

Perfiles poblacionales de enfermedades crónicas no transmisibles, hábitos, características sociodemográficas, vivienda y entorno residencial en Chile: análisis factorial de datos mixtos y conglomerados por K-Prototypes en la ENCAVI 2023-2024

Population profiles of noncommunicable diseases, health behaviours, sociodemographic characteristics, housing and residential environment in Chile: factor analysis of mixed data and K-Prototypes clustering in the 2023-2024 ENCAVI

Carlos González-Torres¹

¹ Programa de Doctorado en Ciencias e Ingeniería para la Salud, Universidad de Valparaíso, Valparaíso, Chile. ORCID 0000-0002-7765-6388. carlos.gonzaleztp@postgrado.uv.cl. Teléfono: 9 8640 5786



Fecha Recepción: 19 de septiembre de 2026

Fecha Revisión: 19 de octubre de 2026

Fecha Aceptación: 20 de octubre de 2026

DOI: <https://doi.org/10.22370/syc.2.1.2026.6077>

Resumen

Introducción. El enfoque centrado en variables estima el aporte promedio de cada factor a un desenlace, pero no identifica subpoblaciones con combinaciones características de exposiciones y enfermedad.

Objetivo. Identificar y caracterizar perfiles poblacionales definidos simultáneamente por enfermedades crónicas no transmisibles (ENT), hábitos, estado nutricional, características sociodemográficas, vivienda y entorno residencial en Chile, y evaluar si la vivienda y el entorno residencial definen un perfil propio.

Métodos. Estudio transversal de fuente secundaria con la Encuesta Nacional de Calidad de Vida y Salud (ENCAVI) 2023-2024. Se analizaron 28 variables, 5 continuas y 23 categóricas, en 10.677 casos completos, mediante análisis factorial de datos mixtos (FAMD) y conglomerados por K-Prototypes. El número de conglomerados se eligió según la función de costo, la silueta, los índices de Calinski-Harabasz y Davies-Bouldin, y la estabilidad (índice de Rand ajustado). Las diferencias se evaluaron con chi-cuadrado, ANOVA o Kruskal-Wallis, y tamaños de efecto con intervalos de confianza bootstrap (2.000 remuestreos).

Resultados. El 38,6% reportó dos o más ENT. Frente a la multimorbilidad, los mayores tamaños de efecto correspondieron a situación ocupacional, edad y nivel educacional; vivienda y entorno mostraron efectos pequeños. Se retuvieron seis conglomerados: un perfil joven con ventaja socioeconómica, dos perfiles de vejez diferenciados por duración del sueño y carga de dolor, uno de precariedad habitacional y ambiental urbana, uno de obesidad y uno de actividad física declarada muy alta.

Conclusiones. Variables de vivienda y entorno con efectos pequeños sobre la multimorbilidad definieron un conglomerado propio, lo que muestra la complementariedad de ambos enfoques.

Palabras clave: Multimorbilidad; Enfermedades no transmisibles; Análisis por conglomerados; Vivienda; Chile.

Abstract

Introduction. Variable-centred approaches estimate the average contribution of each factor to an outcome, but do not identify subpopulations with characteristic combinations of exposures and disease.

Objective. To identify and characterise population profiles defined simultaneously by noncommunicable diseases (NCDs), health behaviours, nutritional status, sociodemographic characteristics, housing and residential environment in Chile, and to assess whether housing and residential environment define a profile of their own.

Methods. Cross-sectional secondary analysis of the 2023-2024 National Quality of Life and Health Survey (ENCAVI). Twenty-eight variables, 5 continuous and 23 categorical, were analysed in 10,677 complete cases. Factor analysis of mixed data (FAMD) and K-Prototypes clustering were applied. The number of clusters was chosen using the cost function, silhouette, Calinski-Harabasz and Davies-Bouldin indices, and stability (adjusted Rand index). Differences were assessed with chi-square, ANOVA or Kruskal-Wallis tests, and effect sizes with bootstrap confidence intervals (2,000 resamples).

Results. Two or more NCDs were reported by 38.6%. Against multimorbidity, the largest effect sizes were occupational status, age and educational level; housing and environment variables showed small effects. Six clusters were retained, separating a young socioeconomically advantaged profile, two old-age profiles differing in sleep duration and pain burden, an urban housing and environmental deprivation profile, an obesity profile and a very high self-reported physical activity profile.

Conclusions. Housing and environment variables with small effects on multimorbidity nonetheless defined a cluster of their own, showing the complementarity of both approaches.

Keywords: Multimorbidity; Noncommunicable Diseases; Cluster Analysis; Housing; Chile.

INTRODUCCIÓN

Este estudio examina si la vivienda y el entorno residencial, habitualmente analizados como exposiciones aisladas, caracterizan subpoblaciones identificables en la Encuesta Nacional de Calidad de Vida y Salud (ENCAVI) 2023-2024, junto con las enfermedades crónicas no transmisibles (ENT) y los hábitos. Las condiciones que requieren cuidado crónico concentran cerca de dos tercios de los años de vida ajustados por discapacidad a nivel global⁽¹⁾. En Chile, la Encuesta Nacional de Salud 2016-2017 situó la prevalencia de multimorbilidad en 54,9%⁽²⁾, y el patrón de agrupación de esas

condiciones se asocia de manera diferencial con la mortalidad⁽³⁾.

La investigación epidemiológica sobre este problema se organiza mayoritariamente en torno a variables: se estima el aporte promedio de cada exposición a un desenlace, con las demás mantenidas constantes. Ese marco responde bien a preguntas de atribución, pero no identifica subpoblaciones, porque una exposición con un tamaño de efecto pequeño en el conjunto de la población puede ser la característica que define a un grupo minoritario. El enfoque centrado en la persona invierte la unidad de análisis y busca

precisamente esas configuraciones. En Chile se ha aplicado mediante clases latentes sobre condiciones crónicas⁽³⁾ y mediante conglomerados sobre factores de riesgo de demencia⁽⁴⁾; en ambos casos el conjunto de variables fue homogéneo en su escala de medición.

Esa homogeneidad es una restricción práctica. Los determinantes sociales y materiales de la salud se miden de forma mixta: edad e índice de masa corporal son continuos, mientras que educación, tipo de combustible de calefacción o diagnóstico autorreportado son categóricos. Convertir todo a categorías descarta información y convertir todo a escalas numéricas impone distancias arbitrarias entre categorías. El análisis factorial de datos mixtos (FAMD) resuelve la representación, al tratar las variables continuas como en un análisis de componentes principales y las categóricas como en un análisis de correspondencias múltiples, con ponderaciones que equilibran ambos bloques⁽⁵⁾. En términos prácticos, resume en pocos ejes la información de variables de distinto tipo. El algoritmo K-Prototypes resuelve la partición, al combinar la distancia euclidiana para el bloque continuo con una medida de discordancia para el categórico, ponderadas por un parámetro que fija el peso relativo de cada bloque⁽⁶⁾. En términos prácticos, forma grupos internamente semejantes, usando promedios para las variables numéricas y la categoría más frecuente para las cualitativas. Esta combinación se ha utilizado recientemente para identificar subgrupos clínicos en datos de mundo real^(7,8), pero ha tenido poca aplicación en encuestas poblacionales de salud latinoamericanas.

La relación entre vivienda y salud en Chile tiene una larga trayectoria. En La realidad médico-social chilena (1939), Salvador Allende, entonces ministro de Salubridad, vinculó la situación sanitaria de los asalariados con sus condiciones de vida^(9,10), analizó la relación de la vivienda con la mortalidad y la morbilidad⁽⁹⁾, y propuso un programa nacional de vivienda⁽¹⁰⁾.

El objetivo primario fue identificar y caracterizar perfiles poblacionales definidos simultáneamente por ENT autorreportadas, hábitos, estado nutricional, características sociodemográficas, condiciones de la vivienda y entorno residencial

percibido en población chilena. Los objetivos secundarios fueron describir la asociación de esas variables con el número de ENT (multimorbilidad) y evaluar si la vivienda y el entorno definen un perfil propio. La hipótesis de trabajo, derivada del marco conceptual y no registrada previamente, fue que estas últimas variables podrían constituir el rasgo distintivo de un subgrupo aun cuando su asociación promedio con la multimorbilidad fuese pequeña.

MATERIAL Y MÉTODO

Diseño y fuente de datos

Estudio transversal de fuente secundaria con la base pública de la ENCAVI 2023-2024 del Ministerio de Salud de Chile, encuesta poblacional con muestreo probabilístico, estratificado y por conglomerados, dirigida a personas de 15 años o más residentes en viviendas particulares del territorio nacional⁽¹¹⁾. La ENCAVI explora la calidad de vida relacionada con la salud y sus determinantes, incluidos estilos de vida y entorno⁽¹¹⁾. El reporte sigue la guía STROBE para estudios transversales⁽¹²⁾.

Variables

Se seleccionaron 28 variables organizadas en cinco dominios. Cinco se mantuvieron en escala continua: edad en años; índice de masa corporal calculado a partir de peso y talla autorreportados; horas de sueño de lunes a viernes; razón de personas por dormitorio de uso exclusivo; y actividad física en MET-minuto por semana, calculada según el protocolo del Cuestionario Internacional de Actividad Física, con truncamiento de cada dominio a 180 minutos diarios. Las 23 restantes son categóricas: sexo, zona, macrozona, nivel educacional, quintil de ingreso del hogar, situación ocupacional y previsión de salud; material de muros y de piso, dicotomizados en aceptable o deficiente según la clasificación del instrumento, y combustible principal de calefacción; presencia autorreportada de contaminación del aire, ruidos molestos y microbasurales o malos olores en el barrio, e inseguridad al caminar solo de noche; consumo de tabaco y consumo de alcohol en

tres niveles derivados del AUDIT; y diagnóstico médico autorreportado de hipertensión arterial, diabetes mellitus, enfermedad respiratoria crónica, depresión o ansiedad, artrosis o artritis, dolor crónico y enfermedad cardiovascular. La multimorbilidad, definida como el número de estas siete condiciones en tres categorías, se usó como variable de interés para las comparaciones de la [Tabla 2](#). Las respuestas "no sabe" y "no responde" se trataron como datos faltantes y el análisis se restringió a los casos completos.

Análisis

Las variables continuas se estandarizaron antes del agrupamiento. Se ajustó un FAMD con diez dimensiones para obtener una representación euclidiana común de las 28 variables. La partición se obtuvo con K-Prototypes, con inicialización de Huang y cinco inicios aleatorios por solución, para valores de K entre 2 y 8. La elección de K combinó cinco criterios: la función de costo del algoritmo, el coeficiente de silueta (si cada persona está más cerca de su propio conglomerado que del más próximo), los índices de Calinski-Harabasz y Davies-Bouldin sobre las cinco primeras dimensiones del FAMD, y la estabilidad de la partición, estimada con el índice de Rand ajustado entre la solución completa y doce soluciones obtenidas en submuestras aleatorias del 80%⁽¹³⁾.

Las diferencias de cada variable según multimorbilidad y según conglomerado se evaluaron con la prueba de chi-cuadrado y V de Cramér para las categóricas (magnitud de la asociación). Para las continuas se verificó la asimetría y la curtosis dentro de cada grupo: cuando ambas se mantuvieron por debajo de 1 en valor absoluto se aplicó ANOVA de una vía con eta cuadrado, y en caso contrario la prueba de Kruskal-Wallis con epsilon cuadrado. Todos los tamaños de efecto se acompañan de un intervalo de confianza del 95% bootstrap percentil basado en 2.000 remuestreos. El análisis se realizó sin ponderar, de modo que los resultados describen la muestra analítica y no constituyen estimaciones poblacionales. Los cálculos se ejecutaron en Python 3.10 con pandas, NumPy, scikit-learn,

prince y kmodes. La Tabla suplementaria S1 define estos indicadores en lenguaje no técnico.

Aspectos éticos

Se trabajó con una base pública y anonimizada, sin variables identificatorias, por lo que no se requirió aprobación de un comité de ética. El estudio respetó los principios de la Declaración de Helsinki.

RESULTADOS

De los 16.590 registros, 10.677 (64,4%) presentaron información completa en las 28 variables. El índice de masa corporal y las horas de sueño concentraron la mayor proporción de datos faltantes. La [Tabla 1](#) presenta la distribución de todas las variables en la muestra analítica: el 67,1% fueron mujeres y la edad media fue 52,1 años (DE 17,6). El 37,0% calefaccionaba su vivienda con biomasa y el 47,2% percibió inseguridad al caminar solo de noche. La hipertensión arterial fue la condición más frecuente (35,7%), seguida de artrosis o artritis (22,6%), depresión o ansiedad y dolor crónico (20,7% cada una) y diabetes mellitus (19,7%). El 38,6% reportó dos o más condiciones y el 36,7% ninguna.

Diferencias según multimorbilidad

Todas las variables salvo el material de muros y de piso difirieron según el número de condiciones crónicas ([Tabla 2](#)). Los mayores tamaños de efecto correspondieron a la situación ocupacional ($V = 0,253$; IC 95% 0,235 a 0,271), la edad ($\eta^2 = 0,234$; 0,219 a 0,248) y el nivel educacional ($V = 0,215$; 0,204 a 0,228). Las variables de entorno mostraron efectos pequeños pero consistentes en su dirección: la inseguridad percibida alcanzó 51,9% entre quienes reportaron dos o más condiciones frente a 42,7% entre quienes no reportaron ninguna ($V = 0,080$; 0,062 a 0,098), y los microbasurales o malos olores 49,9% frente a 41,4% ($V = 0,074$; 0,057 a 0,093).

Selección del número de conglomerados

Las dos primeras dimensiones del FAMD explicaron 8,6% y 5,8% de la varianza. La silueta, el índice de Calinski-Harabasz y el de Davies-Bouldin fueron óptimos en $K = 2$, solución que además resultó muy estable (índice de Rand ajustado medio 0,949; DE 0,055). La función de costo mostró su cambio de pendiente entre $K = 5$ y $K = 6$, y la estabilidad presentó un máximo secundario en $K = 6$ (0,821; DE 0,056; mínimo 0,734) con una caída marcada en $K = 7$ (0,666) ([Figura suplementaria S1](#)). Se retuvo la solución de seis conglomerados, porque la de dos reproduce un gradiente de edad y carga de enfermedad ya visible en la Tabla 1 y no aporta diferenciación adicional.

Perfiles identificados

La Figura 1 muestra la proyección de los participantes en el plano factorial y la desviación estandarizada de cada variable respecto de la muestra completa. El conglomerado C1 ($n = 2.872$; 26,9%) reunió a los más jóvenes (media 36,3 años), con educación superior en 52,9%, ocupados en 65,0% y con la menor carga de enfermedad (12,0% con dos o más condiciones). C2 ($n = 2.273$; 21,3%) y C3 ($n = 2.233$; 20,9%) agruparon a personas mayores (68,0 y 65,3 años) con alta prevalencia de hipertensión (59,0% y 58,3%), pero se diferenciaron por la duración del sueño, mediana de 8,0 horas en C2 y de 5,0 horas en C3, y por la carga de dolor crónico (22,3% frente a 30,1%) y de depresión o ansiedad (18,0% frente a 27,5%). C4 ($n = 1.305$; 12,2%) se definió por el hacinamiento (mediana de 2,0 personas por dormitorio), la concentración en el quintil I (78,9%) y en la macrozona Metropolitana (35,2%), y el deterioro del entorno percibido, con microbasurales en 60,0% e inseguridad en 61,8%; su edad media fue 42,7 años y el 27,2% reportó dos o más condiciones. C5 ($n = 1.245$; 11,7%) correspondió a un perfil de obesidad, con índice de masa corporal mediano de 37,0 kg/m², 78,9% de mujeres y 28,0% de diabetes. C6 ($n = 749$; 7,0%) fue el único con predominio masculino (60,5%) y se caracterizó por una actividad física declarada muy alta (mediana 9.360 MET-minuto por semana) junto con la mayor proporción de consumo riesgoso de alcohol (7,6%).

Todas las variables difirieron entre conglomerados ([Tabla 3](#)). Los mayores tamaños de efecto correspondieron a la edad ($\eta^2 = 0,555$; 0,544 a 0,566), la hipertensión ($V = 0,455$; 0,440 a 0,471), las horas de sueño ($\epsilon^2 = 0,391$; 0,380 a 0,403), la situación ocupacional ($V = 0,380$; 0,364 a 0,397) y las personas por dormitorio ($\epsilon^2 = 0,328$; 0,314 a 0,342). La proporción con dos o más condiciones varió entre 12,0% en C1 y 62,8% en C3.

DISCUSIÓN

En 10.677 personas de 15 años o más en Chile, un análisis conjunto de 28 variables mixtas identificó seis perfiles que se organizan en torno a tres ejes: la edad con su carga acumulada de enfermedad, la adiposidad, y la precariedad habitacional y ambiental. Los dos primeros reproducen agrupaciones ya descritas en el país mediante clases latentes y conglomerados^(3,4). El tercero constituye el aporte de este trabajo.

El contraste entre las [Tablas 2](#) y [3](#) es el resultado central. Frente a la multimorbilidad, el hacinamiento, los materiales de la vivienda y el entorno percibido mostraron tamaños de efecto pequeños, en todos los casos con un V inferior a 0,09, lo que en un análisis centrado en variables llevaría a considerarlos poco relevantes. Sin embargo, esas mismas variables definieron íntegramente a C4, un grupo que reúne al 12,2% de la muestra. La explicación es que su distribución no es gradual sino concentrada: la precariedad no aumenta de manera continua con el número de condiciones crónicas, sino que se acumula en un subgrupo joven, urbano y del quintil más bajo. Los dos enfoques no compiten, describen aspectos distintos de la misma estructura, y el enfoque centrado en la persona recupera configuraciones que el promedio poblacional diluye.

C4 merece atención por lo que no muestra. Su proporción con dos o más condiciones crónicas, 27,2%, es inferior a la de la muestra completa, algo esperable dada su edad media de 42,7 años. Si la exposición a condiciones materiales y ambientales adversas opera de manera acumulativa a lo largo del curso de vida, como sugieren los estudios longitudinales de calidad de la vivienda⁽¹⁴⁾ y las

revisiones sobre hacinamiento y ventilación deficiente⁽¹⁵⁾, este perfil identificaría a una población todavía sin la carga de enfermedad que su exposición haría prever. Un diseño transversal no permite verificar esa hipótesis, pero sí delimita el grupo en el que convendría probarla. Este perfil recuerda que la vivienda de los sectores de menores ingresos fue identificada tempranamente como problema sanitario en Chile^(9,10); un corte transversal no reconstruye cómo se configuró, lo que requiere el diálogo con la historia y las ciencias sociales.

La separación entre C2 y C3 también resulta informativa. Ambos agrupan a personas mayores con prevalencias casi idénticas de hipertensión, pero difieren en tres horas de sueño mediano y en una carga sustancialmente mayor de dolor crónico y de depresión o ansiedad en C3. El hallazgo sugiere que la duración del sueño discrimina dentro de la población mayor mejor que la mayoría de las variables sociodemográficas.

C6 debe interpretarse con cautela. Una mediana de 9.360 MET-minuto por semana equivale a casi tres horas diarias de actividad vigorosa declarada, cifra inverosímil que refleja la conocida tendencia del Cuestionario Internacional de Actividad Física al sobrerreporte; el truncamiento a 180 minutos diarios por dominio lo limita, pero no lo elimina.

Este conglomerado probablemente identifica un patrón de respuesta más que de conducta.

Las limitaciones son relevantes. El diseño transversal impide establecer secuencia temporal. Todas las variables provienen de autorreporte, incluidos peso y talla, y el diagnóstico autorreportado depende del acceso a servicios de salud, de modo que la prevalencia puede subestimarse en los grupos con menor cobertura. El análisis por casos completos excluyó al 35,6% de los registros. El registro de sueño de fin de semana se descartó porque, sin dato de días hábiles, se concentraba en valores pares de 8 a 16 horas, incompatibles con el resto de la muestra. Los índices internos favorecieron dos grupos y la silueta de la solución retenida fue cercana a cero: los conglomerados particionan una nube continua, y su justificación descansa en la estabilidad y la interpretabilidad, no en la separación geométrica. Por último, la ausencia de ponderación limita la interpretación a la muestra analítica.

Estos perfiles podrían orientar la segmentación de acciones de salud pública. La verificación de su valor pronóstico requiere seguimiento longitudinal, particularmente en el grupo definido por precariedad habitacional y ambiental.

Declaraciones

Financiamiento: Sin financiamiento

Conflictos de interés: El autor declara no tener conflictos de interés

Agradecimientos: El autor agradece al Departamento de Epidemiología del Ministerio de Salud de Chile por poner a disposición pública los datos de la Encuesta Nacional de Calidad de Vida y Salud (ENCAVI), los cuales hicieron posible la realización de este estudio. Asimismo, se reconoce el trabajo y compromiso de los equipos técnicos y profesionales involucrados en el diseño, levantamiento, procesamiento y difusión de la encuesta, así como la valiosa participación de las personas encuestadas que contribuyeron con su tiempo e información.

Declaración sobre la IA generativa y las tecnologías asistidas por IA en el proceso de redacción: Durante la elaboración de este trabajo, el autor de correspondencia utilizó Microsoft Copilot, una herramienta de inteligencia artificial generativa desarrollada por Microsoft, para apoyar la revisión gramatical y mejorar la legibilidad del texto. Todo el contenido científico y las interpretaciones fueron elaborados exclusivamente por el autor.

Número de tablas: 3 (más una tabla suplementaria). **Número de figuras:** 1 (más una figura suplementaria).

Número de palabras (Introducción a Discusión): 2.486.

Tabla 1. Estadísticas descriptivas de las 28 variables analizadas y de la multimorbilidad. Encuesta Nacional de Calidad de Vida y Salud 2023-2024, muestra analítica de casos completos (n = 10.677). Las variables continuas se presentan como media (desviación estándar) y mediana [percentil 25; percentil 75]; las categóricas como n (%) sin ponderar.

Variable	Estadístico o categoría	Muestra analítica (n = 10677)
Edad, años	Media (DE)	52,1 (17,6)
	Mediana [P25; P75]	54,0 [38,0; 66,0]
Índice de masa corporal, kg/m ²	Media (DE)	28,4 (5,3)
	Mediana [P25; P75]	27,7 [24,8; 31,2]
Horas de sueño por día	Media (DE)	6,7 (1,6)
	Mediana [P25; P75]	7,0 [6,0; 8,0]
Personas por dormitorio	Media (DE)	1,1 (0,6)
	Mediana [P25; P75]	1,0 [0,7; 1,3]
Actividad física, MET-minuto por semana	Media (DE)	2 176,6 (2 877,2)
	Mediana [P25; P75]	1 188,0 [240,0; 2 994,0]
Sexo	Mujer	7163 (67,1)
	Hombre	3514 (32,9)
Zona de residencia	Urbano	9777 (91,6)
	Rural	900 (8,4)
Macrozona	Norte	1688 (15,8)
	Centro	2207 (20,7)
	Metropolitana	2533 (23,7)
	Sur	4249 (39,8)
Nivel educacional	Básica	2586 (24,2)
	Media incompleta	1753 (16,4)
	Media completa	3275 (30,7)
	Superior	3063 (28,7)
Quintil de ingreso del hogar	I	6274 (58,8)
	II	1716 (16,1)
	III	1392 (13,0)
	IV-V	1295 (12,1)
Situación ocupacional	Trabaja	4493 (42,1)
	No trabaja	6184 (57,9)
Previsión de salud	FONASA A-B	7467 (69,9)
	FONASA C-D	2028 (19,0)

Variable	Estadístico o categoría	Muestra analítica (n = 10677)
	ISAPRE u otro	1182 (11,1)
Material de muros exteriores	Aceptables	10247 (96,0)
	Deficientes	430 (4,0)
Material del piso	Aceptable	10432 (97,7)
	Deficiente	245 (2,3)
Combustible de calefacción	Biomasa	3954 (37,0)
	Gas o electricidad	5282 (49,5)
	Sin calefacción	1441 (13,5)
Contaminación del aire en el barrio	Sí	2898 (27,1)
Ruidos molestos en el barrio	Sí	3990 (37,4)
Microbasurales o malos olores	Sí	4857 (45,5)
Inseguridad al caminar de noche	Sí	5044 (47,2)
Consumo de tabaco	Nunca fumador	6157 (57,7)
	Exfumador	1825 (17,1)
	Fumador actual	2695 (25,2)
Consumo de alcohol	Sin consumo en 12 meses	4452 (41,7)
	Consumo sin riesgo	5772 (54,1)
	Consumo riesgoso	453 (4,2)
Hipertensión arterial	Sí	3809 (35,7)
Diabetes mellitus	Sí	2103 (19,7)
Enfermedad respiratoria crónica	Sí	808 (7,6)
Depresión o ansiedad	Sí	2213 (20,7)
Artrosis o artritis	Sí	2418 (22,6)
Dolor crónico	Sí	2210 (20,7)
Enfermedad cardiovascular	Sí	1166 (10,9)
Multimorbilidad	Cero	3922 (36,7)
	Una	2638 (24,7)
	Dos o más	4117 (38,6)

Los valores corresponden a n (%) sin ponderar, calculados sobre la muestra analítica de casos completos. DE: desviación estándar. P25 y P75: percentiles 25 y 75. MET: equivalente metabólico. IPAQ: Cuestionario Internacional de Actividad Física, con truncamiento de cada dominio a 180 minutos diarios. AUDIT: Alcohol Use Disorders Identification Test. FONASA: Fondo Nacional de Salud. Muros y piso deficientes: internit, adobe o desechos y radier o tierra, respectivamente. Biomasa: leña, carbón o pellet. Multimorbilidad: número de las siete condiciones crónicas autorreportadas.

Tabla 2. Diferencias de las variables según número de enfermedades crónicas no transmisibles autorreportadas, con tamaño de efecto e intervalo de confianza del 95%. Las variables continuas se resumen como media (DE) cuando se aplicó ANOVA con eta cuadrado y como mediana [P25; P75] cuando se aplicó Kruskal-Wallis con epsilon cuadrado. V corresponde a la V de Cramér. Los intervalos provienen de 2.000 remuestreos bootstrap percentil. Las siete condiciones que componen la variable de agrupación se presentan sin prueba estadística por tratarse de una relación definicional.

Variable	Categoría	Total (n = 10677)	Cero (n = 3922)	Una (n = 2638)	Dos o más (n = 4117)	p	Tamaño de efecto (IC 95%)
Edad, años		52,1 (17,6)	42,4 (15,6)	51,2 (17,0)	61,9 (14,1)	< 0,001	$\eta^2 = 0,234$ (0,219 a 0,248)
Índice de masa corporal, kg/m ²		27,7 [24,8; 31,2]	26,7 [24,1; 29,8]	27,5 [24,9; 31,1]	28,8 [25,6; 32,8]	< 0,001	$\epsilon^2 = 0,034$ (0,027 a 0,040)
Horas de sueño por día		6,7 (1,6)	6,9 (1,4)	6,7 (1,5)	6,5 (1,8)	< 0,001	$\eta^2 = 0,008$ (0,005 a 0,012)
Personas por dormitorio		1,0 [0,7; 1,3]	1,0 [0,7; 1,5]	1,0 [0,7; 1,3]	1,0 [0,7; 1,3]	< 0,001	$\epsilon^2 = 0,013$ (0,009 a 0,017)
Actividad física, MET-minuto por semana		1 188,0 [240,0; 2 994,0]	1 485,0 [462,0; 3 939,0]	1 188,0 [297,0; 2 970,0]	720,0 [99,0; 2 337,0]	< 0,001	$\epsilon^2 = 0,029$ (0,023 a 0,036)
Sexo	Mujer	7163 (67,1)	2389 (60,9)	1722 (65,3)	3052 (74,1)	< 0,001	V = 0,124 (0,106 a 0,143)
	Hombre	3514 (32,9)	1533 (39,1)	916 (34,7)	1065 (25,9)		
Zona de residencia	Urbano	9777 (91,6)	3550 (90,5)	2436 (92,3)	3791 (92,1)	0,011	V = 0,029 (0,013 a 0,049)
	Rural	900 (8,4)	372 (9,5)	202 (7,7)	326 (7,9)		
Macrozona	Norte	1688 (15,8)	654 (16,7)	443 (16,8)	591 (14,4)	< 0,001	V = 0,045 (0,034 a 0,060)
	Centro	2207 (20,7)	766 (19,5)	538 (20,4)	903 (21,9)		
	Metropolitana	2533 (23,7)	835 (21,3)	644 (24,4)	1054 (25,6)		
	Sur	4249 (39,8)	1667 (42,5)	1013 (38,4)	1569 (38,1)		
Nivel educacional	Básica	2586 (24,2)	473 (12,1)	544 (20,6)	1569 (38,1)	< 0,001	V = 0,215 (0,204 a 0,228)
	Media incompleta	1753 (16,4)	544 (13,9)	408 (15,5)	801 (19,5)		
	Media completa	3275 (30,7)	1444 (36,8)	845 (32,0)	986 (23,9)		
	Superior	3063 (28,7)	1461 (37,3)	841 (31,9)	761 (18,5)		
Quintil de ingreso del hogar	I	6274 (58,8)	2032 (51,8)	1549 (58,7)	2693 (65,4)	< 0,001	V = 0,099 (0,086 a 0,113)
	II	1716 (16,1)	655 (16,7)	425 (16,1)	636 (15,4)		

Variable	Categoría	Total (n = 10677)	Cero (n = 3922)	Una (n = 2638)	Dos o más (n = 4117)	p	Tamaño de efecto (IC 95%)
	III	1392 (13,0)	599 (15,3)	327 (12,4)	466 (11,3)		
	IV-V	1295 (12,1)	636 (16,2)	337 (12,8)	322 (7,8)		
Situación ocupacional	Trabaja	4493 (42,1)	2171 (55,4)	1211 (45,9)	1111 (27,0)	< 0,001	V = 0,253 (0,235 a 0,271)
	No trabaja	6184 (57,9)	1751 (44,6)	1427 (54,1)	3006 (73,0)		
Previsión de salud	FONASA A-B	7467 (69,9)	2493 (63,6)	1775 (67,3)	3199 (77,7)	< 0,001	V = 0,097 (0,085 a 0,111)
	FONASA C-D	2028 (19,0)	910 (23,2)	544 (20,6)	574 (13,9)		
	ISAPRE u otro	1182 (11,1)	519 (13,2)	319 (12,1)	344 (8,4)		
Material de muros exteriores	Aceptables	10247 (96,0)	3776 (96,3)	2534 (96,1)	3937 (95,6)	0,323	V = 0,015 (0,004 a 0,035)
	Deficientes	430 (4,0)	146 (3,7)	104 (3,9)	180 (4,4)		
Material del piso	Aceptable	10432 (97,7)	3822 (97,5)	2579 (97,8)	4031 (97,9)	0,376	V = 0,014 (0,004 a 0,035)
	Deficiente	245 (2,3)	100 (2,5)	59 (2,2)	86 (2,1)		
Combustible de calefacción	Biomasa	3954 (37,0)	1453 (37,0)	932 (35,3)	1569 (38,1)	0,088	V = 0,019 (0,011 a 0,036)
	Gas o electricidad	5282 (49,5)	1922 (49,0)	1331 (50,5)	2029 (49,3)		
	Sin calefacción	1441 (13,5)	547 (13,9)	375 (14,2)	519 (12,6)		
Contaminación del aire en el barrio	Sí	2898 (27,1)	1000 (25,5)	725 (27,5)	1173 (28,5)	0,010	V = 0,030 (0,014 a 0,049)
Ruidos molestos en el barrio	Sí	3990 (37,4)	1362 (34,7)	984 (37,3)	1644 (39,9)	< 0,001	V = 0,047 (0,028 a 0,066)
Microbasurales o malos olores	Sí	4857 (45,5)	1623 (41,4)	1182 (44,8)	2052 (49,8)	< 0,001	V = 0,074 (0,057 a 0,093)
Inseguridad al caminar de noche	Sí	5044 (47,2)	1667 (42,5)	1249 (47,3)	2128 (51,7)	< 0,001	V = 0,080 (0,062 a 0,098)
Consumo de tabaco	Nunca fumador	6157 (57,7)	2288 (58,3)	1475 (55,9)	2394 (58,1)	< 0,001	V = 0,055 (0,042 a 0,068)
	Exfumador	1825 (17,1)	549 (14,0)	471 (17,9)	805 (19,6)		
	Fumador actual	2695 (25,2)	1085 (27,7)	692 (26,2)	918 (22,3)		
Consumo de alcohol	Sin consumo en 12 meses	4452 (41,7)	1425 (36,3)	1048 (39,7)	1979 (48,1)	< 0,001	V = 0,075 (0,063 a 0,089)

Variable	Categoría	Total (n = 10677)	Cero (n = 3922)	Una (n = 2638)	Dos o más (n = 4117)	p	Tamaño de efecto (IC 95%)
	Consumo sin riesgo	5772 (54,1)	2310 (58,9)	1471 (55,8)	1991 (48,4)		
	Consumo riesgoso	453 (4,2)	187 (4,8)	119 (4,5)	147 (3,6)		
Hipertensión arterial	Sí	3809 (35,7)	0 (0,0)	882 (33,4)	2927 (71,1)		Relación definicional
Diabetes mellitus	Sí	2103 (19,7)	0 (0,0)	305 (11,6)	1798 (43,7)		Relación definicional
Enfermedad respiratoria crónica	Sí	808 (7,6)	0 (0,0)	157 (6,0)	651 (15,8)		Relación definicional
Depresión o ansiedad	Sí	2213 (20,7)	0 (0,0)	538 (20,4)	1675 (40,7)		Relación definicional
Artrosis o artritis	Sí	2418 (22,6)	0 (0,0)	293 (11,1)	2125 (51,6)		Relación definicional
Dolor crónico	Sí	2210 (20,7)	0 (0,0)	364 (13,8)	1846 (44,8)		Relación definicional
Enfermedad cardiovascular	Sí	1166 (10,9)	0 (0,0)	99 (3,8)	1067 (25,9)		Relación definicional

Los valores corresponden a n (%) sin ponderar para las variables categóricas; media (DE) para las continuas evaluadas con ANOVA y mediana [P25; P75] para las evaluadas con Kruskal-Wallis. V: V de Cramér; η^2 : eta cuadrado; ϵ^2 : epsilon cuadrado. Los intervalos de confianza del 95% provienen de 2.000 remuestreos bootstrap percentil. IPAQ: Cuestionario Internacional de Actividad Física. AUDIT: Alcohol Use Disorders Identification Test. FONASA: Fondo Nacional de Salud. Las siete condiciones que componen la variable de agrupación se presentan sin prueba estadística por tratarse de una relación definicional.

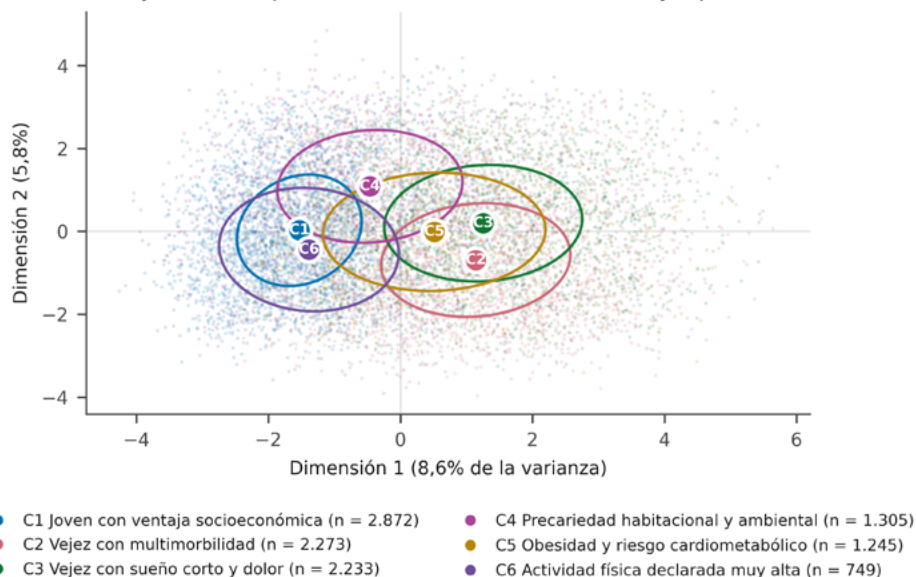
Tabla 3. Diferencias de las variables entre los seis conglomerados obtenidos por K-Prototypes, con tamaño de efecto e intervalo de confianza del 95%. Convenciones y método de cálculo idénticos a los de la Tabla 2.

Variable	Categoría	Total (n = 10677)	C1 (n = 2872)	C2 (n = 2273)	C3 (n = 2233)	C4 (n = 1305)	C5 (n = 1245)	C6 (n = 749)	p	Tamaño de efecto (IC 95%)
Edad, años		52,1 (17,6)	36,3 (11,4)	68,0 (9,8)	65,3 (9,9)	42,7 (14,2)	49,6 (13,4)	45,5 (15,4)	< 0,001	$\eta^2 = 0,555$ (0,544 a 0,566)
Índice de masa corporal, kg/ m ²		27,7 [24,8; 31,2]	25,8 [23,5; 28,3]	27,0 [24,6; 29,6]	28,0 [25,4; 30,8]	28,0 [25,1; 31,2]	37,0 [34,6; 40,0]	27,1 [24,4; 30,4]	< 0,001	$\varepsilon^2 = 0,317$ (0,304 a 0,331)
Horas de sueño por día		7,0 [6,0; 8,0]	7,0 [6,0; 8,0]	8,0 [7,0; 8,0]	5,0 [4,0; 6,0]	7,0 [6,0; 8,0]	7,0 [6,0; 8,0]	7,0 [6,0; 8,0]	< 0,001	$\varepsilon^2 = 0,391$ (0,380 a 0,403)
Personas por dormitorio		1,0 [0,7; 1,3]	1,0 [0,7; 1,3]	0,8 [0,5; 1,0]	1,0 [0,5; 1,0]	2,0 [1,7; 2,0]	1,0 [0,9; 1,3]	1,0 [0,7; 1,3]	< 0,001	$\varepsilon^2 = 0,328$ (0,314 a 0,342)
Actividad física, MET- minuto por semana		1 188,0 [240,0; 2 994,0]	1 651,5 [578,6; 3 365,2]	693,0 [99,0; 1 920,0]	693,0 [66,0; 1 848,0]	1 155,0 [264,0; 2 772,0]	660,0 [0,0; 2 010,0]	9 360,0 [7 758,0; 12 204,0]	< 0,001	$\varepsilon^2 = 0,235$ (0,222 a 0,249)
Sexo	Mujer	7163 (67,1)	1846 (64,3)	1523 (67,0)	1531 (68,6)	985 (75,5)	982 (78,9)	296 (39,5)	< 0,001	V = 0,191 (0,173 a 0,212)
	Hombre	3514 (32,9)	1026 (35,7)	750 (33,0)	702 (31,4)	320 (24,5)	263 (21,1)	453 (60,5)		
Zona de resi- dencia	Urbano	9777 (91,6)	2652 (92,3)	2052 (90,3)	2084 (93,3)	1204 (92,3)	1138 (91,4)	647 (86,4)	< 0,001	V = 0,063 (0,046 a 0,086)
	Rural	900 (8,4)	220 (7,7)	221 (9,7)	149 (6,7)	101 (7,7)	107 (8,6)	102 (13,6)		
Macrozona	Norte	1688 (15,8)	460 (16,0)	311 (13,7)	370 (16,6)	248 (19,0)	181 (14,5)	118 (15,8)	< 0,001	V = 0,101 (0,093 a 0,114)
	Centro	2207 (20,7)	602 (21,0)	433 (19,0)	528 (23,6)	263 (20,2)	245 (19,7)	136 (18,2)		
	Metropoli- tana	2533 (23,7)	619 (21,6)	469 (20,6)	619 (27,7)	460 (35,2)	256 (20,6)	110 (14,7)		
	Sur	4249 (39,8)	1191 (41,5)	1060 (46,6)	716 (32,1)	334 (25,6)	563 (45,2)	385 (51,4)		
Nivel educa- cional	Básica	2586 (24,2)	116 (4,0)	942 (41,4)	871 (39,0)	212 (16,2)	325 (26,1)	120 (16,0)	< 0,001	V = 0,260 (0,251 a 0,270)
	Media in- completa	1753 (16,4)	296 (10,3)	443 (19,5)	482 (21,6)	208 (15,9)	193 (15,5)	131 (17,5)		
	Media com- pleta	3275 (30,7)	942 (32,8)	543 (23,9)	530 (23,7)	532 (40,8)	460 (36,9)	268 (35,8)		
	Superior	3063 (28,7)	1518 (52,9)	345 (15,2)	350 (15,7)	353 (27,0)	267 (21,4)	230 (30,7)		
Quintil de in- greso del hogar	I	6274 (58,8)	1270 (44,2)	1357 (59,7)	1357 (60,8)	1029 (78,9)	875 (70,3)	386 (51,5)	< 0,001	V = 0,149 (0,139 a 0,161)

Variable	Categoría	Total (n = 10677)	C1 (n = 2872)	C2 (n = 2273)	C3 (n = 2233)	C4 (n = 1305)	C5 (n = 1245)	C6 (n = 749)	p	Tamaño de efecto (IC 95%)
	II	1716 (16,1)	511 (17,8)	399 (17,6)	378 (16,9)	145 (11,1)	160 (12,9)	123 (16,4)		
	IV-V	1295 (12,1)	619 (21,6)	197 (8,7)	211 (9,4)	54 (4,1)	91 (7,3)	123 (16,4)		
	III	1392 (13,0)	472 (16,4)	320 (14,1)	287 (12,9)	77 (5,9)	119 (9,6)	117 (15,6)		
Situación ocupacional	Trabaja	4493 (42,1)	1868 (65,0)	427 (18,8)	599 (26,8)	572 (43,8)	520 (41,8)	507 (67,7)	< 0,001	V = 0,380 (0,364 a 0,397)
	No trabaja	6184 (57,9)	1004 (35,0)	1846 (81,2)	1634 (73,2)	733 (56,2)	725 (58,2)	242 (32,3)		
Previsión de salud	FONASA A-B	7467 (69,9)	1622 (56,5)	1806 (79,5)	1720 (77,0)	954 (73,1)	881 (70,8)	484 (64,6)	< 0,001	V = 0,144 (0,132 a 0,158)
	FONASA C-D	2028 (19,0)	771 (26,8)	269 (11,8)	311 (13,9)	236 (18,1)	264 (21,2)	177 (23,6)		
	ISAPRE u otro	1182 (11,1)	479 (16,7)	198 (8,7)	202 (9,0)	115 (8,8)	100 (8,0)	88 (11,7)		
Material de muros exteriores	Aceptables	10247 (96,0)	2776 (96,7)	2192 (96,4)	2139 (95,8)	1222 (93,6)	1196 (96,1)	722 (96,4)	< 0,001	V = 0,047 (0,031 a 0,071)
	Deficientes	430 (4,0)	96 (3,3)	81 (3,6)	94 (4,2)	83 (6,4)	49 (3,9)	27 (3,6)		
Material del piso	Aceptable	10432 (97,7)	2820 (98,2)	2237 (98,4)	2186 (97,9)	1250 (95,8)	1212 (97,3)	727 (97,1)	< 0,001	V = 0,055 (0,038 a 0,081)
	Deficiente	245 (2,3)	52 (1,8)	36 (1,6)	47 (2,1)	55 (4,2)	33 (2,7)	22 (2,9)		
Combustible de calefacción	Biomasa	3954 (37,0)	978 (34,1)	1062 (46,7)	720 (32,2)	329 (25,2)	509 (40,9)	356 (47,5)	< 0,001	V = 0,114 (0,102 a 0,130)
	Gas o electricidad	5282 (49,5)	1541 (53,7)	940 (41,4)	1199 (53,7)	714 (54,7)	586 (47,1)	302 (40,3)		
	Sin calefacción	1441 (13,5)	353 (12,3)	271 (11,9)	314 (14,1)	262 (20,1)	150 (12,0)	91 (12,1)		
Contaminación del aire en el barrio	Sí	2898 (27,1)	780 (27,2)	497 (21,9)	629 (28,2)	472 (36,2)	324 (26,0)	196 (26,2)	< 0,001	V = 0,091 (0,074 a 0,113)
Ruidos molestos en el barrio	Sí	3990 (37,4)	1038 (36,1)	737 (32,4)	903 (40,4)	595 (45,6)	447 (35,9)	270 (36,0)	< 0,001	V = 0,083 (0,067 a 0,103)
Microbasurales o malos olores	Sí	4857 (45,5)	1176 (40,9)	898 (39,5)	1092 (48,9)	783 (60,0)	588 (47,2)	320 (42,7)	< 0,001	V = 0,131 (0,113 a 0,151)
Inseguridad al caminar de noche	Sí	5044 (47,2)	1148 (40,0)	998 (43,9)	1285 (57,5)	806 (61,8)	535 (43,0)	272 (36,3)	< 0,001	V = 0,174 (0,156 a 0,194)
Consumo de tabaco	Nunca fumador	6157 (57,7)	1612 (56,1)	1372 (60,4)	1308 (58,6)	728 (55,8)	714 (57,3)	423 (56,5)	< 0,001	V = 0,085 (0,074 a 0,100)
	Exfumador	1825 (17,1)	389 (13,5)	477 (21,0)	437 (19,6)	196 (15,0)	219 (17,6)	107 (14,3)		
	Fumador actual	2695 (25,2)	871 (30,3)	424 (18,7)	488 (21,9)	381 (29,2)	312 (25,1)	219 (29,2)		

Variable	Categoría	Total (n = 10677)	C1 (n = 2872)	C2 (n = 2273)	C3 (n = 2233)	C4 (n = 1305)	C5 (n = 1245)	C6 (n = 749)	p	Tamaño de efecto (IC 95%)
Consumo de alcohol	Sin consumo en 12 meses	4452 (41,7)	858 (29,9)	1227 (54,0)	1143 (51,2)	478 (36,6)	510 (41,0)	236 (31,5)	< 0,001	V = 0,146 (0,135 a 0,162)
	Consumo sin riesgo	5772 (54,1)	1854 (64,6)	993 (43,7)	1013 (45,4)	765 (58,6)	691 (55,5)	456 (60,9)		
	Consumo riesgoso	453 (4,2)	160 (5,6)	53 (2,3)	77 (3,4)	62 (4,8)	44 (3,5)	57 (7,6)		
Hipertensión arterial	Sí	3809 (35,7)	216 (7,5)	1341 (59,0)	1301 (58,3)	288 (22,1)	505 (40,6)	158 (21,1)	< 0,001	V = 0,455 (0,440 a 0,471)
Diabetes mellitus	Sí	2103 (19,7)	139 (4,8)	676 (29,7)	677 (30,3)	184 (14,1)	348 (28,0)	79 (10,5)	< 0,001	V = 0,278 (0,263 a 0,294)
Enfermedad respiratoria crónica	Sí	808 (7,6)	122 (4,2)	195 (8,6)	254 (11,4)	80 (6,1)	113 (9,1)	44 (5,9)	< 0,001	V = 0,100 (0,084 a 0,119)
Depresión o ansiedad	Sí	2213 (20,7)	482 (16,8)	409 (18,0)	615 (27,5)	268 (20,5)	329 (26,4)	110 (14,7)	< 0,001	V = 0,115 (0,098 a 0,137)
Artrosis o artritis	Sí	2418 (22,6)	150 (5,2)	764 (33,6)	908 (40,7)	190 (14,6)	318 (25,5)	88 (11,7)	< 0,001	V = 0,331 (0,316 a 0,347)
Dolor crónico	Sí	2210 (20,7)	353 (12,3)	506 (22,3)	672 (30,1)	239 (18,3)	311 (25,0)	129 (17,2)	< 0,001	V = 0,159 (0,142 a 0,179)
Enfermedad cardiovascular	Sí	1166 (10,9)	100 (3,5)	383 (16,8)	408 (18,3)	79 (6,1)	144 (11,6)	52 (6,9)	< 0,001	V = 0,197 (0,181 a 0,213)

A. Proyección en el plano factorial del FAMD, con centroides y elipses de una DE



B. Desviación estandarizada de cada variable respecto de la muestra completa

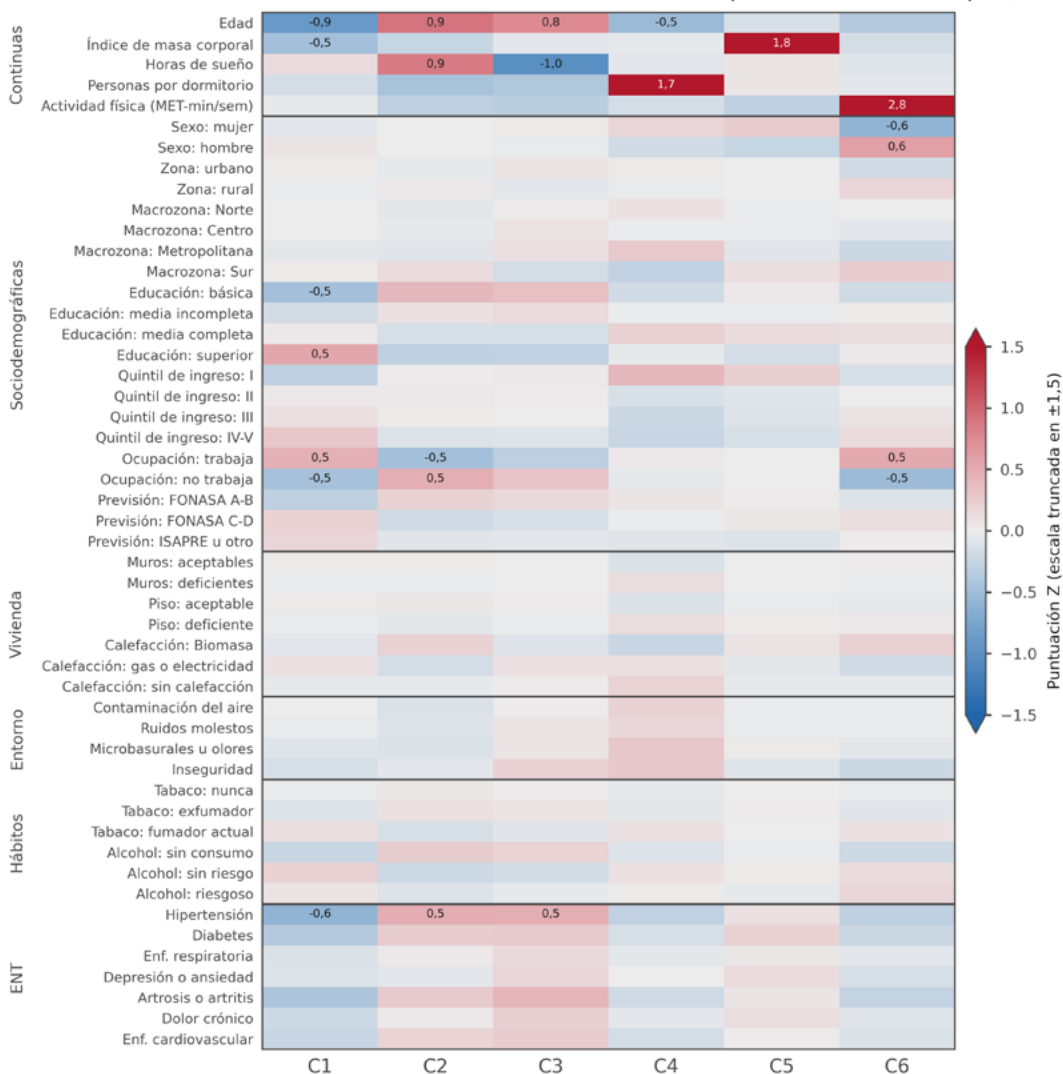


Figura 1. Estructura de los seis conglomerados. A. Proyección de los 10.677 participantes en el plano definido por las dos primeras dimensiones del análisis factorial de datos mixtos, con el centroide de cada conglomerado y su elipse de una desviación estándar. B. Desviación estandarizada de cada variable respecto de la media de la muestra completa, en cada conglomerado. Para las variables continuas corresponde a la diferencia de medias dividida por la desviación estándar muestral; para cada categoría de las variables categóricas, a la diferencia de proporciones dividida por la desviación estándar de la proporción. La escala se truncó en $\pm 1,5$ y las celdas con desviación igual o superior a 0,45 en valor absoluto muestran el valor exacto.

Referencias

- GBD 2019 Acute and Chronic Care Collaborators. Characterising acute and chronic care needs: insights from the Global Burden of Disease Study 2019. *Nat Commun.* 2025;16(1):4235. doi: 10.1038/s41467-025-56910-x
- Santos-López M, Gómez-San Martín P, Margozzini P, Ortuño D. Multimorbidity and tooth loss: data from Chilean National Health Survey 2016-2017. *BMC Oral Health.* 2024;24(1):1417. doi: 10.1186/s12903-024-05184-8
- Nazar G, Díaz-Toro F, Concha-Cisternas Y, Leiva-Ordoñez AM, Troncoso-Pantoja C, Celis-Morales C, et al. Latent class analyses of multimorbidity and all-cause mortality: a prospective study in Chilean adults. *PLoS One.* 2023;18(12):e0295958. doi: 10.1371/journal.pone.0295958
- Mariman JJ, Vergara RC, San Martín C, Zapata V, Arteaga O, Delano PH, et al. Modifiable dementia risk factors in Chilean adults are distinctively associated with social determinants of health. Cross-sectional study. *BMC Public Health.* 2025;25(1):1117. doi: 10.1186/s12889-025-22220-6
- Pagès J. Analyse factorielle de données mixtes. *Rev Stat Appl.* 2004;52(4):93-111.
- Huang Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Min Knowl Discov.* 1998;2(3):283-304. doi: 10.1023/A:1009769707641
- Anderson W, Gould R, Patil N, Mohr N, Dodd K, Boyce D, et al. Unveiling sub-populations in critical care settings: a real-world data approach in COVID-19. *Front Public Health.* 2025;13:1544904. doi: 10.3389/fpubh.2025.1544904
- Deng Y, Cheng S, Huang H, Liu X, Yu Y, Gu M, et al. Machine learning-based phenomapping in patients with heart failure and secondary prevention implantable cardioverter-defibrillator implantation: a proof-of-concept study. *Rev Cardiovasc Med.* 2023;24(2):37. doi: 10.31083/j.rcm2402037
- Allende S. La realidad médico-social chilena. Santiago: Ministerio de Salubridad, Previsión y Asistencia Social; 1939.
- Tedeschi SK, Brown TM, Fee E. Salvador Allende: physician, socialist, populist, and president. *Am J Public Health.* 2003;93(12):2014-5. doi: 10.2105/ajph.93.12.2014
- Ministerio de Salud. Manual de uso de base de datos y libro de códigos. Encuesta Nacional de Calidad de Vida y Salud 2023-2024. Santiago de Chile; 2025. Disponible en: <https://epi.minsal.cl/bases-de-datos>
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol.* 2008;61(4):344-9. doi: 10.1016/j.jclinepi.2007.11.008
- Hennig C. Cluster-wise assessment of cluster stability. *Comput Stat Data Anal.* 2007;52(1):258-71. doi: 10.1016/j.csda.2006.11.025
- Kim J, Park GR, Jang H, Son H. Poor housing conditions in adolescence and adult health outcomes: an outcome-wide longitudinal approach. *J Epidemiol Community Health.* 2025;79(5):317-23. doi: 10.1136/jech-2024-222378
- Rana K, Kent JL, Page A. Housing inequalities and health outcomes among migrant and refugee populations in high-income countries: a mixed-methods systematic review. *BMC Public Health.* 2025;25(1):1098. doi: 10.1186/s12889-025-22186-5

Figura suplementaria S1. Criterios para la selección del número de conglomerados. A. Función de costo de K-Prototypes. B. Coeficiente de silueta. C. Índice de Calinski-Harabasz. D. Índice de Rand ajustado medio entre la solución completa y doce soluciones obtenidas en submuestras aleatorias del 80%, con barras de desviación estándar. La línea vertical discontinua marca la solución retenida en A y D, y el óptimo de cada índice en B y C.

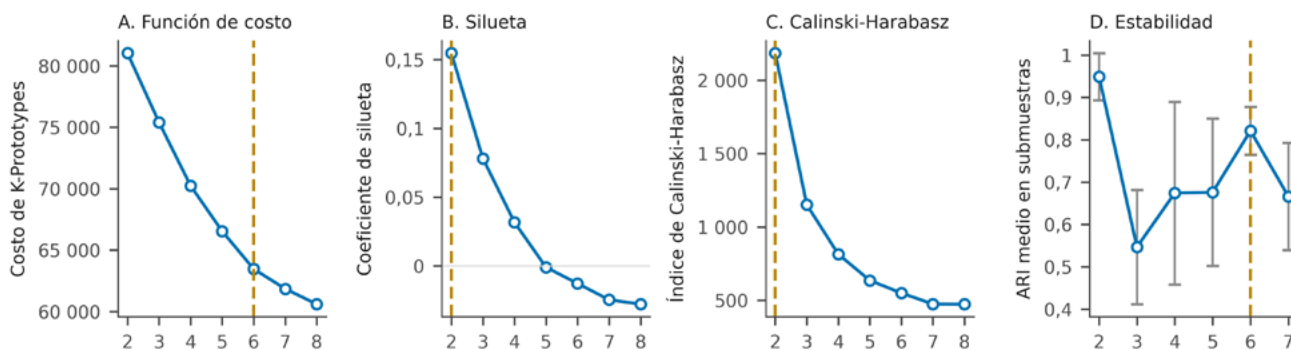


Tabla suplementaria S1. Glosario de términos metodológicos y estadísticos en lenguaje no técnico: Perfiles poblacionales de enfermedades crónicas no transmisibles, hábitos, características sociodemográficas, vivienda y entorno residencial en Chile (ENCAVI 2023-2024). Las definiciones son descriptivas y buscan facilitar la lectura a quienes no provienen de la estadística.

Término	Qué es, en lenguaje sencillo	Cómo se lee en este estudio
Enfermedades crónicas no transmisibles (ENT)	Problemas de salud de larga duración que no se contagian, como hipertensión, diabetes o artrosis.	Se consideraron siete, todas por diagnóstico médico autorreportado.
Multimorbilidad	Presencia simultánea de dos o más enfermedades crónicas en una misma persona.	Se clasificó en tres categorías: ninguna, una, dos o más condiciones.
Perfil o conglomerado	Grupo de personas que se parecen entre sí en el conjunto de sus características y se diferencian de las de otros grupos.	Se identificaron seis (C1 a C6). Son descripciones de la muestra, no categorías diagnósticas.
Análisis factorial de datos mixtos (FAMD)	Técnica que resume en pocos ejes la información de muchas variables, numéricas y cualitativas a la vez.	Permite representar en un plano a las 10.677 personas analizadas (Figura 1A).
K-Prototypes	Algoritmo que forma grupos usando promedios para las variables numéricas y la categoría más frecuente para las cualitativas.	Produjo la partición en conglomerados para valores de K entre 2 y 8.
Función de costo	Medida de cuán dispersas están las personas respecto del centro de su grupo; menor valor indica grupos más compactos.	Se buscó el punto donde la mejora deja de ser sustancial al agregar grupos (entre K = 5 y K = 6).
Coeficiente de silueta	Compara cuán cerca está cada persona de su propio grupo frente al grupo vecino más cercano; va de -1 a 1.	Valores cercanos a cero indican grupos que se superponen; en este estudio la solución retenida mostró ese patrón.
Índices de Calinski-Harabasz y Davies-Bouldin	Comparan la dispersión dentro de los grupos con la distancia entre ellos. En el primero, mayor valor es mejor; en el segundo, menor valor es mejor.	Favorecieron dos grupos, pero esa solución reproducía sobre todo un gradiente de edad y carga de enfermedad.
Índice de Rand ajustado	Mide cuánto coinciden dos agrupaciones de las mismas personas; 1 es coincidencia perfecta y 0 la esperable por azar.	Se calculó entre la solución completa y doce soluciones en submuestras del 80%, como medida de estabilidad.

Término	Qué es, en lenguaje sencillo	Cómo se lee en este estudio
Tamaño de efecto	Indica cuán grande es una diferencia o asociación, más allá de si es estadísticamente detectable.	Se informó para todas las comparaciones, con intervalo de confianza.
V de Cramér	Tamaño de efecto para dos variables cualitativas; va de 0 (sin asociación) a 1 (asociación total).	La interpretación depende del número de categorías; como referencia, en tablas con dos categorías por lado 0,1 es pequeño y 0,3 moderado.
Eta cuadrado (η^2) y epsilon cuadrado (ϵ^2)	Proporción de la variabilidad de una variable numérica que se asocia a la pertenencia a un grupo; van de 0 a 1. ϵ^2 se basa en los rangos de los valores, no en los valores originales.	η^2 se usó con ANOVA y ϵ^2 con Kruskal-Wallis, según la distribución de la variable.
Intervalo de confianza bootstrap	Rango plausible para un valor, obtenido repitiendo el cálculo en muchos remuestreos con reemplazo de los mismos datos.	Se usaron 2.000 remuestreos y el intervalo del 95%.
Casos completos	Análisis restringido a las personas con información en todas las variables.	Quedaron 10.677 de 16.590 registros (64,4%).
Personas por dormitorio	Razón entre el número de habitantes de la vivienda y sus dormitorios de uso exclusivo; indicador de hacinamiento.	Valores mayores indican mayor hacinamiento.
Quintil de ingreso del hogar	División de la población en cinco grupos de igual tamaño según ingreso; el quintil I es el de menor ingreso.	C4 concentró personas del quintil I.
MET-minuto por semana	Unidad que combina la intensidad y el tiempo de la actividad física semanal, según el Cuestionario Internacional de Actividad Física.	Valores muy altos pueden reflejar sobrerreporte (ver Discusión, conglomerado C6).