

# Métodos estadísticos de validación y predicción de medidas de capacidad funcional en personas mayores

## Aportaciones del modelo Rasch y redes bayesianas

Alba Ayala

Unidad de investigación en cuidados y servicios de salud (Investén-ISCIID)

**RICAPPS** - Red de Investigación en Cronicidad, Atención Primaria y Prevención y Promoción de la Salud.

E-mail: [alba.ayala@isciii.es](mailto:alba.ayala@isciii.es)

Seminario en el Centro Nacional de Epidemiología  
13 de Abril de 2023



# Índice

- **Introducción**

- Patient Reported Outcome Measures (PROMs)
- Capacidad funcional
- Calidad de vida relacionada con la salud (CVRS)
- Validación y construcción de escalas
- Estudios de predicción (*mapping*)

- **Hipótesis y objetivos**

- **Métodos**

- **Resultados**

- **Discusión**

- Implicaciones para la investigación, práctica clínica y salud pública

- **Conclusiones**

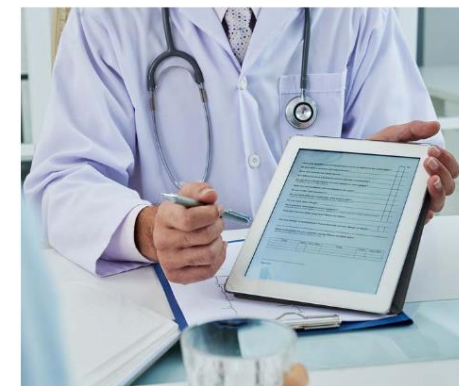
# Introducción

# Patient Reported Outcome Measures (PROMs)

- Las **PROMs** se obtienen a partir de **cuestionarios validados**, respondidos directamente por los pacientes.
- Recogen información desde la **perspectiva del paciente** y no presentan la visión del médico, personal sanitario u otra persona.



- Ejemplos: **capacidad funcional, calidad de vida, gravedad de los síntomas, comportamientos de salud.**



Cella D et al. Types of Patient-Reported Outcomes. RTI Press, 2015.



# Calidad de vida relacionada con la salud (CVRS)

- Evaluación subjetiva sobre la capacidad del individuo para lograr y **mantener** un nivel global de funcionamiento que permite seguir aquellas **actividades** que son **importantes** y que afectan a su estado general de **bienestar**.
- Concepto **multidimensional**: funcionamiento social, físico, y cognitivo; y el bienestar emocional.
- **Índice EQ-5D-5L**: Integra varias dimensiones del constructo de CVRS. Se puede utilizar como **medida de utilidad** en técnicas de evaluación clínica y económica.



Soto M, Failde I. La calidad de vida relacionada con la salud como medida de resultados en pacientes con cardiopatía isquémica. Rev Soc Esp Dolor. 2004

# Calidad de vida relacionada con la salud (CVRS)

EQ-5D-5L



**Introducción P.26:** Entrevistador/a, léale al informante: “Continuando con su estado de salud actual, dígame qué afirmaciones de las siguientes que le voy a leer describe mejor su estado de salud en el día de hoy”:

## (No leer) 26.1- Movilidad

- No tengo problemas para caminar \_\_\_\_\_  1
- Tengo problemas leves para caminar \_\_\_\_\_  2
- Tengo problemas moderados para caminar \_\_\_\_\_  3
- Tengo problemas graves para caminar \_\_\_\_\_  4
- No puedo caminar \_\_\_\_\_  5

## (No leer) 26.2- Autocuidado

- No tengo problemas para lavarme o vestirme \_\_\_\_\_  1
- Tengo problemas leves para lavarme o vestirme \_\_\_\_\_  2
- Tengo problemas moderados para lavarme o vestirme \_\_\_\_\_  3
- Tengo problemas graves para lavarme o vestirme \_\_\_\_\_  4
- No puedo lavarme o vestirme \_\_\_\_\_  5

## (No leer) 26.3- Actividades cotidianas (por ejemplo, trabajar, estudiar, hacer las tareas domésticas, actividades familiares o actividades durante el tiempo libre)

- No tengo problemas para realizar mis actividades cotidianas \_\_\_\_\_  1
- Tengo problemas leves para realizar mis actividades cotidianas \_\_\_\_\_  2
- Tengo problemas moderados para realizar mis actividades cotidianas \_\_\_\_\_  3
- Tengo problemas graves para realizar mis actividades cotidianas \_\_\_\_\_  4
- No puedo realizar mis actividades cotidianas \_\_\_\_\_  5

## (No leer) 26.4- Dolor/malestar

- No tengo dolor ni malestar \_\_\_\_\_  1
- Tengo dolor o malestar leve \_\_\_\_\_  2
- Tengo dolor o malestar moderado \_\_\_\_\_  3
- Tengo dolor o malestar fuerte \_\_\_\_\_  4
- Tengo dolor o malestar extremo \_\_\_\_\_  5

## (No leer) 26.5- Ansiedad/depresión

- No estoy ansioso/a ni deprimido/a \_\_\_\_\_  1
- Estoy levemente ansioso/a o deprimido/a \_\_\_\_\_  2
- Estoy moderadamente ansioso/a o deprimido/a \_\_\_\_\_  3
- Estoy muy ansioso/a o deprimido/a \_\_\_\_\_  4
- Estoy extremadamente ansioso/a o deprimido/a \_\_\_\_\_  5

Herdman M et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011

# Validación y construcción de escalas



- Las medidas de salud, deben ser diseñadas para recoger de manera **válida** y **fiable** la percepción subjetiva del sujeto.
- Las **técnicas** de validación analizan las propiedades psicométricas para garantizar la **reproducibilidad**, validez y **fiabilidad** del instrumento.
- Técnicas más **utilizadas**:  
Teoría Clásica del Test y Teoría de Respuesta al Ítem.



Fiable  
No válida



Baja validez  
y fiabilidad



No válida  
No fiable

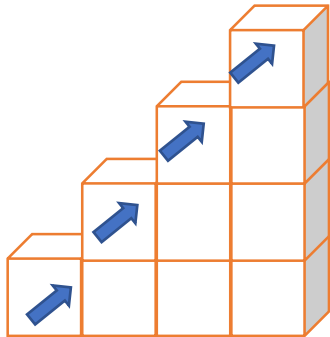


Válida  
Fiable

# Construcción de escalas

- La Teoría Clásica del Test método más utilizado y sencillo de aplicar.

$$X_v = T_v + E_v$$

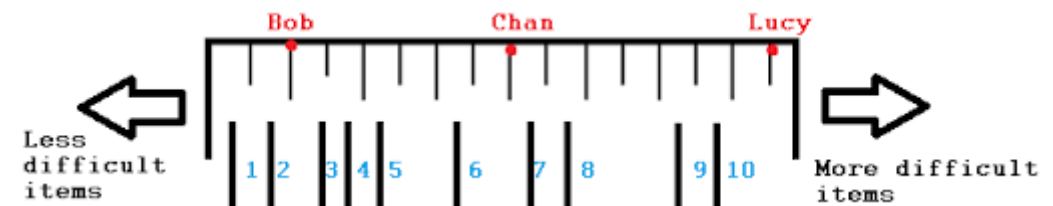


- El escalamiento Guttman es fácil de aplicar, permite ordenar ítems y personas, pero solo puede ser utilizado en cuestionarios de respuesta binaria.

- El modelo Rasch (Teoría de Respuesta al Ítem).

Robusto a valores perdidos y permite obtener una escala lineal de intervalo.

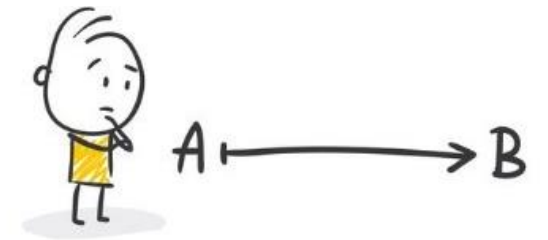
**Inconveniente:** complejo de aplicar.



# Estudios de *mapping*



- Estudios de *mapping*: ayudan a **predecir** el índice de utilidad utilizando la medida específica.
- Muchos estudios solo incluyen medidas **específicas** de una enfermedad.
- Los investigadores tratan de evaluar las **intervenciones** en términos de **coste-efectividad** utilizando **medidas de utilidad** que combinan la **calidad** de vida y la **cantidad** de vida en un solo índice.
- Métodos **directos**: regresión lineal, datos censurados
- Métodos **indirectos** (*response mapping*): logístico multinomial, redes bayesianas



Petrou S et al. The MAPS Reporting Statement for Studies Mapping onto Generic Preference-Based Outcome Measures. *PharmacoEconomics*. 2015

# Hipótesis y objetivos

# Hipótesis

Los métodos estadísticos de validación y predicción complejos proporcionan mejores resultados que métodos más sencillos.

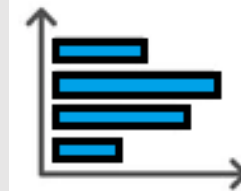


## Objetivos

- 1 Comparar el análisis Rasch con otros métodos de validación de escalas utilizando la escala de capacidad funcional AVD-EDAD2008 (**Estudio 1**).
- 2 Comparar las **redes bayesianas** con otros métodos de mapping para predecir el índice EQ-5D-5L a través de medidas del estado funcional WOMAC, en pacientes con artrosis de cadera y rodilla (**Estudio 2**).
- 3 Comparar las **redes bayesianas** con otros métodos de predicción del índice EQ-5D-5L, a partir de datos de población mayor de la Encuesta Nacional de Salud en España (**Estudio 3**).

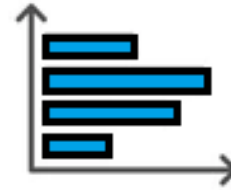
# Métodos

# Métodos



	Estudio 1	Estudio 2	Estudio 3
	<b>Comparación métodos de Validación</b>	<b>Comparación métodos indirectos</b>	<b>Comparación de métodos directos e indirectos</b>
<b>Modelo</b>	Validación	Predicción (response mapping)	Predicción (mapping)
<b>Instrumento de capacidad funcional</b>	<u>AVD-EDAD2008</u> 14 ítems Respuesta binaria 4 opciones de respuesta	<u>WOMAC</u> Dolor (5 ítems), rigidez (2 ítems) y función física (17 ítems) 5 opciones de respuesta	<u>Escala de capacidad funcional de la ENSE</u> 27 ítems 3 opciones de respuesta
<b>Métodos</b>	Análisis Rasch	Redes bayesianas	Redes bayesianas
	Teoría Clásica del Test	Logística multinomial	Logística multinomial
	Análisis Guttman		Regresión logística ordinal
			Regresión lineal
			Tobit

# Métodos

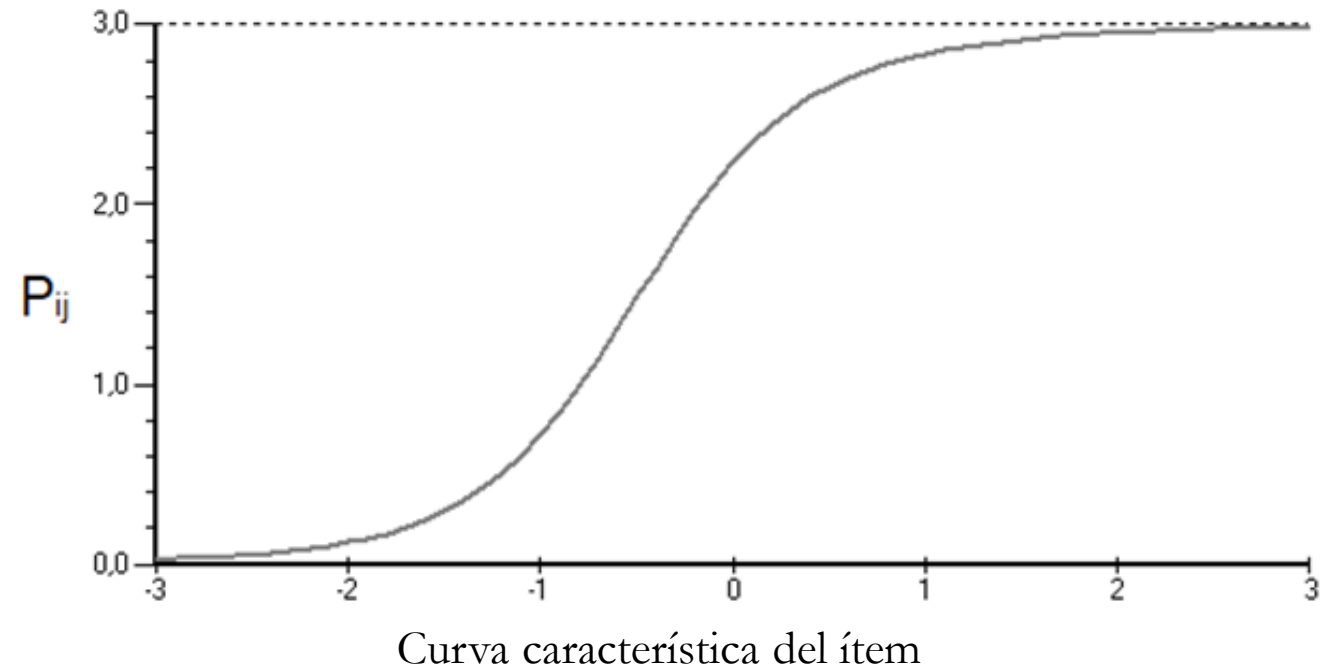


	Estudio 1	Estudio 2	Estudio 3
	<b>Comparación métodos de Validación</b>	<b>Comparación métodos indirectos de mapping</b>	<b>Comparación de métodos directos e indirectos de mapping</b>
<b>Fuente de datos</b>	Encuesta sobre discapacidad, autonomía personal y situaciones de dependencia, INE	Cohorte de hospitales y centros de tres comunidades autónomas	Encuesta Nacional de Salud en España, INE
<b>Periodo de recogida</b>	Noviembre 2007-Febrero 2008	Noviembre 2014-Diciembre 2015	Julio 2011-Junio 2012
<b>Tamaño muestral</b>	n=39.430	n=758	n=5.708
<b>Porcentaje de mujeres</b>	55,9%	61,6%	62,8%
<b>Edad muestra</b>	mayores de 65 años	mayores de 60 años	mayores de 65 años

# Análisis Rasch

El modelo Rasch estima en una escala lineal la habilidad de las personas para contestar el test y la dificultad de los ítems.

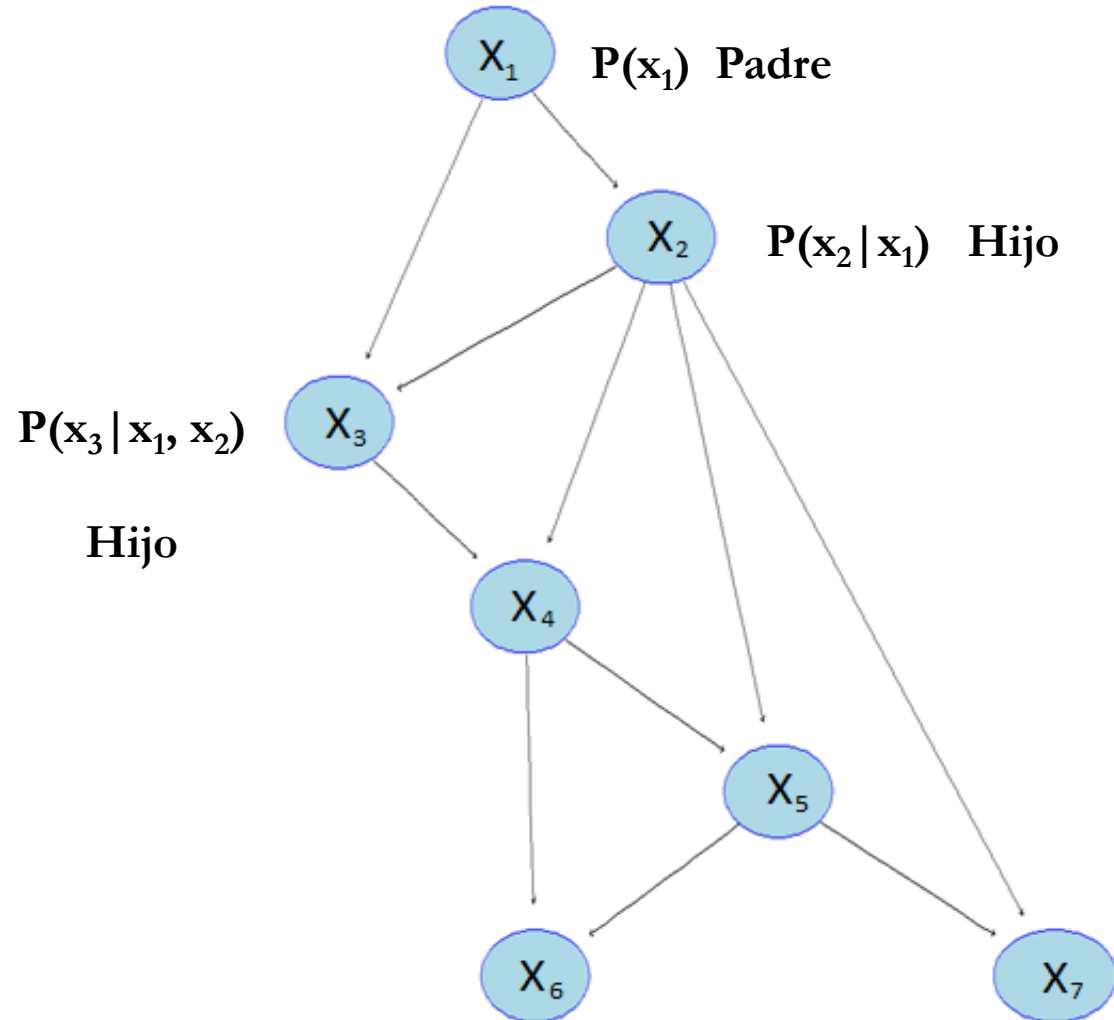
$$P_{ij} = \frac{e^{(\theta_i - \delta_j)}}{1 + e^{(\theta_i - \delta_j)}}$$



Propiedades: Ajuste al modelo, unidimensionalidad, independencia local.

# Redes bayesianas

Grafo dirigido acíclico (DAG)



- Diagrama visual que muestra las relaciones de dependencia e independencia entre las variables
- A cada nodo hijo se le asigna una distribución condicionada dados sus padres.

Distribución de probabilidad conjunta

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i | pa(x_i))$$

# Resultados

# Estudio 1: Comparación métodos de validación

## Objetivo 1:

Comparar el análisis Rasch con otros métodos de validación de escalas utilizando la escala de capacidad funcional AVD-EDAD2008

Gac Sanit. 2019;33(6):511-516

Original

Comparación de métodos de escalamiento de actividades de la vida diaria en personas mayores

Alba Ayala<sup>a,b,\*</sup>, Rogelio Pujol<sup>c</sup>, Maria João Forjaz<sup>b,d</sup> y Antonio Abellán<sup>a</sup>

<sup>a</sup> Instituto de Economía, Geografía y Demografía, Centro de Ciencias Humanas y Sociales, Consejo Superior de Investigaciones Científicas, Madrid, España

<sup>b</sup> Red de Investigación en Servicios de Salud en Enfermedades Crónicas (REDISSEC), España

<sup>c</sup> Instituto Nacional de Estadística, Madrid, España

<sup>d</sup> Departamento de Epidemiología y Bioestadística, Escuela Nacional de Sanidad, Instituto de Salud Carlos III, Madrid, España

---

**Factor de impacto (2021): 2.479 (53/88, Q3, Health Policy & Services)**

**Instrumento capacidad funcional:** AVD-EDAD2008.

14 ítems (respuesta binaria y 4 opciones de respuesta)

Correlaciones entre los diferentes métodos de escalamiento (n = 39.430)

	Rasch dicotómico	Rasch politómico	Guttman	Suma dicotómico	Suma politómico
Correlaciones Pearson					
Rasch dicotómico	1				
Rasch politómico	0,979	1			
Guttman	0,800	0,765	1		
Suma dicotómico	0,993	0,968	0,805	1	
Suma politómico	0,979	0,984	0,789	0,976	1

**Variable de referencia:** “¿Se ha visto limitado para realizar las actividades que la gente habitualmente hace, debido a un problema de salud o discapacidad?”

Opciones de respuesta: gravemente limitado, limitado pero no gravemente, y no limitado.

Escala	Valores Z test Mann Whitney (tamaño del efecto)		
	No limitado vs. no gravemente	No limitado vs. gravemente	No gravemente vs. gravemente
Rasch politómico	105,77 (0,56)	153,31 (0,84)	58,64 (0,59)
Rasch dicotómico	105,86 (0,56)	153,28 (0,84)	57,89 (0,59)
Guttman	80,27 (0,43)	142,60 (0,78)	45,22 (0,46)
Suma dicotómica	105,86 (0,56)	153,28 (0,84)	57,89 (0,59)
Suma politómica	105,75 (0,56)	153,35 (0,84)	59,12 (0,60)

# Estudio 1: Resultados

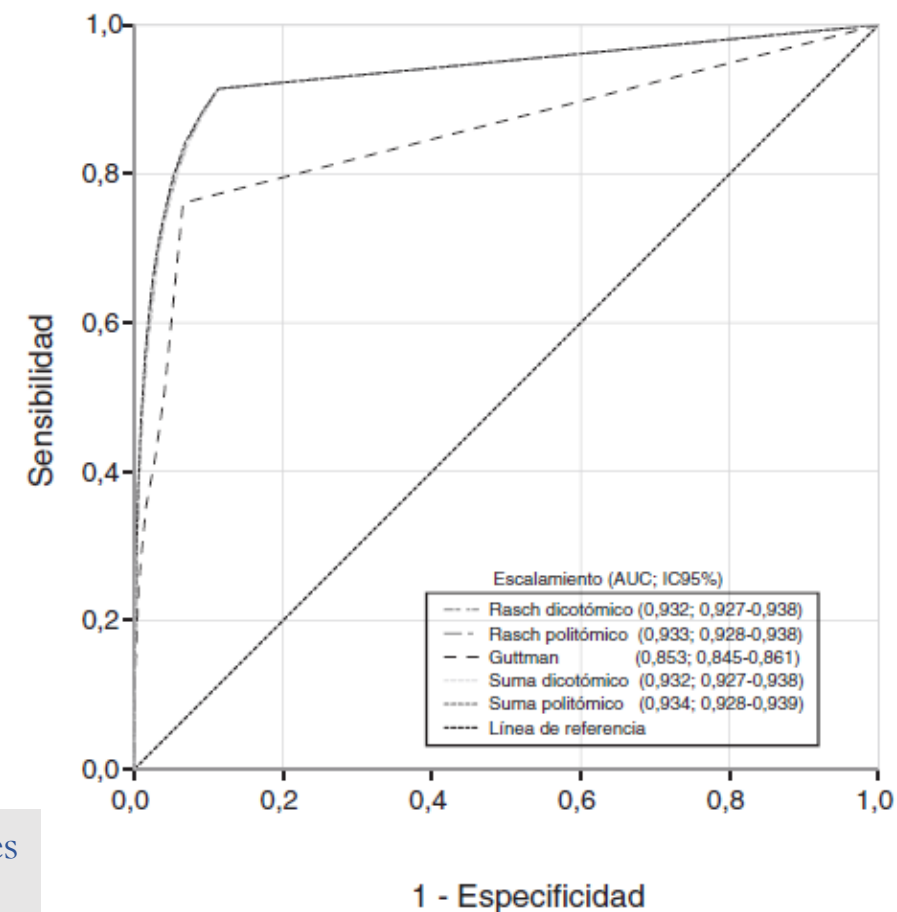
## Poder discriminante

### Precisión relativa

Referencia: escalamiento Rasch politómico

Escala	Precisión relativa (ratio de valores Z) Base: escala Rasch		
	No limitado vs. no gravemente	No limitado vs. gravemente	No gravemente vs. gravemente
Rasch politómico	1,000	1,000	1,000
Rasch dicotómico	1,001	1,000	0,987
Guttman	0,759	0,930	0,771
Suma dicotómica	1,001	1,000	0,987
Suma politómica	1,000	1,000	1,008

### Curva ROC



Vittengl JR et al. Comparative validity of seven scoring systems for the instrumental activities of daily living scale in rural elders. Aging Ment Health. 2006

# Estudio 1: Resultados

Orden de los ítems de menor a mayor carga de discapacidad en distintos métodos de escalamiento

Discapacidad

Menor

Mayor

Rasch dicotómico	Rasch politómico	Guttman
Andar fuera + autobús	Compra + comida + tareas casa	Lavarse + peinarse
Compra + comida + tareas casa	Andar fuera + autobús	Compra + comida + tareas casa
Lavarse + peinarse	Lavarse + peinarse	Andar fuera + autobús
Vestirse	Vestirse	Vestirse
Levantarse + mantenerse	Orinar + defecar	Levantarse + mantenerse
Andar dentro	Levantarse + mantenerse	Andar dentro
Orinar + defecar	Andar dentro	Orinar + defecar
Comer	Comer	Comer

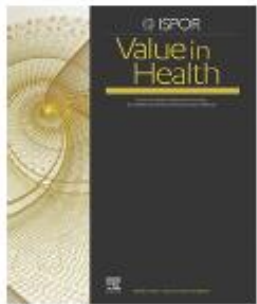
Verbrugge LM et al. Severity, timing, and structure of disability. Soz-Präventivmedizin. 2004

# Estudio 2: Comparación métodos indirectos de mapping

## Objetivo 2:

Comparar las redes bayesianas con otros métodos de mapping para predecir el índice EQ-5D-5L a través de medidas del estado funcional, en pacientes con artrosis de cadera y rodilla.

VALUE HEALTH. 2021; 24(6):874–883

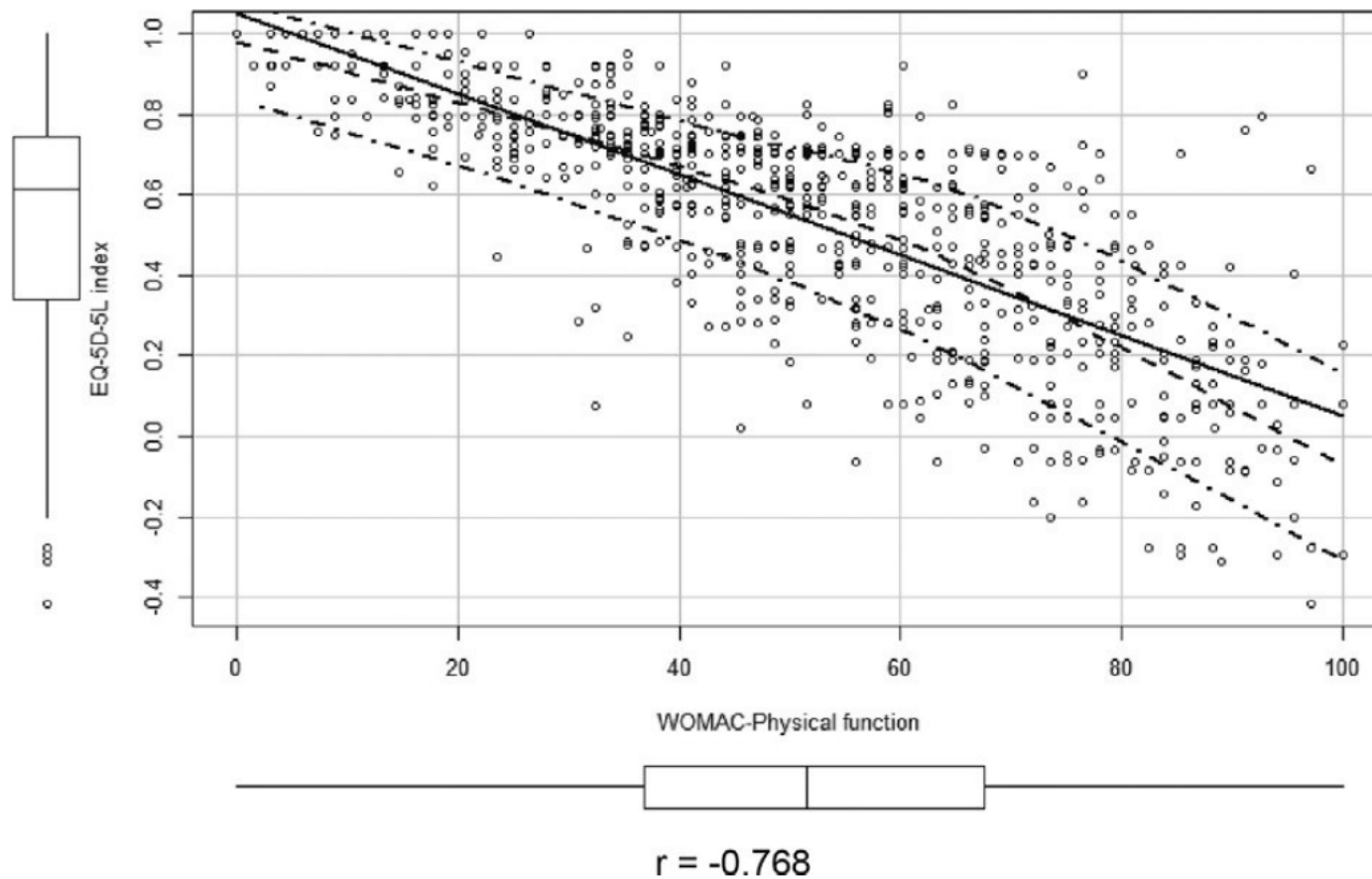


## Response Mapping Methods to Estimate the EQ-5D-5L From the Western Ontario McMaster Universities Osteoarthritis in Patients With Hip or Knee Osteoarthritis

Alba Ayala, MSc, Maria João Forjaz, PhD, Yolanda Ramallo-Fariña, MSc, Jesús Martín-Fernández, MD, PhD, Lidia García-Pérez, MSc, Amaia Bilbao, PhD

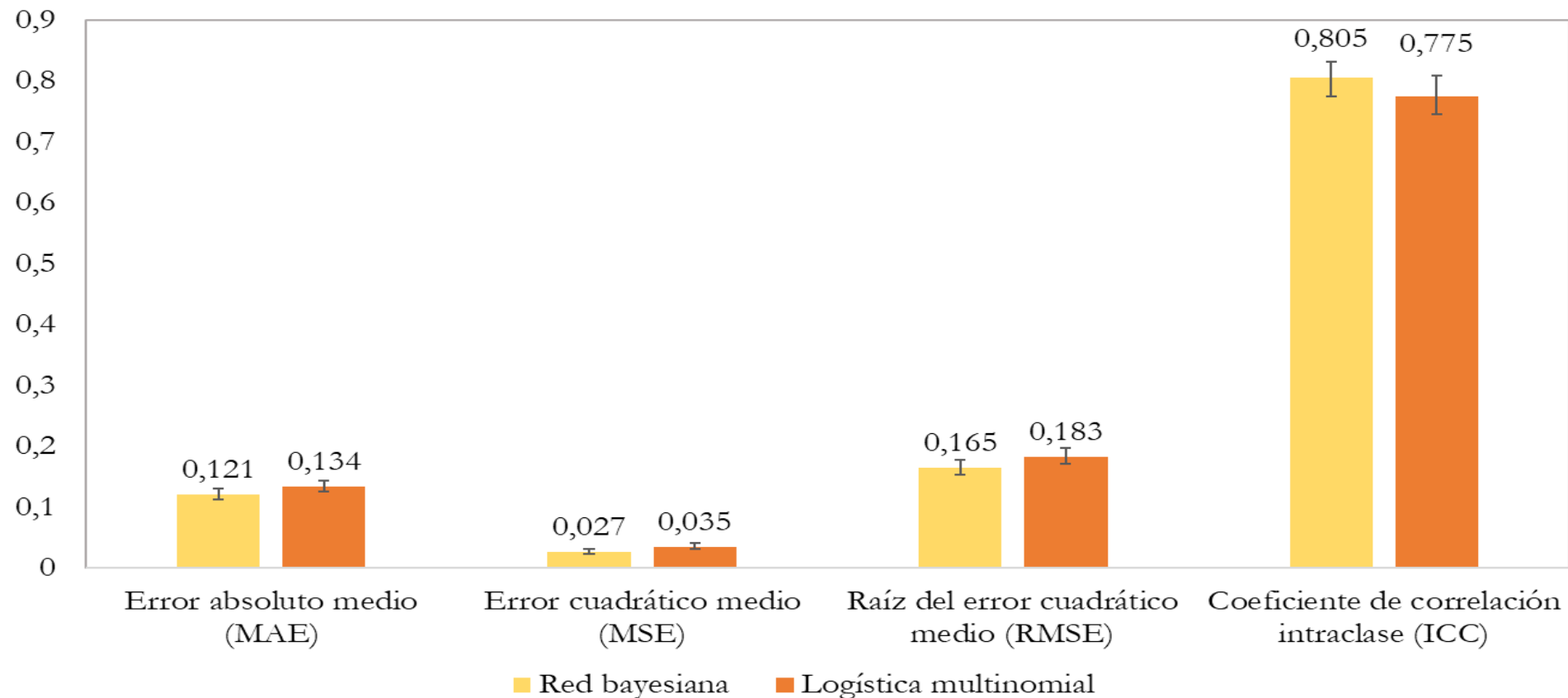
**Factor de impacto (2021): 5.101 (9/88, D1, Health Policy & Services)**

**Instrumento capacidad funcional:** WOMAC (5 opciones de respuesta)  
Dolor (5 ítems), rigidez (2 ítems) y función física (17 ítems).



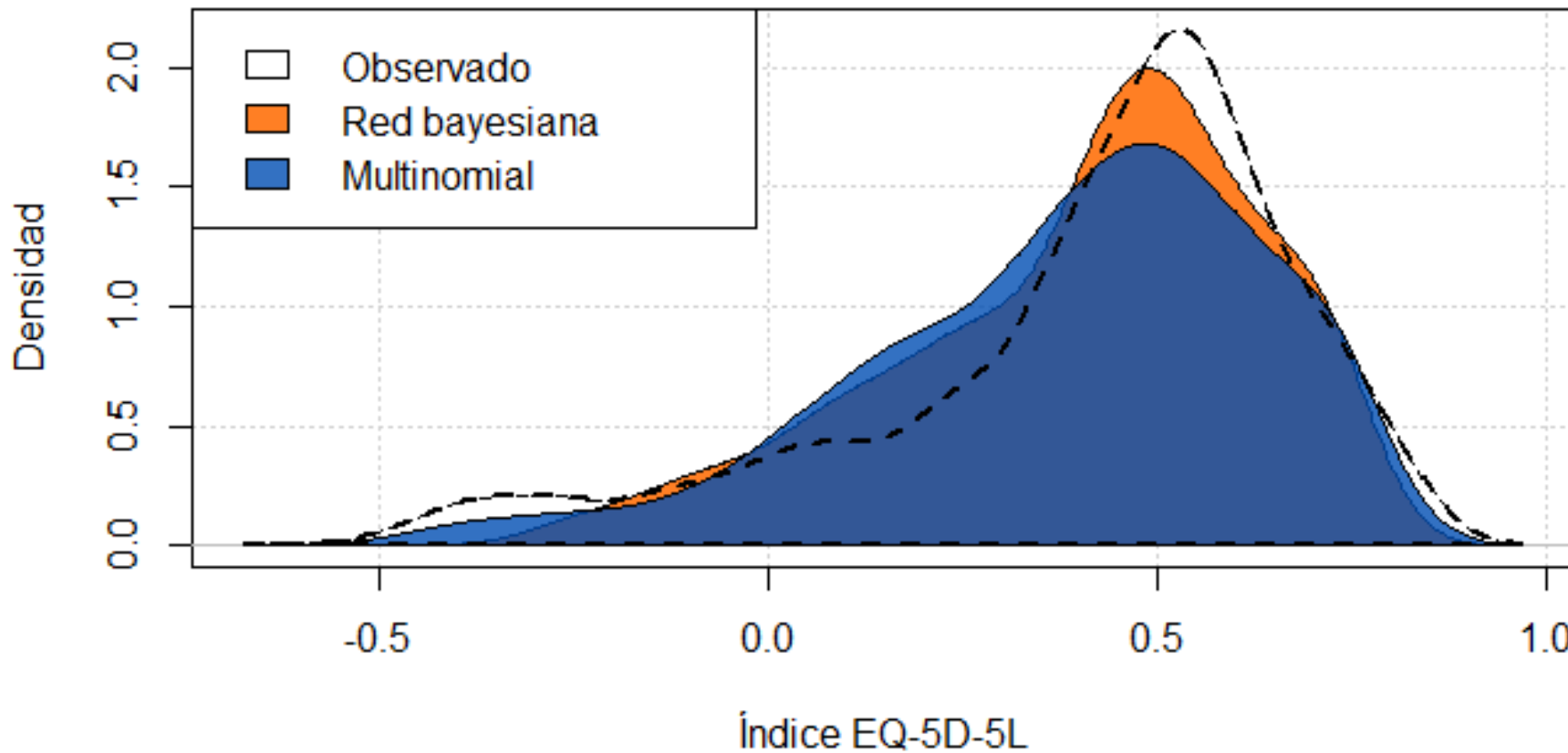
Intervalos de probabilidad- procedimiento bootstrap

Muestra de validación seguimiento (n=663)

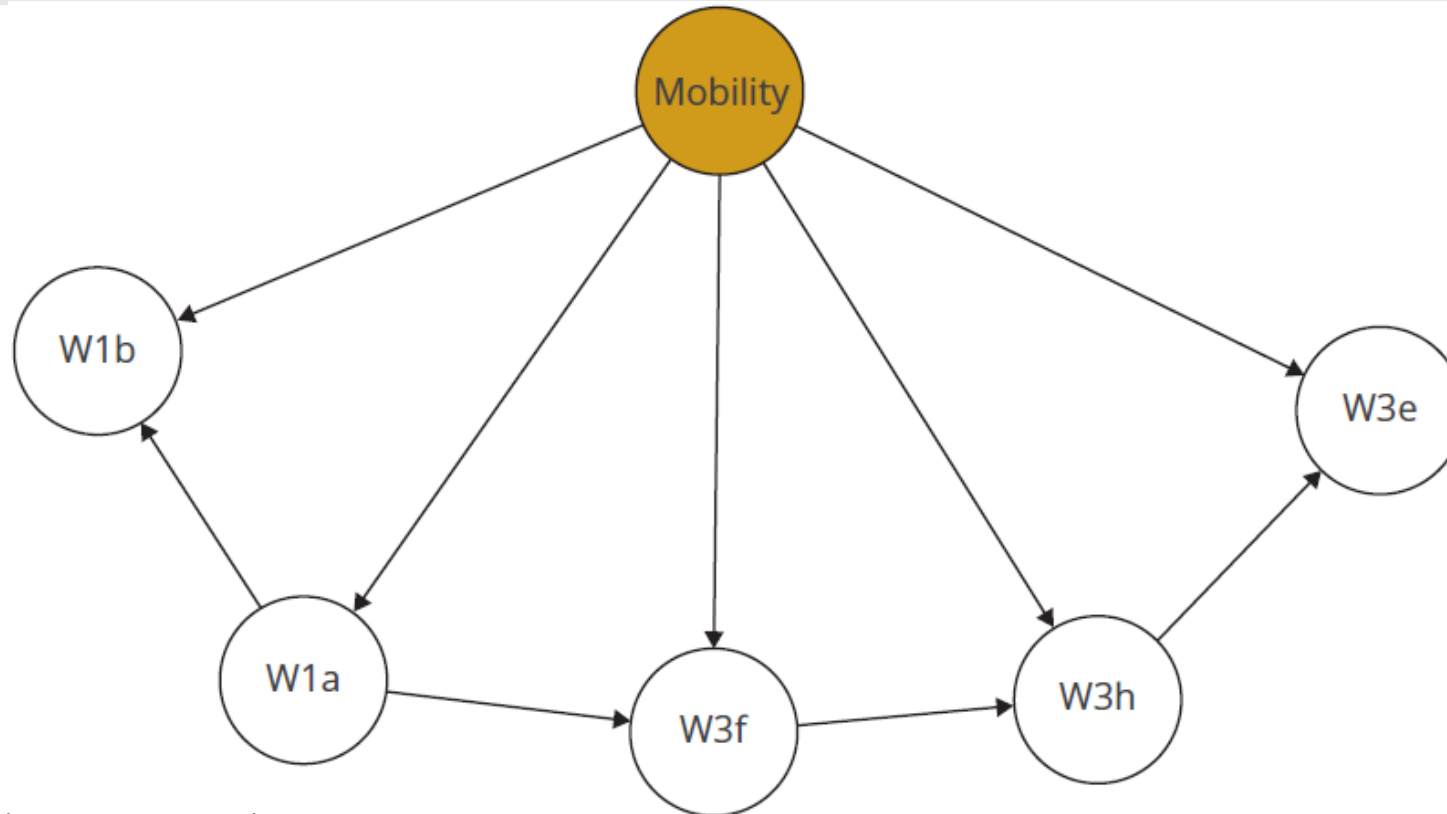


Bilbao A et al. Mapping WOMAC Onto the EQ-5D-5L Utility Index in Patients With Hip or Knee Osteoarthritis. Value in Health, 2020

Muestra de validación seguimiento (n=663)



Kiadaliri AA, Englund M. Assessing the external validity of algorithms to estimate EQ-5D-3L from the WOMAC. Health Qual Life Outcomes, 2016.



¿Cuánto dolor tiene al...?

W1a Andar por un terreno llano

W1b Subir o bajar escaleras

¿Qué grado de dificultad tiene al... ?

W3e Agacharse para coger algo del suelo

W3f Andar por un terreno llano

W3h Ir de compras

# Estudio 2: Resultados

## Red bayesiana



# Estudio 3: Comparación métodos directos e indirectos de mapping

## Objetivo 3:


Comparar las redes bayesianas con otros métodos de predicción del índice EQ-5D-5L, a partir de datos de población mayor de la Encuesta Nacional de Salud en España (ENSE-2012)

Quality of Life Research

<https://doi.org/10.1007/s11136-023-03351-y>

---

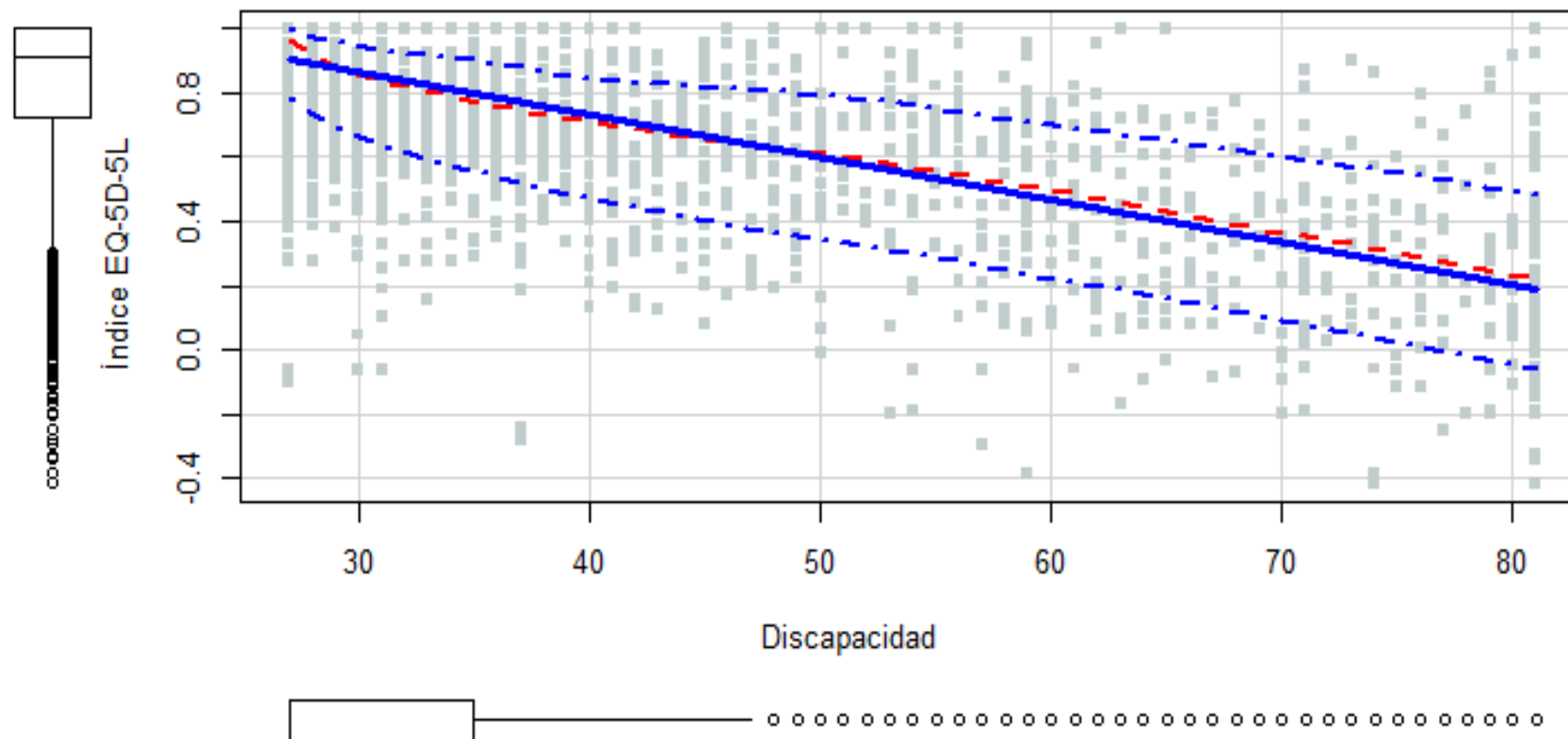
## Mapping the EQ-5D-5L from the Spanish national health survey functional disability scale through Bayesian networks

Alba Ayala<sup>1,6,7</sup>  · Yolanda Ramallo-Fariña<sup>2,6,7</sup> · Amaia Bilbao-Gonzalez<sup>3,4,6,7</sup> · Maria João Forjaz<sup>5,6,7</sup>

Accepted: 13 January 2023

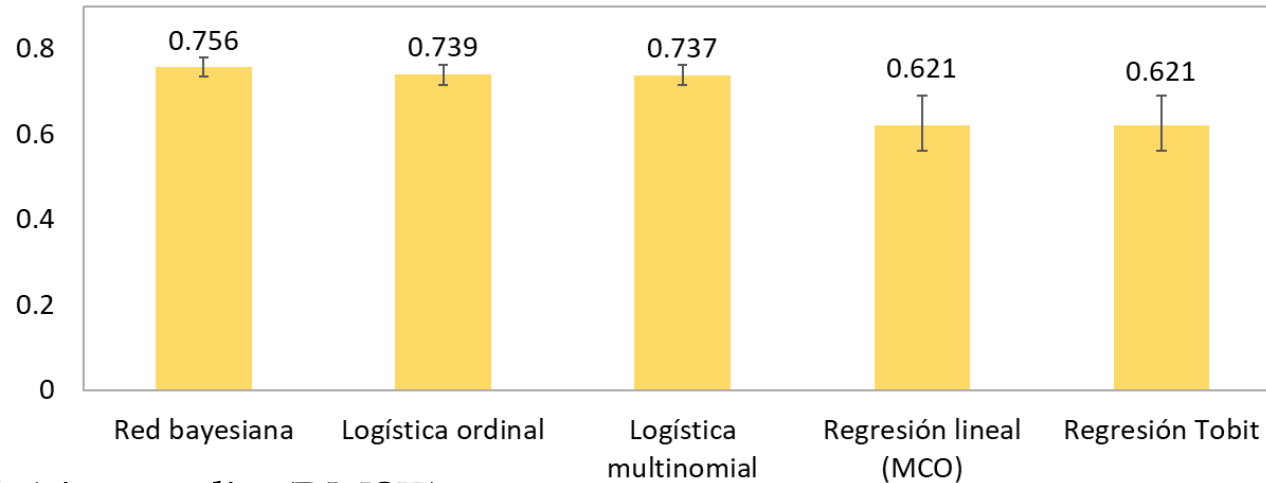
**Factor de impacto (2021): 3.467 (30/159, Q1 Health policy & Services)**

**Instrumento capacidad funcional:** Escala de capacidad funcional de la ENSE  
27 ítems (3 opciones de respuesta)



$$r = -0.534$$

## Coeficiente de correlación intraclase (ICC)



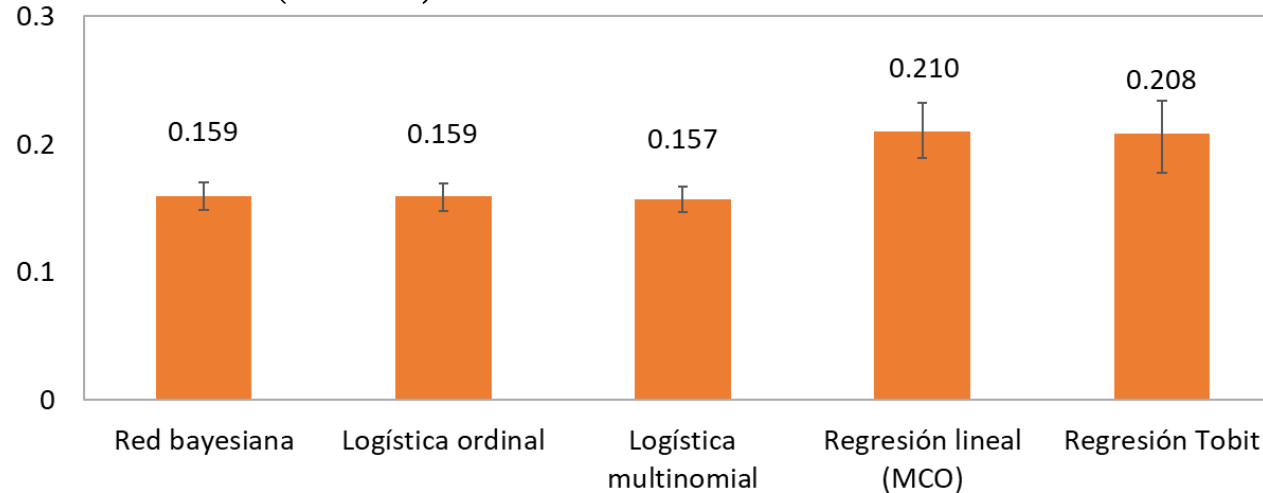
n=5.708

70% muestra de entrenamiento

30% muestra de validación

Intervalos de probabilidad- procedimiento bootstrap

## Error cuadrático medio (RMSE)



Le QA. Probabilistic mapping of the health status measure SF-12 onto the health utility measure EQ-5D using the US-population-based scoring models. Qual Life Res. 2014

# Estudio 3: Resultados

# Red bayesiana

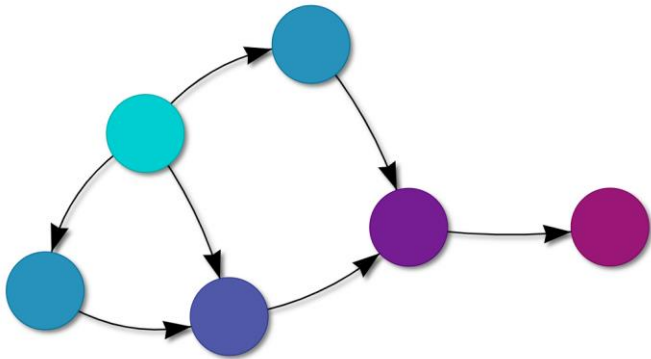


# Discusión

# Discusión



- Se ha confirmado la **hipótesis principal** de que los métodos estadísticos complejos muestran mejores resultados que los más sencillos.
- **Programas** estadísticos actuales tienen capacidad para desarrollar métodos más complejos.
- **Análisis Rasch** y las **redes bayesianas** son más difíciles de aplicar aportan algunas **ventajas** frente al resto de metodologías.



# Implicaciones para práctica clínica y salud pública

- Las medidas de capacidad **funcional** AVD-EDAD2008 pueden ser utilizadas tanto en programas de **intervención** en asistencia **geriátrica** como en la **organización** de servicios de larga duración.
- Las **puntuaciones** directas se pueden **transformar** en puntuaciones Rasch a través de una tabla de **conversión**, que pueden ser utilizada en la **práctica** clínica.



- El orden **jerárquico** de los ítems de la escala de capacidad funcional es de gran **utilidad** en la planificación de **recursos** tanto para los **clínicos** como para los gestores en **salud**.

# Implicaciones para la investigación y salud pública



- Los algoritmos de **mapping** permitirán predecir valores EQ-5D-5L a partir de las **puntuaciones** de la capacidad funcional, y así poder usar el índice de **utilidad** en evaluaciones **económicas**.
- La representación **gráfica** de las redes bayesianas permite **interpretar** las relaciones de **dependencia** entre las dimensiones del instrumento de **CVRS** (EQ-5D-5L) y los ítems de capacidad **funcional**.
- Estas relaciones pueden ser **importantes** para los investigadores y **clínicos** que participan en la **intervención** y el manejo de la **dependencia** y la discapacidad.



# Conclusiones

# Conclusiones



- La medida de capacidad funcional **AVD-EDAD2008** presentó buenas propiedades **psicométricas** con cualquiera de los métodos de **validación** utilizados.
- Utilizar métodos más **complejos** no contribuye necesariamente a la **construcción** de una escala de **capacidad funcional** con mayor poder discriminante, pero sí añaden **ventajas** frente a métodos más sencillos.
- **Métodos indirectos** (*response mapping*) proporcionan **mejores** resultados que los **métodos directos** para estimar el EQ-5D-5L.
- Las **redes bayesianas** son apropiadas para **predecir** el índice **EQ-5D-5L** a través de dos medidas distintas de capacidad funcional.

# Conclusiones



- Las estimaciones obtenidas con las **redes bayesianas** son **robustas** y con mejor validez predictiva que otros métodos de *mapping*.
- La **gráfica** obtenida a través de las redes bayesianas permite interpretar las relaciones de **dependencia** entre las **dimensiones** del instrumento de CVRS (**EQ-5D-5L**) y los ítems de capacidad funcional.
- Se han desarrollado **algoritmos** de *mapping* que permitirán predecir el índice EQ-5D-5L e incluir esta **medida de utilidad** en estudios de **evaluación** económica y clínica que estimen el coste de los **años de vida** ajustados a **calidad**.