility of two anthropometric indices to identify metabolic syndrome, insulin resistance and inflammatory factors in obese and overweight adults. *Nutrition* 2019;57:194–201. <https://doi.org/10.1016/J.NUT.2018.05.004>.
- [42] Zaid M, Ameer F, Munir R, Rashid R, Farooq N, Hasnain S, et al. Anthropometric and metabolic indices in assessment of type and severity of dyslipidemia. *J Physiol Anthropol* 2017;36. <https://doi.org/10.1186/S40101-017-0134-X>.
- [43] Calderón-García JF, Roncero-Martín R, Rico-Martín S, Nicolás-Jiménez JM De, López-Espuela F, Santano-Mogena E, et al. Effectiveness of Body Roundness Index (BRI) and a Body Shape Index (ABSI) in Predicting Hypertension: A Systematic Review and Meta-Analysis of Observational Studies. *Int J Environ Res Public Health* 2021;18:11607. <https://doi.org/10.3390/IJERPH182111607>.
- [44] Chang Y, Guo X, Guo L, Li Z, Li Y, Sun Y. The feasibility of two new anthropometric indices to identify hypertension in rural China: A cross-sectional study. *Medicine (Baltimore)* 2016;95. <https://doi.org/10.1097/MD.0000000000005301>.
- [45] Zhang N, Chang Y, Guo X, Chen Y, Ye N, Sun Y. A Body Shape Index and Body Roundness Index: Two new body indices for detecting association between obesity and hyperuricemia in rural area of China. *Eur J Intern Med* 2016;29:32–6. <https://doi.org/10.1016/J.EJIM.2016.01.019>.
- [46] Wilcox G. Insulin and Insulin Resistance. *Clin Biochem Rev.* 2005; 26(2):19. Available from: </pmc/articles/PMC1204764/>

- [47] Singh B, Saxena A. Surrogate markers of insulin resistance: A review. *World J Diabetes*. 2010;1(2):36. Available from: [/pmc/articles/PMC3083884/](https://pubmed.ncbi.nlm.nih.gov/PMC3083884/)
- [48] Björntorp P. Visceral obesity: a “civilization syndrome.” *Obes Res*. 1993; 1(3):206–22. Available from: <https://pubmed.ncbi.nlm.nih.gov/16350574/>
- [49] Janochova K, Haluzik M, Buzga M. Visceral fat and insulin resistance - what we know? *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2019; 163(1):19–27. Available from: <https://pubmed.ncbi.nlm.nih.gov/30398218/>
- [50] Boyko EJ, Fujimoto WY, Leonetti DL, Newell-Morris L. Visceral adiposity and risk of type 2 diabetes: a prospective study among Japanese Americans. *Diabetes Care*. 2000 ;23(4):465–71. Available from: <https://pubmed.ncbi.nlm.nih.gov/10857936/>
- [51] Chan DC, Barrett HPR, Watts GF. Dyslipidemia in visceral obesity: mechanisms, implications, and therapy. *Am J Cardiovasc Drugs*. 2004 ;4(4):227–46. Available from: <https://pubmed.ncbi.nlm.nih.gov/15285698/>
- [52] Yilmaz Y, Younossi ZM. Obesity-associated nonalcoholic fatty liver disease. *Clin Liver Dis*. 2014 Feb; 18(1):19–31. Available from: <https://pubmed.ncbi.nlm.nih.gov/24274862/>
- [53] Mathieu P, Poirier P, Pibarot P, Lemieux I, Després JP. Visceral obesity: the link among inflammation, hypertension, and cardiovascular disease. *Hypertens*. 2009 Apr; 53(4):577–584. Available from: <https://pubmed.ncbi.nlm.nih.gov/19237685>
- [54] Giovannucci E, Michaud D. The role of obesity and related metabolic disturbances in cancers of the colon, prostate, and pancreas. *Gastroenterology*. 2007; 132(6):2208–25. Available from: <https://pubmed.ncbi.nlm.nih.gov/17498513/>
- [55] Shuster A, Patlas M, Pinthus JH, Mourtzakis M. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. *Br J Radiol*. 2012 Jan;85(1009):1–10. Available from: <https://pubmed.ncbi.nlm.nih.gov/21937614/>
- [56] Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting

plasma glucose and insulin concentrations in man. *Diabetologia*. 1985 Jul;28(7):412–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/3899825>

- [57] Paulmichl K, Hatunic M, Højlund K, Jotic A, Krebs M, Mitrakou A, et al. Modification and validation of the triglyceride-to-HDL cholesterol ratio as a surrogate of insulin sensitivity in white juveniles and adults without diabetes mellitus: The single point insulin sensitivity estimator (SPISE). *Clin Chem* 2016;62:1211–9. <https://doi.org/10.1373/clinchem.2016.257436>.
- [58] Katz A, Nambi SS, Mather K, Baron AD, Follmann DA, Sullivan G, et al. Quantitative insulin sensitivity check index: A simple, accurate method for assessing insulin sensitivity in humans. *J Clin Endocrinol Metab* 2000;85:2402–10. <https://doi.org/10.1210/jcem.85.7.6661>.
- [59] Bello-Chavolla OY, Almeda-Valdes P, Gomez-Velasco D, Viveros-Ruiz T, Cruz-Bautista I, Romo-Romo A, et al. METS-IR, a novel score to evaluate insulin sensitivity, is predictive of visceral adiposity and incident type 2 diabetes. *Eur J Endocrinol* 2018;178:533–44. <https://doi.org/10.1530/EJE-17-0883>.
- [60] Giannini C, Santoro N, Caprio S, Kim G, Lartaud D, Shaw M, et al. The triglyceride-to-HDL cholesterol ratio: Association with insulin resistance in obese youths of different ethnic backgrounds. *Diabetes Care* 2011;34:1869–74. <https://doi.org/10.2337/dc10-2234>.
- [61] Guerrero-Romero F, Villalobos-Molina R, Jiménez-Flores JR, Simental-Mendia LE, Méndez-Cruz R, Murguía-Romero M, et al. Fasting Triglycerides and Glucose Index as a Diagnostic Test for Insulin Resistance in Young Adults. *Arch Med Res* 2016;47:382–7. <https://doi.org/10.1016/j.arcmed.2016.08.012>.
- [62] Kahn HS. The “lipid accumulation product” performs better than the body mass index for recognizing cardiovascular risk: A population-based comparison. *BMC Cardiovasc Disord* 2005;5:1–10. <https://doi.org/10.1186/1471-2261-5-26>.
- [63] Amato MC, Giordano C, Galia M, Criscimanna A, Vitabile S, Midiri M, et al. Visceral adiposity index: A reliable indicator of visceral fat function associated with cardiometabolic risk. *Diabetes Care* 2010;33:920–2. <https://doi.org/10.2337/dc09-1825>.

- [64] Bello-Chavolla OY, Antonio-Villa NE, Vargas-Vázquez A, Viveros-Ruiz TL, Almeda-Valdes P, Gomez-Velasco D, et al. Metabolic Score for Visceral Fat (METS-VF), a novel estimator of intra-abdominal fat content and cardio-metabolic health. *Clin Nutr* 2020;39:1613–21. <https://doi.org/10.1016/j.clnu.2019.07.012>.
- [65] Schumacker RE, Lomax RG. A beginner's guide to structural equation modeling. Mahwah, NJ:Lawrence Erlbaum Associates 2004.
- [66] Bollen KA. Structural Equations with Latent Variables. New York: John Wiley; 1989.
- [67] Manzano PA, Zamora MS. Sistema de ecuaciones estructurales: una herramienta de investigación. Cuaderno técnico 4. Centro Nacional de Evaluación para la Educación Superior, CENEVAL, México 2013.
- [68] Ullman JB. Structural equation modeling: reviewing the basics and moving forward. *J Pers Assess*. 2006; 87(1):35–50. Available from: <https://pubmed.ncbi.nlm.nih.gov/16856785/>
- [69] Browne M. Generalized least squares estimators in the analysis of covariance structures. *South African Stat J*.1974; 8:1–24
- [70] Steiger JH. Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behav Res*. 1990 ;25:173–80.
- [71] Browne M, Cudeck R: Alternative ways of assessing model fit. Testing structural equation models. Edited by: Bollen KA, Long JS. 1993, Newbury Park, California: Sage, 136-162.
- [72] Mejía M, Cornejo C. Aplicación del modelo de ecuaciones estructurales a la gestión del conocimiento. *Lat Am Caribb J Eng Educ*. 2010;4(1):1–9.
- [73] Hooper D, Coughlan J, Mullen M. Structural equation modelling: guidelines for determining model fit. *Electron J Bus Res Methods*. 2008;6:53–60.
- [74] Leal Kaymaliz C. Coeficiente de Correlación de Concordancia bajo la distribución t multivariada [Tesis Doctoral]. Pontificia Universidad Católica de Chile; 2016.

- [75] Barnhart HX, Haber MJ, Lin LI. An overview on assessing agreement with continuous measurements. *J Biopharm Stat.* 2007 Jul;17(4):529–69.
- [76] Choudhary P, Nagaraja H: *Measuring Agreement: Methodology and Applications.* Wiley Series in Probability and Statistics, New York 2015.
- [77] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999 Apr; 8(2):135–60.
- [78] Lin L: A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989; 45:225–268.
- [79] Swiger LA, Harvey LR, Everson DO, Gregory KE. The variance of intraclass correlation involving groups with one observation. *Biometrics* 1964; 20:818–826.
- [80] Van Belle S: Agreement between raters and groups of raters. Unpublished doctoral dissertation. Liège, Belgium: University of Liège, Department of Mathematics; 2009.
- [81] Hajian-Tilaki K: Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp J Intern Med.* 2013;4(2):627–35.
- [82] Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF: Youden Index and Optimal Cut-Point Estimated from Observations Affected by a Lower Limit of Detection. *Biom J.* 2008 Jun;50(3):419.
- [83] Swets, JA: ROC analysis applied to the evaluation of medical imaging techniques. *Investigative Radiology.* 1979; 14:109–121.
- [84] Swets, JA: Measuring the accuracy of diagnostic systems. *Science.* 1988; 240:1285–1293.
- [85] Bamber D: The area above the ordinal dominance graph and the area below the receiver operating graph. *J Math Psych.* 1975; 12: 387-415.
- [86] Lopez-Ratón M., Cadarso-Suárez C, Lado MJ: *Técnicas de estimación e inferencia de las curvas ROC* 2012. Editorial Academica Española.

- [87] Walsh JS: Limitations to the robustness of binormal ROC curves: Effects of model misspecification and location of decision thresholds on bias, precision, size and power. *Statistics in Medicine*. 1997;16(6):669–679.
- [88] Metz CE, Herman, BA, Shen, JH: Maximum likelihood estimation of Receiver Operating Characteristic (ROC) curves from continuously-distributed data. *Statistics in Medicine*. 1998; 17:1033–1053.
- [89] Robertson EA, Zweig MH, Van Steirteghem AC: Evaluating the Clinical Efficacy of Laboratory Tests. *Am J Clin Pathol*. 1983 Jan 1;79(1):78–86.
- [90] Campbell G: Advances in statistical methodology for the evaluation of diagnostic and laboratory tests. *Statistics in Medicine*. 1994;13:499–508.
- [91] Zou KH, Hall WJ, Shapiro DE: Smooth non-parametric Receiver Operating Characteristic (ROC) curves for continuous diagnostic tests. *Statistics in Medicine*. 1997; 16:2143–2156.
- [92] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29–36.
- [93] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*. 1988 Sep;44(3):837.
- [94] Romero-Martínez M, Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Mojica-Cuevas J, Cuevas-Nasu L, et al. National Health and Nutrition Survey 2018-19: Methodology and perspectives. *Salud Publica Mex* 2019;61:917–23. <https://doi.org/10.21149/11095>.
- [95] Instituto Nacional de Estadística y Geografía. Encuesta Nacional de Salud y Nutrición Diseño conceptual. *Inst Nac Salud Pública* 2018:100–20.
- [96] Stephens CR, Easton JF, Robles-Cabrera A, Fossion R, de la Cruz L, Martínez-Tapia R, et al. The Impact of Education and Age on Metabolic Disorders. *Front Public Heal* 2020;8. <https://doi.org/10.3389/fpubh.2020.00180>.

- [97] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model* 1999;6:155. <https://doi.org/10.1080/10705519909540118>.
- [98] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016;15:155–63. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- [99] McBride GB. A proposal for strength-of-agreement criteria for Lin's Concordance Correlation Coefficient. *NIWA Client Rep* 2005;45:307–10.
- [100] Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012;24:69–71.
- [101] Rosseel Y. Lavaan: An R package for structural equation modeling. *J Stat Softw* 2012;48. <https://doi.org/10.18637/jss.v048.i02>.
- [102] Signorell A et mult. al. DescTools: Tools for descriptive statistics. R package version 0.99.40. 2021.
- [103] Gamer M, Lemon J, Fellows I, Puspendra S. irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1. 2019.
- [104] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12. <https://doi.org/10.1186/1471-2105-12-77>.
- [105] Gutiérrez-Solis AL, Datta Banik S, Méndez-González RM. Prevalence of Metabolic Syndrome in Mexico: A Systematic Review and Meta-Analysis. *Metab Syndr Relat Disord* 2018;16:395–405. <https://doi.org/10.1089/met.2017.0157>.
- [106] Bello-Chavolla OY, Vargas-Vázquez A, Antonio-Villa NE, Del Razo-Olvera FM, Elías-López D, Aguilar-Salinas CA. A high incidence of metabolic syndrome traits in mexicans points at obesity-related metabolic dysfunction. *Diabetes, Metab Syndr Obes Targets Ther* 2021;14:1073–82. <https://doi.org/10.2147/DMSO.S266568>.

- [107] Rico-Martín S, Calderón-García JF, Sánchez-Rey P, Franco-Antonio C, Martínez Alvarez M, Sánchez Muñoz-Torrero JF. Effectiveness of body roundness index in predicting metabolic syndrome: A systematic review and meta-analysis. *Obes Rev.* 2020;21(7):1–23.
- [108] Zhang N, Chang Y, Guo X, Chen Y, Ye N, Sun Y. A Body Shape Index and Body Roundness Index: Two new body indices for detecting association between obesity and hyperuricemia in rural area of China. *Eur J Intern Med.* 2016 Apr 1 ;29:32–6.
- [109] Zaid M, Ameer F, Munir R, Rashid R, Farooq N, Hasnain S, et al. Anthropometric and metabolic indices in assessment of type and severity of dyslipidemia. *J Physiol Anthropol.* 2017 Feb 28;36(1):1–10.
- [110] Calderón-García JF, Roncero-Martín R, Rico-Martín S, Nicolás-Jiménez JM De, López-Espuela F, Santano-Mogena E, et al. Effectiveness of Body Roundness Index (BRI) and a Body Shape Index (ABSI) in Predicting Hypertension: A Systematic Review and Meta-Analysis of Observational Studies. *Int J Environ Res Public Health.* 2021 Nov 4;18(21):11607.
- [111] Chang Y, Guo X, Guo L, Li Z, Li Y, Sun Y. The feasibility of two new anthropometric indices to identify hypertension in rural China: A cross-sectional study. *Medicine (Baltimore).* 2016;95(44).
- [112] Liu Y, Liu X, Guan H, Zhang S, Zhu Q, Fu X, et al. Body Roundness Index Is a Superior Obesity Index in Predicting Diabetes Risk Among Hypertensive Patients: A Prospective Cohort Study in China. *Front Cardiovasc Med.* 2021 Nov 18;0:1467.
- [113] Chiu TH, Huang YC, Chiu H, Wu PY, Chiou HYC, Huang JC, et al. Comparison of various obesity-related indices for identification of metabolic syndrome: A population-based study from taiwan biobank. *Diagnostics* 2020;10. <https://doi.org/10.3390/diagnostics10121081>.
- [114] Yu X, Wang L, Zhang W, Ming J, Jia A, Xu S, et al. Fasting triglycerides and glucose index is more suitable for the identification of metabolically unhealthy individuals in the Chinese adult population: A nationwide study. *J Diabetes Investig* 2019;10:1050–8. <https://doi.org/10.1111/jdi.12975>.

- [115] Shin KA, Kim YJ. Usefulness of surrogate markers of body fat distribution for predicting metabolic syndrome in middle-aged and older Korean populations. *Diabetes, Metab Syndr Obes Targets Ther* 2019;12:2251–9. <https://doi.org/10.2147/DMS0.S217628>.
- [116] Guo Y, Musani SK, Sims M, Pearson TA, Deboer MD, Gurka MJ. Assessing the added predictive ability of a metabolic syndrome severity score in predicting incident cardiovascular disease and type 2 diabetes: The Atherosclerosis Risk in Communities Study and Jackson Heart Study. *Diabetol Metab Syndr* 2018;10. <https://doi.org/10.1186/s13098-018-0344-3>.
- [117] Rotter I, Rył A, Szylińska A, Pawlukowska W, Lubkowska A, Laszczyńska M. Lipid Accumulation Product (LAP) as an Index of Metabolic and Hormonal Disorders in Aging Men. *Exp Clin Endocrinol Diabetes* 2017;125:176–82. <https://doi.org/10.1055/s-0042-116071>.
- [118] Wang LX, Filipp SL, Urbina EM, Gurka MJ, DeBoer MD. Longitudinal Associations of Metabolic Syndrome Severity between Childhood and Young Adulthood: The Bogalusa Heart Study. *Metab Syndr Relat Disord* 2018;16:208–14. <https://doi.org/10.1089/met.2017.0160>.
- [119] Espósito RC, De Medeiros PJ, De Souza Silva F, Oliveira AG, Soares Aragão CF, Oliveira Rocha HA, et al. Prevalence of the metabolic syndrome according to different criteria in the male population during the Blue November Campaign in Natal, RN, Northeastern Brazil. *Diabetes, Metab Syndr Obes Targets Ther* 2018;11:401–8. <https://doi.org/10.2147/DMS0.S168430>.
- [120] Stefanescu A, Revilla L, Lopez T, Sanchez SE, Williams MA, Gelaye B. Using A Body Shape Index (ABSI) and Body Roundness Index (BRI) to predict risk of metabolic syndrome in Peruvian adults. *J Int Med Res* 2020;48. <https://doi.org/10.1177/0300060519848854>.
- [121] Maessen MFH, Eijsvogels TMH, Verheggen RJHM, Hopman MTE, Verbeek ALM, De Vegt F. Entering a New Era of Body Indices: The Feasibility of a Body Shape

Index and Body Roundness Index to Identify Cardiovascular Health Status. PLoS One 2014;9:e107212. <https://doi.org/10.1371/JOURNAL.PONE.0107212>.

- [122] Chang Y, Guo X, Chen Y, Guo L, Li Z, Yu S, et al. A body shape index and body roundness index: two new body indices to identify diabetes mellitus among rural populations in northeast China. BMC Public Health 2015;15.<https://doi.org/10.1186/S12889-015-2150-2>.
- [123] Tian S, Zhang X, Xu Y, Dong H. Feasibility of body roundness index for identifying a clustering of cardiometabolic abnormalities compared to BMI, waist circumference and other anthropometric indices: the China Health and Nutrition Survey, 2008 to 2009. Medicine (Baltimore) 2016;95. <https://doi.org/10.1097/MD.0000000000004642>.
- [124] Zhao Q, Zhang K, Li Y, Zhen Q, Shi J, Yu Y, et al. Capacity of a body shape index and body roundness index to identify diabetes mellitus in Han Chinese people in Northeast China: a cross-sectional study. Diabet Med 2018;35:1580–7.<https://doi.org/10.1111/DME.13787>.
- [125] Liu Y, Liu X, Guan H, Zhang S, Zhu Q, Fu X, et al. Body Roundness Index Is a Superior Obesity Index in Predicting Diabetes Risk Among Hypertensive Patients: A Prospective Cohort Study in China. Front Cardiovasc Med 2021;0:1467.<https://doi.org/10.3389/FCVM.2021.736073>.
- [126] Soldatovic I, Vukovic R, Culafic D, Gajic M, Dimitrijevic-Sreckovic V. siMS Score: Simple Method for Quantifying Metabolic Syndrome. PLoS One 2016;11:e0146143. <https://doi.org/10.1371/JOURNAL.PONE.0146143>.

Anexos

En este Cuadro se presentan los CCC y CCI de los índices MetSx-IMC y MetSx-IRC para hombres y mujeres.

Cuadro 1: Coeficientes de concordancia y correlación intraclase entre MetSx-IRC y MetSx-IMC.

	MetSx-IMC			
	CCI (95 % CI)		CCC (95 % CI)	
	Mujeres	Hombres	Mujeres	Hombres
MetSx-IRC	0.793 (0.761, 0.822)	0.782 (0.734, 0.823)	0.811 (0.791, 0.829)	0.800 (0.769, 0.825)