

UNIVERSIDAD PANAMERICANA

**FACULTAD DE EMPRESARIALES
DOCTORADO EN CIENCIAS EMPRESARIALES
INCORPORADO A LA SECRETARÍA DE EDUCACIÓN
PÚBLICA RVOE 20181652**

**“Impacto de las enfermedades no transmisibles en la
productividad laboral en México, 2022.”**

**QUE PARA OBTENER EL GRADO ACADÉMICO DE:
DOCTOR EN CIENCIAS EMPRESARIALES**

PRESENTA:

Sergio Arturo Domínguez Miranda

Director:

Dr Román Rodríguez Aguilar

Codirector:

Dra Marisol Velázquez Salazar

Revisor/Lector:

Dr Salvador Rivas Aceves

Ciudad de México, 2026

DEDICATORIA

Este trabajo está dedicado, en primer lugar, a mi padre, Sergio Domínguez. Aunque ya no puedo tener contigo las conversaciones que tanto extraño, sé que sigues acompañándome desde el cielo junto a mi hijo Iker. Sé que están presentes, que forman parte de cada paso y que este logro también es suyo. Gracias, papá, por lo que fui, por lo que soy y por todo lo que sembraste en mí. Siempre serás mi orgullo... y yo sé que hoy también lo soy para ti. Esto también es para ti, hijito mío.

A mi mamá Lupita, a mi hermano el Mtro. Marco Antonio Domínguez y mi hermana la Dra. Mayra Domínguez, mis compañeros de vida. Gracias por caminar conmigo en cada etapa, por sus palabras constantes, por su fuerza silenciosa y por su amor incondicional. Todo su apoyo, sus ánimos y su presencia están profundamente grabados en mi corazón. Nada de esto habría sido posible sin ustedes.

Y, finalmente, a lo más importante de mi vida: mis hijas, Miranda y Valérie.

Ustedes son mi motor, mi energía y la razón más profunda por la que estoy aquí. Cada palabra suya, sin importar la edad en que fue pronunciada, resonó en lo más profundo de mi ser y me dio fuerza en los momentos más difíciles. Ustedes son lo mejor de mí, mi mayor orgullo y mi mayor inspiración.

Esto es por ustedes y para ustedes.

“Lo que hacemos por nosotros mismos muere con nosotros; lo que hacemos por los demás y por el mundo permanece y es inmortal.”

— Albert Einstein

AGRADECIMIENTOS

El desarrollo y culminación de esta tesis doctoral no habrían sido posibles sin el apoyo, la guía y la confianza de diversas personas e instituciones que acompañaron de manera decisiva mi formación académica y profesional a lo largo de este exigente, pero enriquecedor proceso.

En primer lugar, deseo expresar mi más profundo agradecimiento a Siemens Healthineers, institución que confió en mí y en mi potencial desde el inicio de este camino. A lo largo de estos años, encontré en la empresa no solo un espacio de crecimiento profesional, sino también un entorno que fomentó de manera genuina mi desarrollo académico y humano. Agradezco especialmente el respaldo institucional brindado para continuar con mi formación doctoral, el apoyo recibido para disponer de los espacios necesarios para el desarrollo de mis investigaciones y la oportunidad de participar en congresos internacionales y presentar los resultados de este trabajo más allá de las fronteras de mi país. Mi gratitud a Antonio Ribeiro, Alejandro Paolini y Francisco Rosas, cuyo apoyo, confianza y liderazgo fueron fundamentales para hacer posible este proyecto.

Agradezco de manera muy especial a la Università degli Studi di Firenze por integrarme a su comunidad académica como parte de mi programa doctoral. La experiencia formativa y académica vivida en esta institución amplió de manera decisiva mi perspectiva como investigador, fortaleciendo mi pensamiento crítico y mi comprensión de la investigación científica en un contexto internacional.

Agradezco a la Cámara Nacional de la Industria Farmacéutica y de Dispositivos Médicos por el acompañamiento durante estos cinco años, la inclusión en sus proyectos y la oportunidad de compartir avances de investigación y vincularme con expertos del sector; en especial, a Jorge Daniel por su apoyo constante y apertura académica.

Mi reconocimiento y agradecimiento a la Universidad Panamericana por su orientación, esfuerzo y constante dedicación hacia los estudiantes del Doctorado en Ciencias Empresariales. Gracias por brindar un entorno académico riguroso, humano y exigente, así como las facilidades necesarias para el desarrollo de esta investigación. Esta institución fue un pilar fundamental en mi formación doctoral.

A todos los profesores del programa doctoral, mi sincera gratitud. Fueron muchos quienes influyeron en la construcción de mi visión como investigador, mostrando, con su ejemplo, rigor y compromiso, el camino de la investigación académica. Agradezco profundamente a cada uno de los docentes que, directa o indirectamente, dejaron huella en este proceso formativo.

A mis familiares, amigas y amigos queridos, colegas, médicos y doctores, gracias por su apoyo constante, por sus palabras de aliento y, sobre todo, por su comprensión ante las ausencias y sacrificios que implicó este proyecto. Durante cinco años, el enfoque y la disciplina necesarios para concluir este doctorado demandaron tiempo y energía, y su respaldo fue siempre una fuente invaluable de motivación. También agradezco, de manera muy especial, a quienes, con su presencia o con muestras que trascienden las palabras, hicieron que este camino tuviera sentido más allá de lo académico.

A mis compañeros del doctorado, con quienes compartí aprendizajes, retos y crecimiento personal. Fuimos varios los que iniciamos este camino y tres quienes lo concluimos, formando una generación excepcional. Gracias especialmente a David y Jorge; fue un verdadero privilegio compartir esta etapa con ustedes. Sin duda, una generación memorable.

Un agradecimiento muy especial a mis asesores de tesis, la Dra. Marisol Velázquez y el Dr. Salvador Rivas, por su guía, orientación, paciencia y palabras de aliento a lo largo de todo el proceso. Su apoyo académico y humano fue clave para avanzar en los momentos más complejos de esta investigación. Son, sin duda, extraordinarios profesionales y grandes seres humanos.

Finalmente, deseo expresar mi más profundo y sincero agradecimiento a mi asesor principal, el Dr. Román Rodríguez Aguilar. En él encontré al mentor idóneo para crecer como investigador. Su amplio conocimiento transformó de manera definitiva mi visión académica y profesional, abriendo horizontes que jamás imaginé. Gracias por la oportunidad de colaborar en múltiples artículos, por el rigor científico, la exigencia constante y el acompañamiento decidido a lo largo de este camino. Eres un gran profesionista, un gran líder y, sobre todo, un gran ser humano. Gracias por ayudarme a llegar hasta donde hoy me encuentro.

“Si he logrado ver más lejos, ha sido porque he estado sobre los hombros de gigantes.”
— Isaac Newton

ÍNDICE

Resumen	8
Palabras Clave.....	10
Abstract.....	11
Keywords.....	13
CAPÍTULO I. Introducción.....	14
1.1 Planteamiento del problema	15
1.2 Justificación del Tema	27
1.3 Preguntas generales y específicas de la Investigación.....	33
1.4 Objetivo de Investigación	33
1.5 Hipótesis.....	33
1.6 Descripción de Capítulos	34
CAPÍTULO II. Revisión de la literatura.....	36
2.1 Dimensiones determinantes de las ENT.....	36
<i>Hipótesis 1: Las ENT tienen un efecto negativo en la productividad laboral en México.....</i>	<i>39</i>
2.2 Los hábitos de salud, prevención y promoción de la salud	39
<i>Hipótesis 2: Existe una asociación significativa entre las dimensiones económicas, hábitos de salud y laborales con las ENT.....</i>	<i>41</i>
2.3 Las ENT, hábitos de salud y la productividad laboral	41
<i>Hipótesis 3: La adopción de hábitos de salud positivos de la PEA tiene un efecto positivo en la productividad laboral en México</i>	<i>46</i>
2.4 Marco Conceptual.....	47
CAPÍTULO III. Diseño de la investigación	50
3.1 Variables y Diseño Metodológico.....	51
3.1.1 Operacionalización de Variables.....	51
3.1.2 Diseño metodológico	63
3.2 Muestra y recolección de datos	68
3.2.2 Muestra	68
3.3 Pruebas paramétricas y no paramétricas	73
CAPÍTULO IV. Resultados	88
4.1 Factores relacionados con las ENT en países en desarrollo durante la edad laboral	88
4.2 Hábitos de salud en la PEA y sus efectos en las ENT en México	110
4.3 Impacto de los hábitos de salud en la productividad laboral en México	121
CAPÍTULO V. Conclusiones y Recomendaciones.....	137
CAPÍTULO VI. Contribución Académica, Empresarial y Nacional	143

6.1 Contribución Académica.....	143
6.2 Contribución Empresarial	143
6.3 Contribución Nacional.....	144
CAPÍTULO VII. Limitantes de la Investigación y Futuras Direcciones de la Investigación.	146
Referencias	149
ANEXOS	163

ÍNDICE DE TABLAS Y FIGURAS

Artefacto	Nombre	Página
Tabla 1	Principales Enfermedades no transmisibles – generalidades y estadísticas	18
Tabla 2	Iniciativas basadas en evidencia sobre salud ocupacional y productividad laboral	30
Tabla 3	Matriz de congruencia	34
Tabla 4	Bases de datos globales	53
Tabla 5	Variables seleccionadas en Factores H y L de las ENT	59
Tabla 6	Variables seleccionadas de la ENSANUT	59
Tabla 7	Variables seleccionadas de la ENIGH	61
Tabla 8	Resumen del diseño metodológico	67
Tabla 9	Varianza explicada acumulada por las dimensiones seleccionadas en el análisis de componentes principales (PCA)	90
Tabla 10	Predicción del modelo por MSE en enfermedad cardiovascular	105
Tabla 11	Regresiones penalizadas – enfermedades cardiovasculares	105
Tabla 12	Predicción del modelo por MSE en enfermedades crónico-respiratorias	106
Tabla 13	Regresiones penalizadas – enfermedades crónico-respiratorias	106
Tabla 14	Predicción del modelo por MSE en enfermedad neoplásica	107
Tabla 15	Regresiones penalizadas – enfermedad neoplásica	108
Tabla 16	Predicción del modelo por MSE en Diabetes <i>Mellitus</i>	109
Tabla 17	Regresiones penalizadas – Diabetes <i>Mellitus</i>	109
Tabla 18	Estadísticas descriptivas de las variables seleccionada	110
Tabla 19	Distribución proporcional de las variables categóricas seleccionadas - Ensanut 2022	111
Tabla 20	Comparación de desempeño de modelos – ENSANUT	113
Tabla 21	Ajuste del modelo SEM	116
Tabla 22	Resultados ecuaciones estructurales ENSANUT	120
Tabla 23	Resultados de la prueba de <i>Welch</i> para comparación de medias	124
Tabla 24	Descripción de covariables para el análisis DML	127
Tabla 25	Resultados del modelo <i>Double Machine Learning</i> según algoritmo utilizado	130
Tabla 26	Penalización estimada de ENT sobre la productividad ajustada (modelos DML)	134
Tabla 27	Diagrama de Inferencia 1. Factores estructurales asociados a ENT – Dimensión económica (PCA)	165
Tabla 28	Diagrama de Inferencia 2. Patrones de consumo alimentario asociados a ENT (PCA)	166
Tabla 29	Diagrama de Inferencia 3. Factores de riesgo H y L asociados a ENT (PCA por enfermedad)	166
Tabla 30	Diagrama de Inferencia 4. Efectos de factores de riesgo sobre mortalidad por ENT	167
Tabla 31	Diagrama de Inferencia 5. Hábitos de salud y ENT en la PEA de México (SEM)	167
Tabla 32	Diagrama de Inferencia 6. Impacto causal de las ENT en la productividad laboral (DML)	168
Figura 1	Principales causas de muerte en México. Casos por cada 100,000 habitantes. 2019.	20
Figura 2	Incidencia y prevalencia de las ENT en México	21
Figura 3	Mortalidad en México de enfermedades cardiovasculares – coordenadas paralelas, 2019	22
Figura 4	Mortalidad en enfermedades crónico-respiratorias en México – coordenadas paralelas, 2019	23
Figura 5	Mortalidad en diabetes en México – coordenadas paralelas, 2019	23
Figura 6	Mortalidad en enfermedades neoplásicas en México – coordenadas paralelas, 2019	24
Figura 7	Un marco para la promoción de la salud	40
Figura 8	Esquema conceptual para el análisis del impacto de las ENT en la productividad laboral	48
Figura 9	Esquema del diseño de la investigación	63
Figura 10	Enfoque de análisis híbrido utilizando aprendizaje automático (ML) y modelos de ecuaciones estructurales (SEM)	65

Figura 11	Proceso Analítico para Evaluar el Impacto de las ENT mediante Técnicas Estadísticas y de <i>Machine Learning</i>	66
Figura 12	Distribución de países y carga de variables económicas en las dos primeras dimensiones	90
Figura 13	Índice de factores económicos	91
Figura 14	Distribución de países y carga de variables alimentarias en las dos primeras dimensiones	92
Figura 15	Índice de consumo alimentario	94
Figura 16	Índice de consumo segmentado	95
Figura 17	Distribución de países y carga de variables entre los factores H en las dos primeras dimensiones	96
Figura 18	Distribución de países y carga de variables entre los factores L en las dos primeras dimensiones	97
Figura 19	Mapa de calor de los Índices de factores de riesgo por enfermedad y país	99
Figura 20	Índices vs mortalidad por ENT	100
Figura 21	Agrupación de índices K-medias	102
Figura 22	Características seleccionadas mediante aprendizaje automático: (a) <i>Random Forest</i> ; (b) Regresión <i>Lasso</i>	114
Figura 23	Resultados del modelo SEM estimado	118
Figura 24	Comparación de productividad entre empleados con y sin ENT	123
Figura 25	Distribución de productividad por grupo	125
Figura 26	Costo anual estimado de pérdida de productividad con intervalos de confianza 95%	135

“Impacto de las enfermedades no transmisibles en la productividad laboral en México, 2022.”

Resumen

Las enfermedades no transmisibles (ENT) representan un desafío creciente para los países en desarrollo, al constituir una de las principales causas de morbilidad, mortalidad y pérdidas económicas en la población económicamente activa (PEA). En este contexto, esta tesis doctoral tuvo como objetivo analizar la relación entre los hábitos de salud y la prevalencia de ENT, así como estimar su impacto sobre la productividad laboral en México. Los resultados centrales de la investigación muestran que la presencia de ENT se asocia con una reducción significativa de la productividad laboral, lo que se traduce en una pérdida económica anual aproximada de 496 mil millones de pesos, evidenciando la magnitud del problema desde una perspectiva tanto sanitaria como económica. Las hipótesis planteadas buscaron demostrar que existe una asociación significativa entre las dimensiones económica, de salud y laborales con las ENT, que estas enfermedades tienen un efecto negativo sobre la productividad laboral en México y que la adopción de hábitos saludables por parte de la PEA contribuye positivamente al desempeño productivo.

El abordaje metodológico se estructuró en tres fases complementarias. En la primera fase se realizó un análisis comparativo internacional utilizando una muestra de 14 países en desarrollo. Se construyeron índices mediante análisis de componentes principales (PCA) a partir de variables económicas, alimentarias y de riesgo, lo que permitió sintetizar la información en dimensiones interpretables. Posteriormente, se aplicaron técnicas de análisis de conglomerados (*k-means*) para agrupar a los países según sus perfiles de salud y desarrollo, y se emplearon modelos de regresión penalizada (*Lasso*, *Ridge* y redes elásticas) para identificar las variables con mayor asociación con la mortalidad por ENT según cada grupo de enfermedades (cardiovasculares, respiratorias crónicas, neoplásicas y diabetes *mellitus*).

La segunda fase se enfocó en el análisis nacional a partir de la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2022, con el propósito de evaluar la relación entre los hábitos de salud negativos y la presencia de ENT en la PEA mexicana. Se emplearon modelos de ecuaciones estructurales (SEM) que integraron múltiples dimensiones latentes (perfil individual,

biomarcadores, bioquímica, consumo, estrés y actividad física) con el fin de estimar los efectos causales directos e indirectos sobre la prevalencia de ENT. Esta fase permitió modelar las interacciones complejas entre prácticas de salud, condiciones biológicas y entorno sociodemográfico.

En la tercera fase se utilizó la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022 para estimar el impacto de las ENT sobre la productividad laboral, definida como el cociente entre los ingresos laborales y las horas trabajadas. Para ello, se implementó un modelo de *Double Machine Learning* (DML), que combina técnicas de regresión y algoritmos de aprendizaje automático para controlar variables confusoras y estimar efectos causales de manera robusta. Esta metodología permitió diferenciar el efecto específico de las ENT en la productividad considerando subgrupos de personas con y sin condiciones crónicas asociadas.

En referencia a la primera etapa de análisis, México mostró un comportamiento intermedio dentro de la muestra de países analizados, destacando negativamente en el caso de la diabetes *mellitus*. El análisis de clúster permitió agrupar a los países en patrones similares de riesgo y desarrollo, mientras que los modelos de regresión penalizada confirmaron la utilidad de las redes elásticas para reducir la multicolinealidad y seleccionar las variables más influyentes en cada enfermedad. Los hallazgos obtenidos revelaron que factores como el bajo consumo de *omega* 3, la elevada ingesta de azúcares y grasas, el sedentarismo, la hipertensión y el exceso de peso corporal están asociados de manera significativa con una mayor mortalidad por ENT.

El modelo SEM confirmó la existencia de una relación causal significativa entre los hábitos de salud negativos y las ENT. La variable latente de hábitos negativos, integrada por seis dimensiones, mostró ser un predictor sólido de la aparición de enfermedades crónicas no transmisibles. Los factores con mayor peso dentro del modelo fueron el perfil individual (edad y escolaridad), los biomarcadores (presión arterial sistólica, IMC, circunferencia de cintura) y el consumo de alcohol y tabaco. Además, se identificó una relación entre ENT y hábitos: las ENT no solo son consecuencia de conductas poco saludables, sino que también contribuyen al deterioro de los hábitos posteriores al diagnóstico.

Los resultados del modelo DML evidenciaron que la presencia de ENT tiene un efecto negativo y significativo sobre la productividad laboral, con una disminución promedio de -11.44% (IC95%: [-14.593, -8.398]) en la productividad laboral ($p = 0.000$). Al aplicar un factor de ajuste

para condiciones crónicas, se observó que las personas con ENT con enfermedades crónicas asociadas presentaron una pérdida de productividad del 18.75% respecto a las que tienen ENT, pero sin condiciones crónicas, lo que podría explicarse por la falta de tratamiento o control adecuado.

Los resultados de esta investigación permiten concluir que las ENT no solo representan un problema de salud pública, sino también una barrera significativa para el desarrollo económico sostenible. La evidencia empírica respalda la necesidad de implementar políticas de salud ocupacional y estrategias de prevención focalizadas que promuevan estilos de vida saludables en la PEA. Asimismo, se resalta la utilidad de modelos híbridos, que integran estadística clásica, aprendizaje automático y modelado estructural como herramientas eficaces para la generación de evidencia técnica para la toma de decisiones. Finalmente, se recomienda avanzar hacia sistemas de monitoreo dinámico que integren dispositivos digitales y algoritmos predictivos, con el fin de prevenir el desarrollo de ENT y mitigar sus efectos sobre la productividad laboral.

Este estudio observacional presenta limitaciones relevantes. Una de ellas es la dificultad de integrar bases de datos con enfoques distintos: mientras la ENIGH mide el gasto en salud, la ENSANUT se centra en condiciones clínicas, lo que genera retos al construir variables comparables. Asimismo, la heterogeneidad en la recolección de datos y la ausencia de seguimiento longitudinal restringen la posibilidad de establecer inferencias causales. Futuras investigaciones podrían beneficiarse de la incorporación de tecnologías *wearables* y modelos de aprendizaje profundo, así como de estudios longitudinales que permitan una comprensión más precisa del impacto progresivo de las ENT en la productividad laboral.

Palabras Clave

Enfermedades no transmisibles, productividad laboral, población económicamente activa, promoción de la salud, estilos de vida saludables, *machine learning*, ecuaciones estructurales, *Double Machine Learning*, análisis de componentes principales, regresiones penalizadas.

"Impact of non-communicable diseases on workers' productivity in Mexico, 2022"

Abstract

Non-communicable diseases (NCDs) represent a growing challenge for developing countries, as they constitute one of the main causes of morbidity, mortality, and economic losses within the economically active population (EAP). In this context, this doctoral thesis aimed to analyze the relationship between health-related habits and the prevalence of NCDs, as well as to estimate their impact on labor productivity in Mexico. The central findings of the study indicate that the presence of NCDs is associated with a significant reduction in labor productivity, resulting in an estimated annual economic loss of approximately 496 billion Mexican pesos, highlighting the magnitude of the problem from both a public health and an economic perspective. The core hypotheses sought to demonstrate that there is a significant association between economic, health, and labor dimensions and NCDs; that NCDs have a negative effect on labor productivity in Mexico; and that the adoption of healthy habits by the EAP contributes positively to productive performance.

The methodological approach was structured in three complementary phases. In the first phase, an international comparative analysis was carried out using a sample of 14 developing countries. Indices were constructed through principal component analysis (PCA) based on economic, dietary, and risk variables, which allowed the information to be synthesized into interpretable dimensions. Subsequently, cluster analysis techniques (k-means) were applied to group countries according to their health and development profiles, and penalized regression models (Lasso, Ridge, and elastic nets) were used to identify the variables most strongly associated with NCD mortality across different groups of diseases (cardiovascular, chronic respiratory, neoplastic, and diabetes mellitus).

The second phase focused on a national-level analysis using the 2022 National Health and Nutrition Survey (ENSANUT) to assess the relationship between negative health habits and the presence of NCDs in Mexico's EAP. Structural equation modeling (SEM) was employed to integrate multiple latent dimensions (individual profile, biomarkers, biochemistry, consumption, stress, and physical activity) to estimate both direct and indirect causal effects on NCD prevalence.

This phase enabled the modeling of complex interactions among health behaviors, biological conditions, and sociodemographic contexts.

In the third phase, the 2022 National Household Income and Expenditure Survey (ENIGH) was used to estimate the impact of NCDs on labor productivity, defined as the ratio between labor income and hours worked. For this purpose, a Double Machine Learning (DML) model was implemented, combining regression techniques and machine learning algorithms to control for confounding variables and estimate causal effects robustly. This methodology allowed for the differentiation of the specific effect of NCDs on productivity, while also considering subgroups of individuals with and without associated chronic conditions.

Concerning the first stage of analysis, Mexico displayed an intermediate profile within the sample of countries but performed poorly in the case of diabetes mellitus. The cluster analysis grouped countries into similar patterns of risk and development, while the penalized regression models confirmed the usefulness of elastic nets in reducing multicollinearity and selecting the most influential variables for each disease. The findings revealed that factors such as low omega-3 intake, high consumption of sugars and fats, sedentary lifestyles, hypertension, and excess body weight are significantly associated with higher mortality from NCDs.

The SEM model confirmed the existence of a significant causal relationship between negative health habits and NCDs. The latent variable of negative habits, composed of six dimensions, proved to be a strong predictor of the onset of chronic non-communicable diseases. The most influential factors in the model were the individual profile (age and education), biomarkers (systolic blood pressure, BMI, waist circumference), and alcohol and tobacco consumption. Moreover, a bidirectional relationship between NCDs and habits was identified: NCDs are not only a consequence of unhealthy behaviors but also contribute to the deterioration of habits following diagnosis.

The DML model results showed that the presence of NCDs has a negative and statistically significant effect on labor productivity, with an average decrease of -11.44% (95% CI: $[-14.593, -8.398]$) in labor productivity ($p = 9.4 \times 10^{-8}$). When applying an adjustment factor for chronic conditions, it was observed that individuals with NCDs and associated chronic diseases experienced an 18.75% productivity loss compared to those with NCDs but without comorbidities, which may be explained by a lack of adequate treatment or disease control.

The results of this research allow us to conclude that NCDs are not only a public health problem but also a significant barrier to sustainable economic development. Empirical evidence supports the need to implement occupational health policies and targeted prevention strategies that promote healthy lifestyles among the EAP. Furthermore, the usefulness of hybrid models integrating classical statistics, machine learning, and structural modeling is highlighted as an effective tool for generating technical evidence to inform decision-making. Finally, the study recommends moving towards dynamic monitoring systems that integrate digital devices and predictive algorithms to prevent the development of NCDs and mitigate their effects on labor productivity.

This observational study presents relevant limitations. One of the challenges is integrating databases with different approaches: while ENIGH measures health expenditures, ENSANUT focuses on clinical conditions, which creates difficulties in constructing comparable variables. Likewise, the heterogeneity in data collection and the absence of longitudinal follow-up restrict the possibility of establishing causal inferences. Future research could benefit from the incorporation of wearable technologies and deep learning models, as well as from longitudinal studies that allow for a more accurate understanding of the progressive impact of NCDs on labor productivity.

Keywords

Non-communicable diseases, labor productivity, economically active population, health promotion, healthy lifestyles, machine learning, structural equation modeling, Double Machine Learning, principal component analysis, penalized regressions.

CAPÍTULO I. Introducción

Las ENT son las principales causas de muerte y el principal problema en salud pública a nivel mundial (Córdova-Villalobos *et al.*, 2008). El incremento de las ENT está fuertemente asociado con un cuadro agudo de dislipidemia que es parte de los principales factores de riesgo junto con fumar, actividades sedentarias, la incorrecta alimentación y factores genéticos que han ocasionado y se potencian enfermedades como el síndrome metabólico (Diabetes, hipertensión y obesidad), enfermedades oncológicas, cardiológicas y neurológicas (Ezzati & Riboli, 2013). La mejora de las condiciones sociales, la calidad de la dieta, el peso corporal, el abandono del hábito de fumar y el aumento de la actividad física (AF) pueden reducir en gran medida la carga de este problema de salud (Christian *et al.*, 2025).

Las ENT tienen importantes implicaciones económicas tanto a nivel global como nacional debido a sus elevados costos directos e indirectos. A escala mundial, las ENT provocan pérdidas económicas significativas derivadas de gastos médicos elevados y reducciones sustanciales en la productividad laboral (OPS, 2023). Según la Organización Mundial de la Salud (2022), estas enfermedades podrían costarle a la economía global aproximadamente 47 billones de dólares en los próximos 20 años si no se toman medidas adecuadas de prevención y control. En el contexto mexicano, el impacto económico es particularmente notable; de acuerdo con Arredondo *et al.* (2018), las ENT generan altos costos para el sistema de salud pública y afectan negativamente el crecimiento económico al disminuir la productividad de la PEA, lo cual se refleja en incrementos en los días de ausentismo laboral y una reducción general de la capacidad productiva de las empresas mexicanas.

Tomando en cuenta estos elementos, las empresas han iniciado la implementación de políticas de prevención y promoción de la salud para mantener sana a la PEA. Haldar y Mallik (2010) afirman que una mejor salud aumenta la productividad de la fuerza laboral al reducir la incapacidad, la debilidad y el número de días perdidos debido a bajas por enfermedad. Schultz (1997) enfatiza que un buen estado de salud de la población ayuda a forjar la producción con cualquier combinación de habilidades, capital físico y conocimiento tecnológico. Sin embargo, es difícil analizar qué iniciativas pudieran ser factibles y eficientes en la prevención de ENT, ya que existe una gran diversidad de estrategias que las empresas están implementando; muchas de ellas requieren altas inversiones y para otras el tiempo de implementación puede ser prolongado.

El presente documento plantea una investigación cuyo principal objetivo es evaluar el impacto de las enfermedades no transmisibles (ENT) en la productividad laboral, a partir del análisis conjunto de patrones de consumo alimentario, estilos de vida y variables económicas, utilizando información proveniente de diversas bases de datos nacionales e internacionales. El estudio se desarrolla bajo un enfoque cuantitativo, con énfasis en la construcción y aplicación de una metodología para el análisis integrado de bases de datos con indicadores de salud, económicos y laborales, teniendo como población objetivo a la población económicamente activa (PEA).

La investigación se estructura en dos niveles analíticos complementarios con alcances temporales diferenciados y explícitamente justificados. En una primera etapa se realiza un análisis internacional comparativo del comportamiento de indicadores asociados a las ENT en un conjunto de países en desarrollo, utilizando información correspondiente al periodo previo a la pandemia por COVID-19. Este análisis tiene como finalidad establecer un marco de referencia estructural pre-pandemia, que permita identificar patrones generales, relaciones entre dimensiones económicas, alimentarias y de salud, así como regularidades comunes en la prevalencia de las ENT, evitando sesgos derivados de la disrupción sanitaria y económica observada a partir de 2020.

En una segunda etapa se analiza el caso particular de México mediante un corte transversal correspondiente al año 2022, empleando información de la Encuesta Nacional de Salud y Nutrición y de la Encuesta Nacional de Ingresos y Gastos de los Hogares. En esta fase se caracteriza el comportamiento de los principales factores de riesgo asociados a las ENT y se aplican métodos cuantitativos multivariados y técnicas de aprendizaje automático para identificar patrones de comportamiento y determinar las variables más relevantes que inciden en la productividad laboral de la PEA. De este modo, el análisis nacional se apoya conceptualmente en la evidencia estructural identificada en la etapa internacional, pero se enfoca en capturar el impacto empírico específico de las ENT en el contexto mexicano post-pandemia.

1.1 Planteamiento del problema

1.1.1 Panorama de las ENT

Se definen como enfermedades no transmisibles (ENT) aquellas que no son causadas principalmente por una infección aguda y solo en 2021 representaron cerca del 75% del total de las muertes en el mundo (WHO, 2021a). Entre las principales ENT se incluyen enfermedades

cardíacas, accidentes cerebrovasculares, cáncer, enfermedades respiratorias crónicas o diabetes, y que frecuentemente dan como resultado una necesidad de tratamiento y cuidados a largo plazo. Si bien las ENT pueden resultar de la interacción de factores genéticos y ambientales, la predisposición para padecerlas se hace evidente cuando el individuo se expone a un estilo de vida no propicio, aumento en el consumo de calorías, azúcares simples y grasas y reducción en la actividad física (Aguilar, 1999).

La tendencia hacia el envejecimiento de la población mundial está en constante aumento, mientras que el ritmo acelerado de vida y trabajo contribuye a esta dinámica. Como resultado, la incidencia de enfermedades crónicas sigue una trayectoria ascendente año tras año, con un correspondiente incremento en el gasto en salud (Baig *et al.*, 2017). Este fenómeno se ve exacerbado por los estilos de vida poco saludables, los cuales pueden generar discapacidad, deteriorar la calidad de vida individual, reducir la productividad y aumentar los costos asociados con la atención médica (Al-Qurishi *et al.*, 2015). En este contexto, es crucial abordar de manera integral los factores que contribuyen a estas tendencias, tanto a nivel individual como a nivel de políticas públicas, con el fin de promover una población más saludable y sostenible.

Desde hace poco más de 20 años, las ENT ocupan los primeros lugares en mortalidad, padecimientos como enfermedades del corazón, accidentes cerebrovasculares y diabetes *mellitus*, estando en el primer, segundo y noveno lugar respectivamente (WHO, 2020).

Es relevante mencionar que cuando existe una concentración de glucosa en el ayuno mayor o igual a 126 mg/dl, la probabilidad de contraer diabetes *mellitus* es aún mayor, así como el incremento de la probabilidad de sufrir las complicaciones crónicas de la enfermedad (CDC, 2022). Como parte de las complicaciones más frecuentes y de mayor impacto se encuentran el infarto al miocardio, la insuficiencia cardíaca, el infarto cerebral y la insuficiencia arterial en miembros inferiores. Otras complicaciones como el daño renal y ocular pueden aparecer varios años después de la aparición de la enfermedad.

Hablando de hipertensión arterial, las guías de práctica clínica para el diagnóstico y tratamiento desarrolladas por el Centro Nacional de Excelencia Tecnológica en Salud conceptualizan la hipertensión arterial sistémica (HTA) como un padecimiento multifactorial que se caracteriza por la elevación persistente de las cifras de presión arterial sistólica y diastólica a cifras mayores o iguales a 140/90 mmHg (CENETEC, 2021). Existe una asociación entre la

tensión arterial sistólica (TAS) y el riesgo de desarrollar un evento cardiovascular cuando la TAS es de forma constante mayor que 115 mmHg (Mancia *et al.*, 2013). Una elevación de 2 mm/Hg en la tensión arterial sistólica (TAS) se asocia con un aumento de 7% con el riesgo de mortalidad por cardiopatía isquémica y 10% en el riesgo de mortalidad por accidente cerebrovascular, por lo que la HTA es considerada como un importante factor de riesgo de infarto al miocardio, ictus isquémico y hemorrágico, insuficiencia cardíaca, enfermedad renal crónica, deterioro cognitivo y muerte prematura, todos considerados dentro de las enfermedades cardiovasculares (NICE, 2022).

De acuerdo con Shamah *et al.* (2020), el término enfermedad cardiovascular se refiere a enfermedades del corazón y vasos sanguíneos, producidas principalmente por la acumulación de placas de colesterol en las paredes de las arterias, haciéndolas estrechas. Otras causas de las enfermedades cardiovasculares son: niveles elevados de ácido úrico, diabetes, obesidad, tabaquismo o estrés, y se pueden catalogar como: cardiopatía coronaria, insuficiencia cardíaca, arritmias, valvulopatías, hipertensión, accidente vascular cerebral o cardiopatía congénita (WHO, 2021c).

La mortalidad en las enfermedades crónico-respiratorias de acuerdo con Scichilone *et al.* (2015) se refiere a las afecciones respiratorias crónicas más comunes, que se caracterizan por inflamación persistente de las vías respiratorias y obstrucción bronquial y de acuerdo con la Organización Mundial de la Salud (WHO, 2022a) es la tercera causa de muerte en el mundo causando más de 3.23 millones de muertes en 2019. La principal causa es el consumo de tabaco y la predisposición genética, pero también existen otras causas frecuentemente de índole laboral como la exposición a polvo, gases o vapor (Chapman *et al.*, 2006).

Otra de las ENT relevantes es la diabetes *mellitus*, que es una enfermedad metabólica crónica caracterizada por niveles elevados de glucosa en la sangre, lo que conduce con el tiempo a daños graves en el corazón, los vasos sanguíneos, los ojos, los riñones y los nervios, siendo la más común la diabetes tipo 2, generalmente adquirida en adultos, que ocurre cuando el cuerpo se vuelve resistente a la insulina o no produce suficiente cantidad de dicha hormona (WHO, 2022b). La diabetes no tratada conduce a lesiones multiorgánicas y sistémicas, incluidos el corazón, los riñones, los nervios y los vasos sanguíneos, lo que deteriora la calidad de vida y aumenta la tasa de mortalidad causada por las complicaciones de la diabetes como lo son: la enfermedad cardiovascular, enfermedades renales, neoplasias e inclusive Alzheimer (Li *et al.*, 2019).

Respecto a las neoplasias o tumorales, es la segunda causa de muerte en el mundo, siendo los más comunes el de pulmón, próstata, colon, estómago e hígado en hombres y en mujeres el cáncer de mama, colon, pulmón, cervicouterino y tiroides (WHO, 2022c).

Shamah *et al.* (2020) mostraron que los principales motivos de consulta por grupos de edad para el grupo de 20 a 49 años son enfermedades y síntomas respiratorios agudos, seguidos de diabetes, enfermedad cardiovascular y obesidad. En la Tabla 1 se muestra un resumen de la información sobre las principales ENT y su impacto en la sociedad.

Diversos investigadores en conjunto con la Organización Panamericana de la Salud (OPS, 2022) estimaron que desde el 2010 hasta el 2030 se espera un gasto de 47 billones de dólares en el tratamiento de las ENT, lo que es equivalente a una pérdida del 48% del PIB global anual e impacta en el 4% del PIB anual en países de medianos a bajos ingresos.

Tabla 1.

Principales Enfermedades no transmisibles – generalidades y estadísticas

Enfermedad	Generalidades	Estadísticas
Diabetes	La diabetes es una enfermedad crónica y compleja que demanda atención médica constante. Su manejo efectivo implica más que simplemente controlar los niveles de glucosa en sangre (WHO, 2021b). Para mitigar su impacto, es necesario adoptar un enfoque integral que aborde diversos aspectos de la salud y el bienestar del paciente.	En 2017 (IDF, 2017) se estimó que, en todo el mundo, aproximadamente 425 millones de adultos, de entre 20 y 79 años, vivían con diabetes. Para 2045, se espera que este número aumente a 629 millones.
Obesidad	Es el mayor problema de salud global y la principal causa del aumento de la morbilidad y la mortalidad. Es uno de los mayores desafíos de salud pública de este siglo (Alwashmi <i>et al.</i> , 2019). La promoción de la AF y el cambio de estilo de vida es crucial (Wong <i>et al.</i> , 2014),	Las reducciones en el peso corporal (5-10%) y el IMC reducen los riesgos para la salud asociados con las enfermedades crónicas (Garvey & Arathuzik, 2016).
Hipertensión	Uno de los factores de riesgo más importantes para la mortalidad por todas las causas y la principal causa de mortalidad, morbilidad y	La presión arterial alta afecta a mil millones de personas, y se proyecta que para 2025, su prevalencia será superior

	discapacidad cardiovascular (Mancia <i>et al.</i> , 2013).	a 1.500 millones de personas en el mundo (Parati <i>et al.</i> , 2017).
Enfermedades cardiovasculares	La frecuencia cardíaca se considera el signo vital estándar que indica cambios en los ciclos cardíacos (Sunny <i>et al.</i> , 2022). Una frecuencia cardíaca alta en reposo se correlaciona con un mayor riesgo de enfermedad arterial coronaria e insuficiencia cardíaca (De la Torre Díez <i>et al.</i> , 2016). La mayoría de sus factores de riesgo, como la presión arterial alta, el perfil lipídico anormal, el tabaquismo y la inactividad física, deben prevenirse (Zhang <i>et al.</i> , 2017).	Las enfermedades cardiovasculares son la principal causa de muerte en el mundo, responsables de aproximadamente 17.9 millones de muertes al año, lo que representa el 32% de todas las muertes globales (WHO, 2020)
Accidente cerebro vascular	Se clasifica como una de las principales causas de muerte y discapacidad en todo el mundo. Varias limitaciones como la accesibilidad a los exámenes clínicos pueden ser relevantes debido a restricciones geográficas o de recursos (Chen & Sawan, 2021).	El 80% de los accidentes cerebrovasculares se pueden prevenir atendiendo los factores de riesgo como la presión arterial alta (Benjamin <i>et al.</i> , 2019).
Neoplásicas	Es la principal causa de muerte en el mundo siendo los principales el cáncer de mama, pulmón, colono rectal y próstata (WHO, 2022d)	82% de las muertes prematuras por ENT ocurren en países de ingresos bajos y medianos (Williams <i>et al.</i> , 2018). Es la principal causa de muerte en cerca de 50 países de altos ingresos y ocupa el segundo lugar en otros 40 países en desarrollo (Pinerós <i>et al.</i> , 2017).
Enfermedades Respiratorias	Las enfermedades respiratorias representan un importante problema de salud pública debido a su alta frecuencia, carga económica e impacto social y la mayoría de las afecciones respiratorias graves son incurables; ese tipo de enfermedades incluyen asma, enfermedad pulmonar obstructiva crónica (EPOC), enfermedad pulmonar intersticial, fibrosis quística y la hipertensión arterial pulmonar contribuyen sustancialmente a	En todo el continente americano en 2018 las enfermedades respiratorias causaron más de 530 mil defunciones (OPS, 2021).

	la mortalidad y la carga relacionada con la enfermedad (Rose <i>et al.</i> , 2017).	
--	---	--

Fuente: Elaboración propia

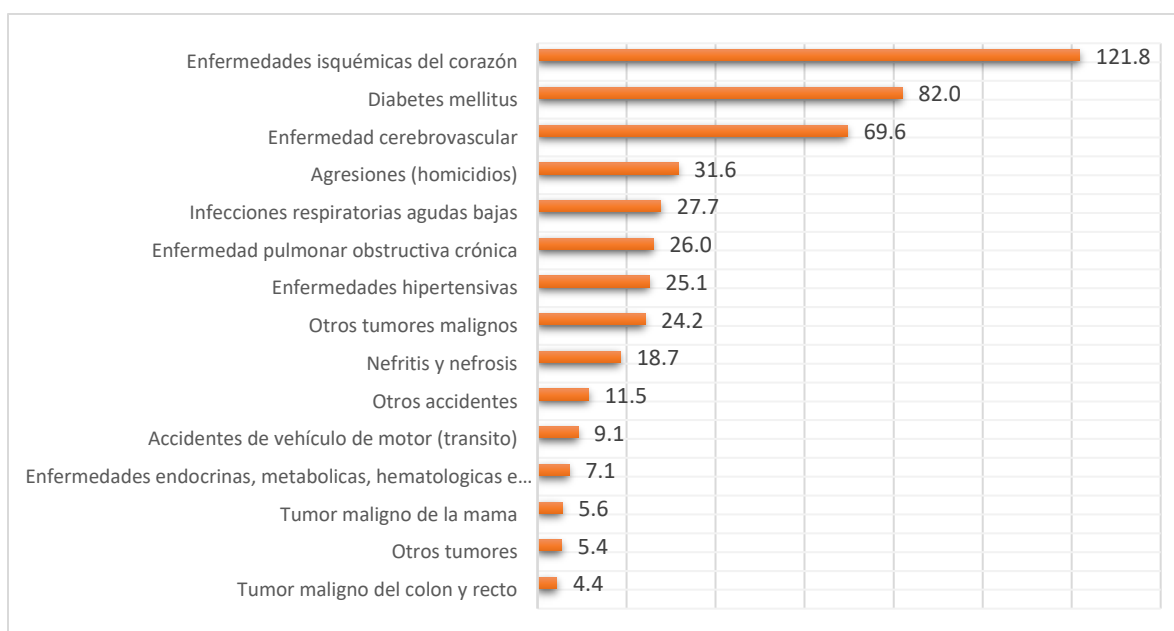
Jaspers *et al.* (2015) identificaron que en el 60% de las encuestas realizadas, el gasto de bolsillo representa del 20% al 30% de sus ingresos. Chaker *et al.* (2015) estimaron que en los Estados Unidos de América existe una pérdida de productividad por persona por año en un intervalo de 8,539 a 10,175 dólares por enfermedades cardiovasculares y de 1,962,314 dólares anuales debido a la diabetes *mellitus* con una mortalidad prematura del 49%. Esta evidencia nos muestra la necesidad de un sistema de salud que dé seguimiento al personal con enfermedades crónicas junto con un enfoque de prevención basado en un abordaje en el monitoreo continuo de la salud de los empleados (Chaker *et al.*, 2015).

1.1.2 Diagnóstico situacional en México

En el caso de México, para 2019, de las primeras 16 causas más frecuentes de muerte están dominadas por las ENT, siendo las enfermedades cardiovasculares, las relacionadas con la diabetes *mellitus*, las neoplasias y las enfermedades respiratorias las más relevantes (Ver Figura 1).

Figura 1.

Principales causas de muerte en México. Defunciones por cada 100,000 habitantes. 2019.



Fuente: Sistema de información de la Secretaría de Salud. DGIS. Elaboración propia.

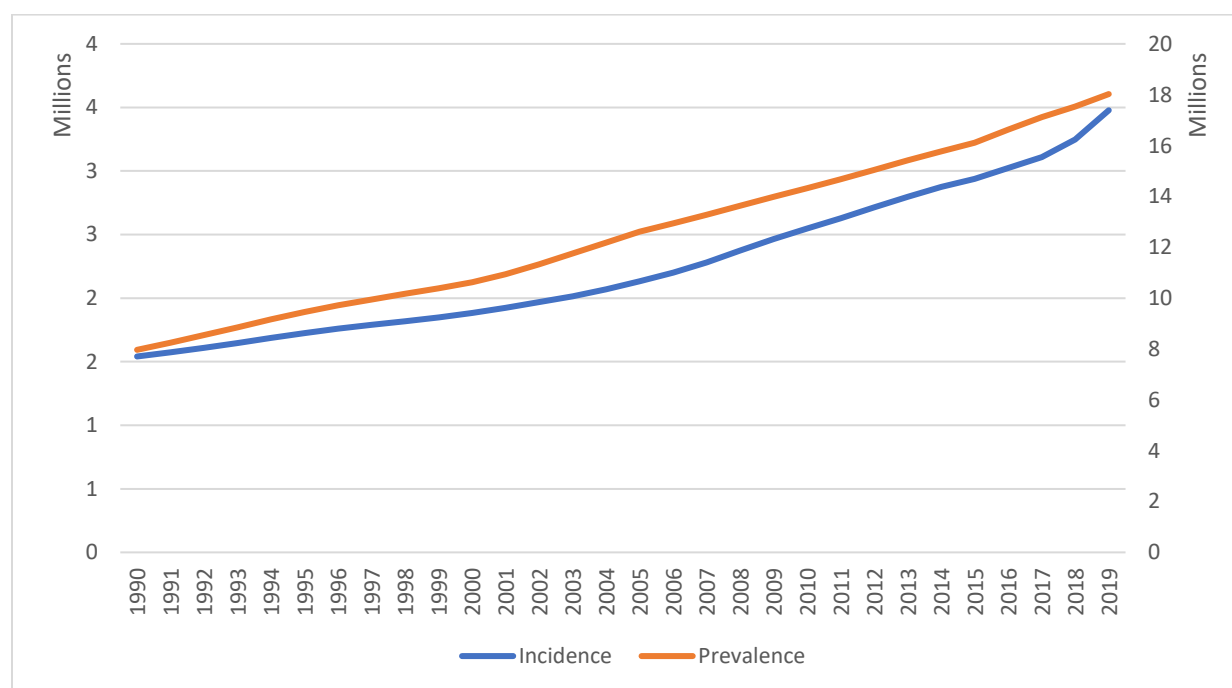
Aguilar *et al.* (2005) argumentaron que un alto porcentaje de la mortalidad en México tiene una gran influencia derivada de los primeros años de la vida adulta cuando existen concentraciones anormales de colesterol, triglicéridos, colesterol LDL y ácido úrico, lo que puede significar una incidencia de enfermedades.

Incidencia, prevalencia y mortalidad de las ENT en México

De acuerdo con el sistema epidemiológico de México (2023), la incidencia se define como el número de casos de enfermedad que comienzan o de personas que se enferman durante un período dado en una población determinada. La prevalencia es el número de casos existentes en una población en un momento determinado, sin distinguir si son casos nuevos o antiguos. En México, la incidencia y prevalencia de ENT han presentado un crecimiento acelerado en los últimos años (Ver Figura 2).

Figura 2

Incidencia y prevalencia de las ENT en México



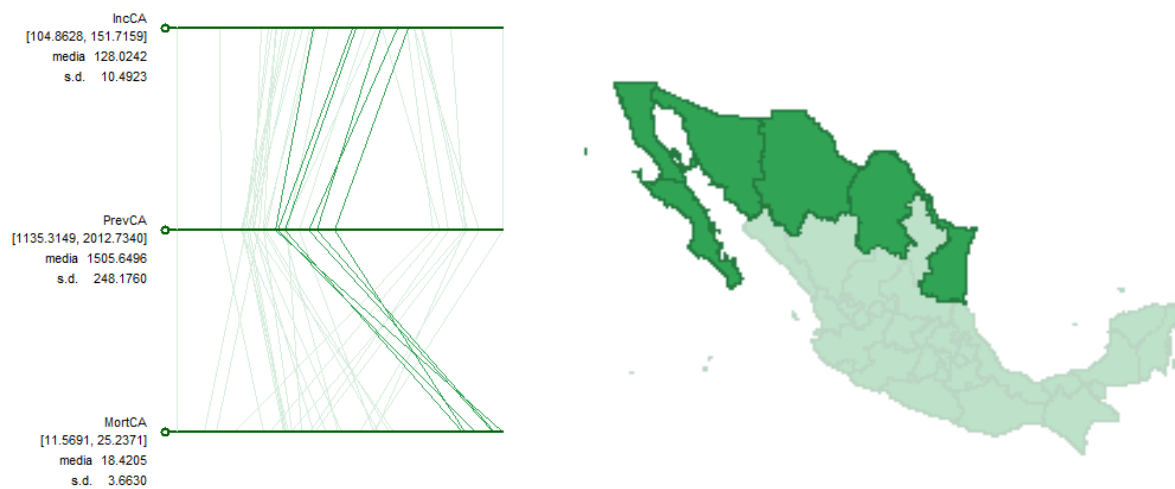
Fuente: IHME, 2020. Elaboración propia.

Por su parte, la tasa de mortalidad por enfermedades cardiovasculares muestra que el promedio de fallecimientos por ENT es de 18.42 por cada cien mil habitantes con una desviación estándar de 3.66 casos. Los datos muestran que los estados al norte de México muestran el mayor

número de casos por cien mil habitantes, es de notar que, si bien los estados de Baja California, Baja California Sur, Sonora, Chihuahua, Coahuila y Tamaulipas se ubicaban alrededor de la media de los datos en la incidencia, así como en la prevalencia, esto no se refleja en los números de mortalidad (Ver Figura 3).

Figura 3

Mortalidad en México de enfermedades cardiovasculares – coordenadas paralelas, 2019



Fuente: IHME, 2020. Elaboración propia.

Por cada 100,000 casos de incidencia o permanencia, hay una mortalidad de 136 y 1 casos respectivamente (IHME, 2020).

Dentro de las ENT en México mantienen una media menor de las enfermedades cardiovasculares con un valor de 7.70 casos por cada cien mil habitantes y una desviación estándar de 1.52. Los estados de Sonora y Chihuahua muestran el mayor número de casos con valores de 10.80 y 10.98 respectivamente, cabe mencionar que dichos estados tanto en la prevalencia como en la incidencia aparecían por debajo del tercer cuartil 75 (IHME, 2020) (Ver Figura 4).

Adicionalmente, por cada 100,000 casos de incidencia (IncRD) o permanencia (PrevDB), hay una mortalidad de apenas 12 y 4 casos respectivamente (IHME, 2020).

Figura 4

Mortalidad en enfermedades crónico-respiratorias en México – coordenadas paralelas, 2019

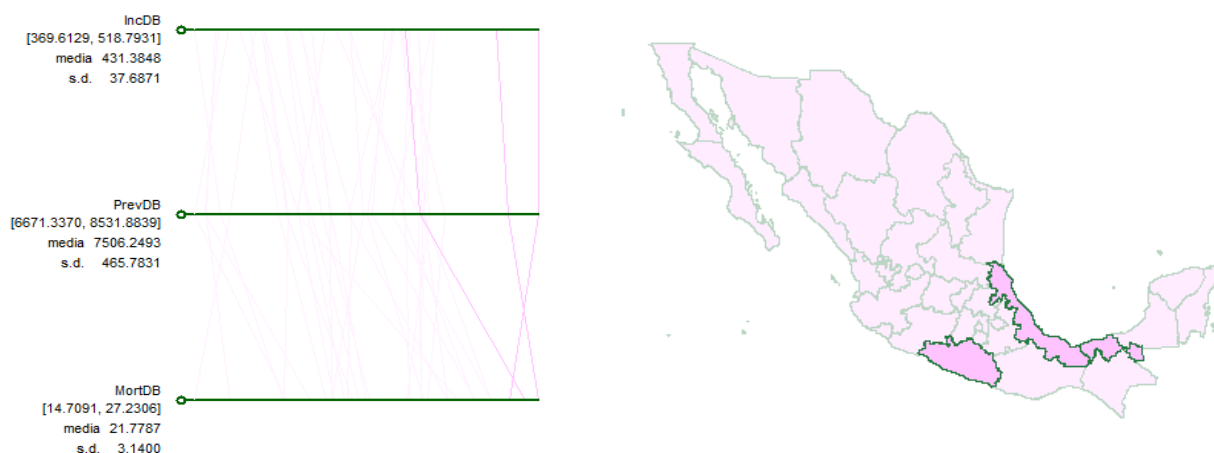


Fuente: IHME, 2020. Elaboración propia.

Respecto a la mortalidad en las enfermedades derivadas de la diabetes, se muestra un promedio en México de 21.78 casos por cada cien mil habitantes y una desviación estándar de 3.14, como se muestra en la Figura 5. Los estados que muestran un mayor número de casos son Guerrero (26.78), Tabasco (26.2) y Veracruz (27.23), siendo los últimos dos estados con mayores casos tanto en incidencia como en prevalencia (IHME, 2020) (Ver Figura 5).

Figura 5

Mortalidad en diabetes en México – coordenadas paralelas, 2019



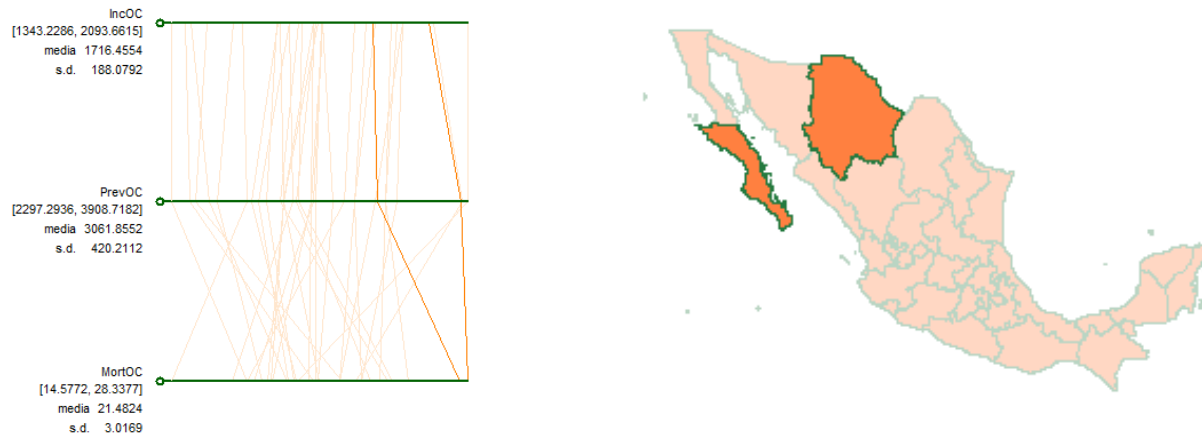
Fuente: IHME, 2020. Elaboración propia.

Por cada 100,000 casos de incidencia o permanencia, hay una mortalidad de 73 y 6 casos respectivamente (IHME, 2020). Esto es porque existe una gran cantidad de enfermedades que desencadenan la diabetes, como enfermedades renales o las mismas enfermedades cardiovasculares, y al final la causa de muerte podría ser variada y no necesariamente explicada solamente por la diabetes.

La cantidad de casos promedio en México de todo tipo de enfermedades oncológicas se ubica en 21.48 casos con una desviación estándar de 3.02. Los estados con el mayor número de casos son Chihuahua (28) y Baja California Sur (28.33), siendo este último estado el único que se mantiene en el valor superior en los casos tanto en la incidencia como en la prevalencia como en la misma mortalidad. Cabe destacar que los estados de Jalisco y México, que se encontraban con el mayor número de casos de incidencia y prevalencia, se muestran ahora dentro del percentil 75 (IHME, 2020) (Ver Figura 6).

Figura 6

Mortalidad en enfermedades neoplásicas en México – coordenadas paralelas, 2019



Fuente: IHME, 2020. Elaboración propia.

Por cada 100,000 casos de incidencia o prevalencia, hay una mortalidad de 8 y 4 casos respectivamente (IHME, 2020).

De acuerdo con la OMS (2022), una gran parte del impacto de las ENT pudo haberse evitado al realizar actividades de prevención en salud como el fomento de la AF, ya que de acuerdo con el INEGI (2020), el 57.9% de la población de 18 años y más no realiza ningún deporte o

ejercicio físico en su tiempo libre, proporción que no ha variado significativamente desde 2013. En adultos, 17.3% de la población no cumplió con los criterios mínimos establecidos por la OMS para ser considerados como físicamente activos; los adultos de 20 a 69 años reportaron un sedentarismo en promedio de 214.8 minutos al día, siendo mayor en hombres. Además, se observó que el grupo de 20 a 29 años pasó mayor tiempo sentado en comparación con los otros grupos (Shamah *et al.*, 2020), aspecto que se ha relacionado con el desarrollo de enfermedades crónicas como obesidad, diabetes *mellitus* e hipertensión.

Por otro lado, en México el gasto total en salud en porcentaje del producto interno bruto representó el 5.6 % en 2019 (INEGI, 2020), lejos de la media de 8.8 % de los países que son parte de la Organización para el Desarrollo Económico (OCDE, 2019). La Asociación Mexicana de la Industria de Seguros (AMIS, 2018) ha indicado que solo el 7% de la población tiene acceso a seguro de gastos médicos mayores y esto también se ve reflejado a través del gasto de bolsillo¹ de las personas en México que refleja el 41% del gasto total en salud (OCDE, 2019) y un crecimiento constante superior respecto a lo observado en los países de la OCDE. Es importante un mayor enfoque en este ámbito, pues el crecimiento de la prevalencia en las ENT puede llevar a un hogar a incurrir en gastos catastróficos por motivos de salud, generando pérdidas en el bienestar social (Rodríguez-Aguilar, 2022).

1.1.3 El estado de salud de los trabajadores y su productividad laboral

Debido al aumento de la esperanza de vida de los mexicanos, de 40 años en 1970 a 76.5 años en 2020 (INEGI, 2020), el cambio en el estilo de vida, tanto en alimentación como en actividad física, y el establecimiento de programas para la prevención y control de las ENT son una necesidad imperante. El diagnóstico temprano juega un papel importante en el impacto del tratamiento, la mejora de la calidad de vida, el tiempo de intervención en salud que realiza el estado. De igual forma, hay efectos económicos debido a afectaciones en la productividad de los trabajadores atribuibles a su estado de salud.

Los efectos de la promoción en la salud pueden llevar a las empresas a prevenir problemas de ausentismo, presentismo, bajo desempeño laboral y pérdida de productividad. Diversas investigaciones han demostrado que los empleados con bajos niveles de actividad física y

¹ El gasto de bolsillo se define como los pagos realizados por el hogar al momento de recibir servicios de salud.
Fuente: Secretaría de Salud, 2018.

comportamiento sedentario son menos productivos en el trabajo, tienen altos niveles de presentismo y una menor capacidad de trabajo (Rongen *et al.*, 2013). Cuando la persona experimenta constantemente una gran carga de trabajo durante un período más largo y no puede recuperarse, entonces el trabajo puede tener efectos negativos prolongados y puede causar enfermedades graves. Muchas descripciones de trabajo han cambiado porque el trabajo ha pasado del trabajo manual a tareas predominantemente sedentarias. Trabajar largas horas durante largos períodos de tiempo se asocia con depresión, ansiedad, trastornos del sueño, enfermedades cardíacas crónicas y estrés (Chaparro Narváez y Guerrero, 2001).

Para el caso de México, de acuerdo con lo marcado en la Ley Federal del Trabajo, se establece que los empleados deben estar registrados en el Instituto Mexicano del Seguro Social (IMSS), que en total representa una afiliación del 51% de los mexicanos (INEGI, 2020). De igual forma, el gobierno publicó la Norma Oficial Mexicana NOM-035-STPS-2018 (2022) que tiene como objetivo establecer los elementos para identificar, analizar y prevenir los factores de riesgo psicosocial, así como promover un entorno organizacional favorable en los centros de trabajo.

Sin embargo, a pesar de los esfuerzos realizados para regular la manera de enfrentar las ENT, las enfermedades cardiovasculares, la diabetes, el cáncer y las enfermedades respiratorias crónicas han tenido un impacto significativo en la productividad laboral en México. Según la Encuesta Nacional de Salud y Nutrición (ENSANUT, 2022), las enfermedades crónicas no transmisibles son responsables del 25% del ausentismo laboral en México.

Un informe de la OCDE (2019) estima que el presentismo laboral debido a enfermedades crónicas representa alrededor del 25% de la pérdida de productividad relacionada con la salud en México. Por su parte, el Instituto Mexicano del Seguro Social (IMSS) reportó que el ausentismo y el presentismo asociados a enfermedades crónicas representaron un impacto equivalente al 3.7 % del PIB nacional (IMSS, 2022). Dentro de estas enfermedades, la diabetes *mellitus* destaca por su impacto económico, representando por sí sola aproximadamente el 6 % del PIB (Lopez-Sanchez *et al.*, 2022).

La creciente prevalencia de condiciones como diabetes, hipertensión, cáncer y enfermedades cardiovasculares genera costos directos en atención médica y costos indirectos asociados a ausentismo, reducción en el desempeño laboral y pérdidas de capital humano capacitado. En México, este panorama se agudiza, ya que las empresas deben afrontar

considerables gastos relacionados con incapacidades prolongadas y reducción en la eficiencia operativa, lo cual implica un desafío estratégico importante para el sector productivo nacional.

1.2 Justificación del Tema

En los últimos años han sucedido cambios importantes en la concepción de la salud y la enfermedad humana de un modelo de enfermedad a un modelo de salud. Bandura (1998) muestra que es tan significativo hablar de niveles de vitalidad como de grados de deterioro. El modelo de salud, por tanto, se centra tanto en la promoción de la salud como en la prevención de enfermedades y los hábitos de vida, los cuales ejercen un gran impacto en la calidad de la salud humana.

Desde una perspectiva económica, estos cambios conceptuales cobran especial relevancia en el ámbito laboral. De acuerdo con la Comisión Económica para América Latina y el Caribe (CEPAL, 2016), la productividad laboral se define como la relación entre la producción generada y la cantidad de trabajo utilizada, generalmente medida como producto por trabajador o por hora trabajada. Esta medida refleja la eficiencia con la que el trabajo se transforma en valor agregado y constituye un determinante clave del crecimiento económico, la competitividad y el bienestar social. En este sentido, la productividad laboral no depende únicamente de factores tecnológicos u organizacionales, sino también de la salud física y mental de la población económicamente activa, que condiciona su capacidad de desempeño y permanencia en el mercado laboral.

Tanto la prevención como la promoción en favor de la salud comprenden elementos para evitar o retrasar los problemas de salud y disminuir la carga que generan las enfermedades. La creciente actividad e interés en los campos de la promoción y prevención de la salud en el entorno laboral, impulsados por una mayor conciencia sobre los costos de la atención médica y el presentismo, junto con el envejecimiento de la población, están generando una demanda cada vez mayor de programas de promoción de la salud, especialmente dirigidos a la población trabajadora (Chapman, 2005). Sin embargo, las preguntas sobre el rendimiento económico asociado con este tipo de esfuerzos también son una realidad continua, pues todas las organizaciones laborales deben justificar sus decisiones de inversión de recursos en la promoción y prevención de la salud en términos principalmente económicos.

La promoción de la salud no solo abarca acciones dirigidas a fortalecer las habilidades y capacidades de los individuos, sino también acciones dirigidas a cambiar los determinantes sociales, ambientales y económicos de la salud para optimizar su impacto positivo en la salud pública y personal. La promoción de la salud es el proceso de capacitar a las personas, individual y colectivamente, para aumentar el control sobre los determinantes de la salud y, por lo tanto, mejorar su salud.

De acuerdo con Zhang *et al.* (2017), la inactividad física y el comportamiento sedentario son riesgos para la salud de los empleados. Ya que los adultos que laboran con enfermedades, principalmente las ENT, presentan un menor rendimiento laboral, ingresos más bajos y pérdida de empleo debido a la disminución de la capacidad laboral relacionada con la enfermedad. También representa una carga económica para los empleadores, ya que las empresas a menudo sufren pérdidas financieras significativas debido a enfermedades y la mala salud de sus empleados. Por ende, el cuidado de la salud es una responsabilidad compartida en donde empleados y empleadores precisarían ocuparse de salvaguardar la integridad de los recursos humanos, por lo que acciones encaminadas a la promoción de entornos saludables para el individuo, con el fin de establecer un concepto amplio y positivo de salud, pudieran exaltar un beneficio para todos los involucrados. Beneficio que podría implicar un incremento de la productividad laboral.

Es posible considerar también que los estilos de vida son un elemento clave en la salud del trabajador, de esta manera, las enfermedades que puedan presentarse se han ligado estrechamente al estilo de que lleve el individuo, y tal como muestra en sus investigaciones epidemiológicas de Reynoso y Seligson (2002) donde demostraron que el crecimiento en la tasa de mortalidad en los últimos años son los factores de riesgo y factores positivos para la salud del trabajador.

El impacto económico de las ENT es sustancial y creciente. De acuerdo con datos de la Organización Mundial de la Salud (2022), se estima que los costos económicos globales derivados de las ENT alcanzan los 47 billones de dólares en el periodo 2011–2030, considerando tanto los costos directos en atención médica como los indirectos relacionados con pérdida de productividad, incapacidad y mortalidad prematura. A nivel de empresas, estas enfermedades afectan negativamente el retorno de la inversión en capital humano, incrementan los gastos por seguros y disminuyen la competitividad. En América Latina, la Organización Panamericana para la Salud (2021) ha advertido que los países de ingresos medios enfrentan un doble desafío: un aumento

sostenido de la prevalencia de ENT y una débil infraestructura preventiva que limita su detección oportuna y tratamiento adecuado, lo cual incrementa los costos a largo plazo.

En este contexto, el cuidado de la salud debe entenderse como una responsabilidad compartida entre empleadores y empleados. Las acciones orientadas a la creación de entornos saludables son necesarias para resguardar el capital humano, fomentar una cultura organizacional proactiva y lograr beneficios para todos los involucrados, incluyendo incrementos sostenibles en la productividad (Kirsten, 2010).

Si bien existen varios métodos que pueden ser útiles al momento de la implementación de un programa de salud empresarial, para que se pueda realizar una buena gestión de la salud en el trabajador es importante considerar un buen diseño y metodología sistemática de un plan estratégico para la empresa considerando los elementos emocionales, físicos, sociales, espirituales e intelectuales y llegar a una gestión adecuada. Promover condiciones de salud y hábitos de vida saludables a fin de lograr el objetivo de tener trabajadores con una cultura dentro de la organización que refuerce y defienda los valores éticos y mejore la productividad, eficiencia y competitividad de la empresa, tal como lo afirma Marqués (2020). De igual forma, no existe un mecanismo que analice los resultados en las inversiones de tiempo y dinero enfocados en mostrar las ventajas de una vida saludable o los programas que se puedan sugerir enfocados en salud dentro y fuera de los lugares de trabajo.

La presente investigación adquiere relevancia no solo por su contribución a la comprensión del impacto de las ENT en la productividad laboral, sino también por su potencial para aportar evidencia empírica que apoye el diseño de estrategias más eficientes de salud ocupacional. En particular, se busca contribuir a cerrar la brecha existente en la literatura sobre la población económicamente activa en México, donde los estudios que vinculan de manera directa las ENT, los hábitos de salud y la productividad laboral siguen siendo escasos.

En la Tabla 2 se muestran diversos artículos que muestran iniciativas realizadas por empresas y diversos métodos para abordar diferentes problemáticas como el sedentarismo, la obesidad, el síndrome metabólico, la enfermedad cardiovascular o el estrés.

Tabla 2.

Iniciativas basadas en evidencia sobre salud ocupacional y productividad laboral.

Iniciativa Programa	Variables medidas	Método	Resultados Relevantes	Referencia
Programa de salud integral en empresa mexicana	IMC, triglicéridos, presión arterial	Intervención de estilo de vida en el lugar de trabajo, datos antes/después	Disminución significativa de IMC y triglicéridos tras 1 año	Méndez-Hernández <i>et al.</i> (2012)
Bienestar corporativo en países de ingresos bajos y medios. (México-India)	Actitudes sobre salud ocupacional, ROI	Estudio cualitativo semiestructurado	Barreras culturales y financieras en implementación	Wipfli <i>et al.</i> (2018)
Bienestar en universidad pública Malasia	Presión sistólica y diastólica	Estudio de cohorte a 6 años	Reducción anual significativa de presión arterial	Eng <i>et al.</i> (2016)
CDSMP online en México y Perú	Depresión (PHQ-8), actividad física	Ensayo controlado (n=114)	Reducción significativa de síntomas depresivos; aumento PA	Ramírez <i>et al.</i> (2023)
Salud laboral holística en el ambiente laboral	Salud holística, burnout, productividad	Encuesta global (30,000 empleados)	Empleados saludables ≈50% más innovadores y menos burnout	Brassey <i>et al.</i> (2023)
Bienestar laboral basada en actividad física	Capacidad laboral, estado físico, fuerza muscular, síntomas musculoesqueléticos	Revisión sistemática de estudios	Mejora en capacidad laboral, sin información sólida sobre costo-efectividad	Anderson <i>et al.</i> (2019)
Evaluaciones económicas de intervenciones en salud ocupacional	Productividad relacionada con salud	Revisión sistemática	Heterogeneidad en medición, escasa evidencia de valor económico	Song & Baicker (2019)
Programa de bienestar en empresas minoristas en E. U.	Comportamientos salud, indicadores clínicos, utilización médica, empleo	Ensayo controlado	Aumento de conductas saludables, sin efecto significativo en salud clínica ni productividad tras 18 meses	Proper & Van Oostrom (2023)
Programa de vuelta al trabajo para trastornos musculoesquelético s/psicológicos	Economicidad de retorno al empleo	Evaluación costo-efectividad	Costos reducidos comparados con tradicional; mejor retorno al empleo	Goetzel <i>et al.</i> (2011)

Nutrición y actividad física en el trabajo	Absentismo, desempeño, rendimiento, productividad	Revisión sistemática (RCT y no RCT)	14 estudios con efectos positivos; mejoras en absentismo y productividad	Jensen <i>et al.</i> (2024)
Intervenciones de salud mental laboral	ROI, ausentismo, rotación, productividad	Estudio con 100 empresas	ROI medio 4:1; menor ausentismo y rotación, mejor productividad	Patel & Saini (2024)

Fuente: Elaboración propia.

El estado del arte en la investigación sobre la productividad laboral y las ENT ha documentado que la prevalencia de estas condiciones tiene efectos negativos significativos en la productividad atribuibles al ausentismo, la reducción en el rendimiento laboral y los costos directos asociados a la atención médica. Una amplia gama de estudios empíricos ha demostrado que intervenciones implementadas en entornos laborales enfocadas en la promoción de hábitos saludables como la actividad física, la alimentación equilibrada y el manejo del estrés pueden contribuir a la mejora de indicadores clínicos como el índice de masa corporal, la presión arterial y los niveles de triglicéridos, así como a una disminución del gasto médico institucional.

Investigaciones recientes han incorporado enfoques analíticos avanzados como el uso de modelos predictivos y técnicas de aprendizaje automático, los cuales han facilitado la identificación de patrones de comportamiento y factores de riesgo asociados con la aparición de ENT y su impacto en el desempeño económico de las personas trabajadoras (Domínguez-Miranda & Rodríguez-Aguilar, 2024a; Domínguez-Miranda *et al.*, 2026). No obstante, persisten vacíos importantes en la literatura, particularmente en lo relativo a la evaluación del retorno sobre la inversión de estos programas de salud empresariales. Esta limitación es especialmente relevante en países en desarrollo como México, donde los estudios que cuantifican de manera directa el impacto económico de las ENT en entornos laborales son prácticamente inexistentes.

Desde el punto de vista metodológico, gran parte de los estudios existentes se basa en modelos de regresión lineal tradicional para estimar la relación entre condiciones de salud y productividad laboral. Si bien estos enfoques resultan útiles como aproximaciones iniciales, presentan limitaciones importantes cuando se aplican a fenómenos complejos como las ENT. En primer lugar, la relación entre salud y productividad está sujeta a problemas de endogeneidad, derivados de la posible causalidad inversa y de la presencia de variables omitidas, ya que la salud influye en la productividad, pero el nivel de ingresos, las condiciones laborales y el tipo de empleo

también afectan el estado de salud. En segundo lugar, el uso de múltiples variables correlacionadas como los hábitos de salud, indicadores clínicos, características socioeconómicas y laborales, genera problemas de multicolinealidad y alta dimensionalidad, que limitan la interpretabilidad y estabilidad de los estimadores en modelos lineales estándar. Asimismo, estas relaciones suelen ser no lineales y heterogéneas, lo que dificulta su adecuada representación mediante especificaciones paramétricas simples.

En este contexto, la adopción de un enfoque metodológico híbrido resulta especialmente pertinente para los objetivos de esta investigación. Por un lado, el uso de modelos de regresión penalizada, como *Lasso*, *Ridge* y redes elásticas, permite manejar conjuntos de datos con alta dimensionalidad y correlación entre variables, facilitando la selección de los factores más relevantes asociados con la mortalidad y prevalencia de ENT, así como la reducción del ruido estadístico. Por otro lado, la implementación de modelos de ecuaciones estructurales (SEM) posibilita capturar relaciones causales complejas entre constructos latentes, hábitos de salud, factores socioeconómicos y condiciones laborales, incorporando efectos directos e indirectos que no pueden ser identificados mediante enfoques puramente reduccionistas.

Finalmente, la estimación del impacto causal de las ENT sobre la productividad laboral requiere metodologías que permitan abordar de manera explícita los problemas de endogeneidad inherentes al uso de datos observacionales. En este sentido, el enfoque de *Double Machine Learning* (DML) ofrece una alternativa robusta, al combinar modelos de aprendizaje automático para el control flexible de variables de confusión con técnicas econométricas para la estimación de efectos causales. Esta aproximación permite obtener estimadores insesgados bajo supuestos más débiles que los modelos tradicionales, resultando particularmente adecuada para el análisis de grandes bases de datos como la ENSANUT y la ENIGH.

La presente investigación contribuye de manera novedosa al combinar de forma coherente herramientas estadísticas clásicas, técnicas de aprendizaje automático y métodos de inferencia causal. Esta estrategia metodológica integral permite:

1. Caracterizar estructuralmente el fenómeno de las ENT en distintos niveles de análisis.
2. Identificar los factores de mayor relevancia asociados a su prevalencia y severidad.
3. Modelar relaciones causales complejas entre hábitos de salud y ENT.

4. Estimar de manera rigurosa el impacto económico de las ENT sobre la productividad laboral.

De este modo, se genera evidencia empírica sólida y contextualizada que resulta útil para la toma de decisiones en materia de salud ocupacional y política pública en México.

1.3 Preguntas generales y específicas de la Investigación

Pregunta general de investigación

- ¿Cuál es el efecto de las ENT en la productividad laboral en México?

Preguntas específicas de investigación

- ¿Qué relación existe entre las dimensiones económicas, hábitos de salud y laborales con las ENT?
- ¿Cuál es el efecto de la adopción de hábitos de salud positivos de la PEA en la productividad laboral en México?

1.4 Objetivo de Investigación

- Estimar el efecto de las ENT en la productividad laboral en México.

Objetivos específicos

- Explorar la relación entre las dimensiones económicas, hábitos de salud y laborales con las ENT.
- Evaluar el efecto de los hábitos de salud positivos de la PEA en la productividad laboral en México.

1.5 Hipótesis

Hipótesis 1: Las ENT tienen un efecto negativo en la productividad laboral en México.

Hipótesis 2: Existe una asociación significativa entre las dimensiones económicas, hábitos de salud y laborales con las ENT.

Hipótesis 3: La adopción de hábitos de salud positivos de la PEA tiene un efecto positivo en la productividad laboral en México.

La Tabla 3 presenta la Matriz de Congruencia desarrollada para esta investigación, la cual permite establecer una alineación lógica entre los elementos clave del diseño metodológico: las preguntas de investigación, las hipótesis correspondientes, los constructos y variables implicadas, los objetivos específicos y las acciones metodológicas planeadas.

Tabla 3

Matriz de congruencia

Preguntas de Investigación	Hipótesis	Constructos - Variables	Objetivos	Acciones (Metodología)
¿Cuál es el efecto de las ENT en la productividad laboral en México?	Las ENT tienen un efecto negativo en la productividad laboral en México.	ENT Productividad	Determinar el efecto de las ENT en la productividad laboral en México.	• Elegir y analizar las bases de datos relevantes para la modelación. • Identificar los factores más relevantes que afectan a la productividad y que puedan ser susceptibles de mejora en los hábitos saludables. • Aplicar herramientas econométricas y de aprendizaje automático, como el <i>Double Machine Learning</i>, para estimar el impacto causal de las ENT sobre la productividad, integrando indicadores relacionados con dimensiones económicas, alimentarias y de hábitos de salud.
¿Qué relación existe entre las dimensiones económicas, hábitos de salud y laborales con las ENT?	Existe una asociación significativa entre las dimensiones económicas, hábitos de salud y laborales con las ENT.	Dimensiones de las ENT ENT	Explorar la relación entre las dimensiones económicas, hábitos de salud y laborales con las ENT.	
¿Cuál es el efecto de la adopción de hábitos de salud positivos de la PEA en la productividad laboral en México?	La adopción de hábitos de salud positivos de la PEA tiene un efecto positivo en la productividad laboral en México.	Hábitos en salud Productividad	Evaluar el efecto de los hábitos de salud positivos de la PEA en la productividad laboral en México.	

1.6 Descripción de Capítulos

El Capítulo I presenta el planteamiento del problema con un diagnóstico situacional en México y se argumenta la relevancia del estudio. Se formulan las preguntas de investigación, los objetivos y las hipótesis que guían el trabajo. Este capítulo culmina con una descripción general de la estructura del documento.

El Capítulo II ofrece una revisión exhaustiva de la literatura relacionada con las dimensiones determinantes de las ENT, los hábitos de salud y su relación con la productividad laboral. Se contextualizan las hipótesis en marcos conceptuales que articulan factores económicos, conductuales y sociales.

En el Capítulo III se detalla el diseño metodológico adoptado. Se describen las variables utilizadas, la estrategia de operacionalización y el enfoque estadístico aplicado. Se discute el proceso de selección de muestra, las fuentes de datos y se clasifican las pruebas realizadas en paramétricas y no paramétricas, vinculadas a los tres ejes de análisis: dimensiones relacionadas con las ENT en países en desarrollo, hábitos de salud en la PEA en México y su impacto en la productividad.

El Capítulo IV presenta los resultados empíricos organizados según los objetivos específicos. Se exploran los factores asociados a las ENT a nivel internacional mediante índices contruidos con análisis de componentes principales y modelos de regresión penalizados, se analiza la relación entre hábitos de salud y prevalencia de ENT en México utilizando modelos de ecuaciones estructurales y se estima el impacto de dichos hábitos sobre la productividad laboral a través de técnicas de inferencia causal como *Double Machine Learning*.

El Capítulo V contiene las conclusiones generales del estudio, derivadas del análisis cuantitativo y de la discusión conceptual.

En el Capítulo VI se explicitan las contribuciones del trabajo desde tres perspectivas: académica, empresarial y nacional, subrayando su potencial para informar políticas públicas y estrategias organizacionales.

El Capítulo VII reconoce las limitaciones metodológicas y empíricas del estudio y propone futuras líneas de investigación que permitan fortalecer y ampliar los hallazgos obtenidos.

Finalmente, el Capítulo VIII reúne las referencias bibliográficas utilizadas a lo largo del trabajo, seguido de los anexos correspondientes que complementan la información metodológica y empírica.

CAPÍTULO II. Revisión de la literatura

Este capítulo tiene como objetivo analizar los fundamentos conceptuales y empíricos que sustentan la relación entre las ENT, los hábitos de salud y la productividad laboral. A partir de un enfoque multidimensional, se revisan los principales determinantes de la salud, incluyendo factores biológicos, conductuales, sociales, ambientales y estructurales, así como su papel en la génesis y evolución de las ENT. Asimismo, se aborda la importancia de la prevención y promoción de la salud en contextos laborales, explorando las implicaciones que tienen los estilos de vida sobre el desempeño económico individual y colectivo. Por último, se examina el vínculo entre los hábitos de salud y la productividad laboral, destacando las contribuciones teóricas y empíricas previas, así como las principales líneas estratégicas propuestas por organismos internacionales y autoridades nacionales para mitigar los efectos de las ENT en la PEA.

2.1 Dimensiones determinantes de las ENT

Para abordar el concepto de salud es primordial entender todos los aspectos que influyen en la calidad de vida y su expectativa. Exponer la relación entre salud y enfermedad requiere, además, precisar la importancia de factores biológicos, conductuales, sociales y ambientales a los que se encuentre expuesto el individuo y enfatizar la importancia de la interacción de estos tres elementos que contribuyen a la creación, mantenimiento y problemas exhibidos en la práctica médica (Erazo *et al.*, 2002). De acuerdo con Lalonde (1974), los determinantes de la salud son el conjunto de factores personales, sociales, económicos y ambientales, los cuales determinan el entorno en el que se desarrollan los habitantes y esto repercute en la salud del individuo y de toda la población. Se pueden categorizar los elementos y factores de la salud en:

- 1) Biología humana
- 2) Estilo de vida y conducta de las personas
- 3) Entorno Social
- 4) Relación con el medio ambiente
- 5) Calidad y cantidad de los servicios de asistencia sanitaria

Dentro de los factores biológicos se encuentran los relacionados con la salud física y mental, respecto a la biología humana, edad de los individuos, herencia genética, enfermedades crónico-degenerativas, trastornos genéticos, malformaciones, entre otras que radican en la biología humana.

Pfeffer *et al.* (2015) refieren que las características genéticas están incorporadas dentro de los factores de riesgo no modificables en un individuo; sin embargo, pueden ser prevenidas y susceptibles de predicciones para reducir el impacto de la carga genética que impacta en las ENT. Por su parte, Herbert y Cohen (1993) muestran que los tipos de reacciones bioquímicas que generan los individuos han demostrado estar intrínsecamente relacionados con la regulación del sistema inmunológico; por lo tanto, la exposición a factores estresantes incontrolables tiende a afectar la función del sistema inmunológico de manera que puede aumentar la susceptibilidad a las enfermedades.

Los factores relacionados con el estilo de vida se refieren a la toma de decisiones del individuo respecto a su salud, las actividades físicas que realiza, sus creencias, costumbres, actividad física, alimentación, sueño y descanso, consumo de alcohol y tabaco, estrés, vida sexual sana y conducción segura, entre otras cosas, que repercuten en la salud (Lalonde, 1974; Marques, 2012). Dependiendo de las decisiones del individuo respecto a su estilo de vida, pudiera impactar positiva o negativamente, poniendo incluso en riesgo la salud. Si bien las características genéticas son elementos no modificables en un individuo, el estilo de vida sí depende de las decisiones que se tomen, incluyendo los hábitos como el sedentarismo y la alimentación inadecuada, que son, al final, factores que determinan muchas de las enfermedades crónicas no transmisibles (Ezzati & Riboli, 2013).

Los factores sociales se refieren a la cohesión social, etnia, nivel económico, mercado laboral, redes de apoyo y participación social (Puyol, 2012). La sociedad tiene una cohesión intrínseca que promueve o deteriora la salud de los individuos y que está mucho relacionada con el estrato social y la posible o nula capacidad de movilidad social. El ser humano es un ser social, que necesita de los demás para constituirse en la sociedad y sentirse parte de esta; por lo tanto, es natural que factores sociales lleguen a intervenir en su vida e influenciar su salud (Rosado *et al.*, 2018; Álvarez-Castaño, 2009).

Las influencias sociales impactan en los factores del estilo de vida tales como fumar, tomar drogas, hacer poco ejercicio físico, seguir una dieta poco saludable, entre otros. De la misma forma, estas variables a nivel social pueden afectar a las empresas, pues el nivel elevado de estrés laboral puede incrementar la práctica de hábitos insanos como fumar, abusar del alcohol o reducir la práctica de hábitos saludables como la frecuencia de ejercicio físico. Lo cual sugiere que el estrés

social no solo puede inducir efectos directos sobre la salud, sino también efectos indirectos a través de la modificación de los estilos de vida (Svedberg *et al.*, 2006).

La salud y el bienestar dependen en gran medida de las características ambientales urbanas y rurales; asimismo, los gradientes sociales se asocian de forma consistente con múltiples enfermedades crónicas (Diez Roux y Mair, 2010; Marmot, 2005). Los factores económicos modifican en gran medida las reacciones individuales y grupales. Por otro lado, modelos consolidados muestran que la organización del trabajo (exigencias y control) y la reciprocidad entre esfuerzo y recompensas (remuneración, seguridad laboral, estatus y estima) se asocian con bienestar/estrés y con riesgos cardio-metabólicos, afectando respuestas individuales y colectivas en el entorno de trabajo (Karasek Jr., 1979; Siegrist, 2010).

Dentro de los factores medioambientales se incluyen elementos físicos y químicos, así como la contaminación del aire, del agua y del suelo. Estos factores impactan la salud de los individuos, quienes carecen de control para minimizar sus efectos. En este sentido, pueden considerarse externalidades negativas que afectan tanto el estado físico como el mental (Marques *et al.*, 2018). En esta categoría puede ser considerable el impacto en las ENT, por ejemplo, las enfermedades respiratorias por la exposición constante a la atmósfera o enfermedades gastrointestinales por el consumo de agua contaminada.

Por último, los factores del servicio de atención incluyen el vasto sistema de atención a la salud, el acceso de la población, su calidad, el financiamiento de los servicios que atienden a los individuos, medicamentos disponibles y unidades médicas, entre otros servicios. En la accesibilidad a los servicios de atención médica interactúan condiciones como la distancia, los horarios de atención, las dotaciones y cualidades individuales para demandar servicios (Sánchez-Torres, 2017). Este rubro se refiere al acceso a los sistemas de salud con los que cuenta una sociedad, a través del ámbito laboral, y se enfoca en el acceso a los servicios de salud que las empresas tengan contratados para atender a los colaboradores, incluyendo el seguro social, de gastos médicos mayores o acuerdos con hospitales y clínicas.

Una buena salud es el mejor recurso para el progreso personal, económico y social. La promoción de la salud constituye un proceso político y social global que abarca acciones dirigidas a modificar las condiciones sociales, ambientales y económicas con el fin de favorecer su impacto positivo en la salud individual y colectiva (OMS, 1986).

Hipótesis 1: Las ENT tienen un efecto negativo en la productividad laboral en México.

2.2 Los hábitos de salud, prevención y promoción de la salud

De acuerdo con la Organización Mundial de la Salud (WHO, 2018) y del Centro de Control y Prevención de Enfermedades de los Estados Unidos (CDC, 2022), se pueden enumerar los hábitos de salud como:

1. Hacer ejercicio regularmente y evitar el sedentarismo.
2. Mantener una dieta equilibrada y nutritiva.
3. Dormir lo suficiente y tener un horario regular de sueño.
4. Evitar el consumo excesivo de alcohol.
5. No fumar ni usar productos de tabaco.
6. Mantenerse hidratado, bebiendo suficiente agua.
7. Practicar técnicas de manejo del estrés, como la meditación o el yoga.
8. Evitar el consumo excesivo de alimentos procesados y ricos en grasas y azúcares.

El concepto de hábitos de salud ha sido evaluado en diversas investigaciones relacionadas, se han promovido diversas intervenciones y llevado a la formulación de políticas (Budreviciute *et al.*, 2020).

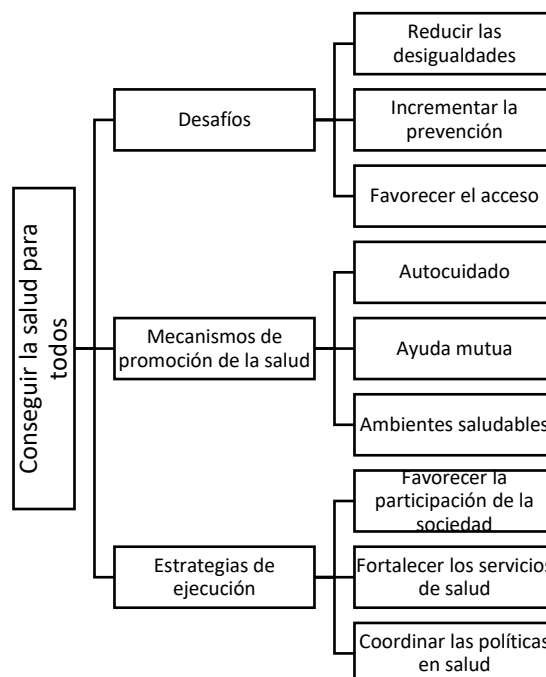
Un modelo de salud efectivo se centra tanto en la promoción de la salud como en la prevención de enfermedades, tomando en cuenta que los hábitos de vida ejercen un gran impacto en la calidad de la salud humana. La investigación guiada por diversas teorías psicosociales del comportamiento en salud ha contribuido a la comprensión de cómo los factores cognitivos y sociales contribuyen a la salud y la enfermedad humanas basados en las teorías como el modelo de las creencias en salud (Rosenstock, 1974), la teoría social cognitiva (Bandura, 1998), las teorías de la acción razonada (Ajzen & Fishbein, 1980), comportamiento planeado (Ajzen, 1991) o la teoría de motivación de protección (Rogers, 1983). De acuerdo con la Organización Mundial de la Salud (WHO, 2021d), la disminución de las ENT está enfocada en la reducción del abuso de sustancias adictivas como el alcohol o el tabaco, el incremento de la actividad física, la mejora de la alimentación y el mayor acceso a servicios de salud.

La promoción de la salud ha sido motivo de análisis y discusión a lo largo del tiempo y aunque la prevención primaria se enmarca principalmente en el entorno sanitario, el desarrollo de

la promoción de la salud debe realizarse dentro del ámbito comunitario (Gil-Girbau *et al.*, 2021), por ende, es importante considerar que puede ser de vital importancia dentro del ambiente laboral. Esta temática se ha utilizado como estrategia para realizar diversos planteamientos relacionados con la formulación de políticas públicas, intervenciones dirigidas a personas y comunidades, y para la propuesta y realización de investigaciones en salud (Budreviciute *et al.*, 2020) (Ver Figura 7).

Figura 7

Un marco para la promoción de la salud



Fuente: Epp, 1996

La promoción de la salud se consigue a través de tres mecanismos intrínsecos o internos de las personas, según lo planteado por Epp (1996) estos son:

1. El autocuidado se refiere a las decisiones y acciones que la persona lleva a cabo en beneficio de su propia salud.
2. La ayuda mutua o acciones que las personas realizan para ayudarse unas a otras de acuerdo con la situación que estén viviendo.

3. Los entornos sanos o la creación de las condiciones ambientales que favorecen la salud, tales como las áreas de trabajo que se constituyen en escenarios potenciadores de las capacidades de los trabajadores.

La buena salud es un proceso que involucra muchos factores. Aunque nacer con buenos genes y tener acceso a la atención médica son importantes, no son garantía ni seguridad para garantizar el bienestar. Los alimentos que comemos, los niveles de actividad física, la exposición al humo del tabaco, las interacciones sociales, el entorno en el que vivimos y muchos otros factores influyen en nuestra salud. Así como la combinación única de componentes biológicos, psicológicos, sociales, intelectuales y espirituales, así como el entorno cultural, económico y político en el que vivimos (Fertman y Allensworth, 2016). El enfoque que explora la interacción entre las personas y su entorno ha sido clave para el progreso de la promoción de la salud a nivel poblacional. La Carta de Ottawa y la evolución de la “nueva salud pública” reorientaron las políticas desde el riesgo individual hacia estrategias que abordan determinantes y ambientes (escuelas, lugares de trabajo, ciudades), consolidando un hito en el campo de la salud pública y la acción intersectorial (Kickbusch, 2003).

Hipótesis 2: Existe una asociación significativa entre las dimensiones económicas, hábitos de salud y laborales con las ENT.

2.3 Las ENT, hábitos de salud y la productividad laboral

La productividad constituye un concepto central en el análisis del crecimiento económico y del desempeño de las economías, particularmente en países en desarrollo. De acuerdo con la Comisión Económica para América Latina y el Caribe (CEPAL, 2016), la productividad laboral se define como la relación entre el nivel de producción generado y la cantidad de trabajo utilizada para producirla, comúnmente medida como el producto por trabajador o por hora trabajada. Esta medida refleja la capacidad de una economía para transformar insumos laborales en valor agregado y está estrechamente vinculada con el nivel de ingresos, la competitividad y el bienestar de la población.

Desde esta perspectiva, la productividad laboral no depende exclusivamente del esfuerzo individual del trabajador, sino que es el resultado de la interacción de múltiples factores estructurales y contextuales. Entre los determinantes más relevantes de la productividad laboral se

encuentran el capital humano (educación, habilidades y experiencia), las condiciones de salud de la fuerza laboral, el acceso a tecnología, la organización del trabajo, el entorno institucional, los incentivos económicos y las condiciones del mercado laboral. En particular, la CEPAL destaca que las brechas estructurales en productividad observadas en América Latina están estrechamente relacionadas con desigualdades en salud, educación y acceso a servicios, lo que limita el potencial productivo de amplios segmentos de la población económicamente activa (PEA).

En este contexto, las ENT y los hábitos de salud adquieren un papel relevante como determinantes de la productividad laboral. Las ENT afectan la capacidad funcional de los individuos, incrementan la probabilidad de ausentismo y presentismo, reducen la eficiencia en el desempeño de las tareas laborales y generan pérdidas económicas tanto a nivel individual como agregado. Por ello, la relación entre salud y productividad no puede entenderse únicamente desde una perspectiva clínica, sino también como un fenómeno económico y social con implicaciones directas sobre el desarrollo.

Diversas organizaciones internacionales han propuesto marcos estratégicos para la prevención y el control de las ENT, reconociendo su impacto transversal en la salud y la productividad. La Organización Panamericana de la Salud (2019) plantea cuatro pilares fundamentales:

1. Fortalecer entornos saludables
2. Participación y empoderamiento de la comunidad
3. Abordaje de los determinantes sociales de la salud
4. Fortalecer los sistemas y servicios de la salud

En el programa de prevención y control de las ENT por parte de la Organización Mundial de la Salud (WHO, 2013), se establecen cuatro líneas estratégicas para la prevención de estas enfermedades, en línea con 25 indicadores y los 9 objetivos contenidos en el marco global de vigilancia integral de la OMS, que son:

- a) Políticas y alianzas multisectoriales para la prevención y control de ENT: Fortalecer y promover acciones multisectoriales con todos los sectores relevantes del gobierno y la sociedad, incluida la integración en agendas económicas, académicas y de desarrollo.

- b) Factores de riesgo y factores protectores de las ENT: Reducir la prevalencia de los principales factores de riesgo de las ENT y fortalecer los factores protectores; emplear estrategias de promoción de la salud basadas en evidencia y herramientas de políticas, incluidas la regulación, la vigilancia y medidas voluntarias; y abordar los determinantes sociales, económicos y ambientales de la salud.
- c) Respuesta de los sistemas de salud a las ENT y sus factores de riesgo: Mejorar la cobertura, el acceso equitativo y la calidad de la atención para las cuatro principales ENT (enfermedades cardiovasculares, cáncer, diabetes y enfermedades respiratorias crónicas) y otras que sean priorizadas a nivel nacional, con énfasis en la atención primaria de salud que incluye la prevención y el mejoramiento del autocontrol.
- d) Vigilancia e investigación de las ENT: Fortalecer la capacidad de los países para la vigilancia e investigación de las ENT, sus factores de riesgo y determinantes, y utilizar los hallazgos de investigación como base para el desarrollo e implementación de políticas basadas en evidencia, programas académicos de desarrollo e implementación.

La evidencia empírica muestra que los hábitos de salud, como la actividad física, la alimentación, el consumo de alcohol y tabaco, y el manejo del estrés, están estrechamente vinculados con la aparición y progresión de las ENT, así como con el desempeño laboral. Estudios como el de Ding et al. (2016) documentan que la inactividad física se asocia con un mayor riesgo de ENT, lo que a su vez tiene consecuencias negativas sobre la productividad. De manera similar, Allen et al. (2021) proponen un marco analítico que integra la carga de ENT, los factores de riesgo asociados a estilos de vida y variables estructurales, mostrando cómo estas dimensiones pueden orientar políticas públicas con efectos potenciales sobre el bienestar y la productividad laboral.

Desde una perspectiva sociodemográfica, Zhao *et al.* (2020) analizaron cómo los factores socioeconómicos influyen en la morbilidad de adultos mayores en China, destacando la relación entre condiciones de salud crónicas, hábitos de vida y su efecto en la capacidad funcional y laboral de esta población. Asimismo, Wang y Wang (2020) aplicaron modelos basados en árboles de decisión para predecir la aparición de ENT en función de variables relacionadas con hábitos de salud, lo que tiene implicaciones relevantes para la prevención y la productividad futura.

Por su parte, Hosseinpoor *et al.* (2012) diseñaron un índice para comprender las desigualdades socioeconómicas en la distribución de las ENT en países de ingresos bajos y medios,

subrayando la importancia de abordar los determinantes sociales de la salud que afectan tanto a la aparición de enfermedades como a la productividad de la población económicamente activa.

Diversos estudios han destacado la necesidad de mejorar la caracterización de las ENT en países en desarrollo, particularmente en poblaciones trabajadoras, señalando que una mejor gestión de los indicadores de salud puede traducirse en una mayor productividad laboral y un impacto positivo en el crecimiento económico (Shikdar & Sawaqed, 2003; Kristen, 2010; Saha, 2013; Siddique *et al.*, 2020)

A fin de proteger la integridad del trabajador, se han incrementado los temas relacionados con la salud ocupacional; hoy en día, hablar de productividad implica reconocer la importancia del trabajador y su salud como tema fundamental del desarrollo organizacional y económico. De acuerdo con Tang (2015), hablar de productividad implica una serie de herramientas de medición existentes que difieren en muchos aspectos, incluida la cobertura de contenido, el alcance y los enfoques para cuantificar las diferentes fuentes de pérdida de productividad como el ausentismo, el presentismo y las ausencias laborales. Teniendo en cuenta la compatibilidad con diferentes métodos de medición y datos psicométricos disponibles, los costos de la productividad cada vez más pueden ser sustanciales y es recomendable su inclusión en las evaluaciones económicas de salud (Drummond *et al.*, 2015).

Como parte de los elementos de la productividad dentro de la salud de los trabajadores aparecen dos conceptos importantes por clarificar: el ausentismo y el presentismo. El ausentismo se refiere al tiempo de trabajo perdido debido a cambios en la situación laboral, como una reducción del tiempo de trabajo de rutina, cese temporal del trabajo, pérdida del trabajo o jubilación anticipada (Zhang *et al.*, 2011). El presentismo, por otro lado, se refiere a la reducción de la productividad en el trabajo debido a problemas de salud (Mayne *et al.*, 2004b). Los riesgos para la salud asociados con el ausentismo y el presentismo incluyen la falta de actividad física, mala nutrición, índice de masa corporal alto, estrés y diabetes (Mansyur, 2021).

Una revisión sistemática de literatura realizada por Tompa (2006) sobre evaluaciones económicas de intervenciones en salud y seguridad ocupacional en el lugar de trabajo destaca que varios estudios atribuyen los aumentos en productividad exclusivamente a las intervenciones implementadas, sin considerar que, en muchos casos, dichos aumentos coincidieron con la introducción simultánea de nuevos esquemas de incentivos salariales dentro de las organizaciones.

Esto plantea dudas sobre la validez de las atribuciones causales y subraya la necesidad de un análisis más riguroso del contexto organizacional.

Koopmanschap *et al.* (1995) estimaron el valor de la producción perdida teniendo en cuenta la existencia de desempleo y la posibilidad de que alguien pueda reemplazar a un empleado enfermo. La duración del período de fricción, que representa el tiempo requerido para que un nuevo empleado esté completamente incorporado al trabajo, varía según la enfermedad y el tipo de ocupación (Zhang *et al.*, 2011). Como regla general, a mayor nivel de calificación del puesto, mayor será el período de fricción. Este método de costos de fricción suele proporcionar estimaciones significativamente más bajas del costo de la mala salud en el entorno laboral.

Por contraste, los costos calculados mediante el método del capital humano fueron casi diez veces mayores. Este enfoque estima el valor económico de la producción perdida suponiendo que el tiempo no trabajado debido a enfermedad o muerte equivale directamente a una pérdida económica para la sociedad o la empresa, considerando los ingresos futuros que el individuo habría generado durante su vida laboral restante. Aunque este método tiende a sobrestimar los costos al no considerar sustituciones laborales, es ampliamente utilizado por su simplicidad y capacidad de capturar el impacto total de la pérdida de productividad. (Koopmanschap *et al.*, 1995). Rozjabek *et al.* (2020) por su parte, realizaron un estudio empírico evaluando componentes físicos, la activación, vitalidad y salud general en el ausentismo, presentismo, comportamiento de la realización de actividades y la pérdida de productividad derivado de la encuesta nacional de salud y bienestar en el 2018 en Estados Unidos con 69,742 adultos evaluados, encontrando que el índice de masa corporal mostró un efecto negativo en el compromiso con la empresa y problemas de actividad.

Los estudios que evalúan la efectividad de los programas para proporcionar intervenciones de prevención primaria para la prevención y control de las ENT en contextos de países de ingresos medios son insuficientes y están poco analizados. Además, los resultados disponibles provienen principalmente de estudios observacionales, y se necesita sintetizar evidencia de ensayos controlados (Jeet *et al.*, 2017).

Diversos estudios internacionales han buscado cuantificar el impacto que tienen las ENT y los hábitos de salud en la productividad laboral, utilizando metodologías tanto cualitativas como cuantitativas. En un estudio aleatorizado con más de 30,000 empleados en Estados Unidos,

Brassey *et al.* (2023) encontraron que un enfoque integral de salud en el lugar de trabajo se asocia con mayor innovación y reducción del *burnout*, lo que demuestra una relación directa entre bienestar integral y productividad. De igual forma, Song y Baicker (2019), en un ensayo clínico controlado, analizaron un programa de bienestar en un entorno corporativo y, aunque se observaron mejoras en ciertos comportamientos de salud, no se identificaron efectos estadísticamente significativos sobre productividad laboral en el corto plazo, destacando la necesidad de intervenciones sostenidas y de largo plazo.

Hipótesis 3: La adopción de hábitos de salud positivos de la PEA tiene un efecto positivo en la productividad laboral en México

Respecto a las ENT específicamente, estudios como el de Eng *et al.* (2016) muestran que programas de salud implementados en instituciones públicas pueden reducir la presión arterial entre trabajadores hipertensos, lo que a su vez puede contribuir a una menor probabilidad de ausentismo y presentismo, factores reconocidos como determinantes de la pérdida de productividad (Zhang *et al.*, 2011; Mayne *et al.*, 2004a).

En el caso específico de México, la evidencia empírica sobre el vínculo directo entre ENT, hábitos saludables y productividad laboral aún es limitada. Existen estudios como el de Méndez-Hernández *et al.* (2012), que evalúan la implementación de programas de actividad física en universidades públicas y demuestran reducciones en el gasto médico derivado de diabetes e hipertensión; sin embargo, pocos trabajos se han enfocado en cuantificar directamente la pérdida de productividad laboral. Asimismo, aunque estudios como el de Ramírez *et al.* (2023) muestran beneficios en salud mental y actividad física con programas de autocuidado en línea, su enfoque no se centra en productividad laboral como variable de resultado.

Por su parte, la Secretaría de Salud del gobierno de México definió tres pilares como parte de la estrategia nacional para la prevención y el control del sobrepeso, la obesidad y la diabetes (SS, 2013a), siendo estos: salud pública, atención médica y regulación sanitaria, junto con política fiscal. Su objetivo es la promoción de la adopción de un estilo de vida saludable, detección oportuna y masiva para su adecuado control, generar una plataforma que permita contar con personal capacitado, así como insumos suficientes para el acceso a los servicios de salud e incrementar el control metabólico. Basado en esta información se pueden tomar elementos

relevantes para un diseño, desarrollo e implantación de un sistema de información que monitoree el comportamiento de las ENT, así como el sobrepeso y la obesidad.

No obstante, a pesar de la amplia evidencia que documenta la relevancia del presentismo como uno de los principales mecanismos a través de los cuales las ENT afectan la productividad laboral, su medición empírica directa presenta limitaciones sustantivas, particularmente en estudios basados en encuestas poblacionales y bases de datos secundarias. Instrumentos específicos para capturar pérdidas de rendimiento asociadas al deterioro funcional, la fatiga o la disminución de la capacidad laboral no suelen estar disponibles de manera sistemática, lo que dificulta su incorporación explícita en modelos econométricos tradicionales, especialmente en contextos de países en desarrollo.

Ante esta dificultad, la presente investigación adopta una estrategia metodológica alternativa orientada a aproximar el efecto del presentismo de manera indirecta. En particular, el Capítulo IV propone un método de estimación basado en un ajuste por cronicidad, que permite capturar el impacto diferencial de las ENT sobre la productividad laboral considerando la coexistencia de enfermedades crónicas y la persistencia de las condiciones de salud en el tiempo. Este enfoque reconoce que las ENT no afectan la productividad únicamente a través de episodios puntuales de ausentismo, sino mediante una reducción sostenida del desempeño laboral, consistente con la literatura sobre presentismo. De esta forma, el ajuste por cronicidad se incorpora como un mecanismo analítico que vincula la evidencia conceptual revisada en este capítulo con la estrategia de estimación causal desarrollada posteriormente, fortaleciendo la coherencia entre el marco teórico, la revisión de literatura y el diseño metodológico del estudio.

2.4 Marco Conceptual

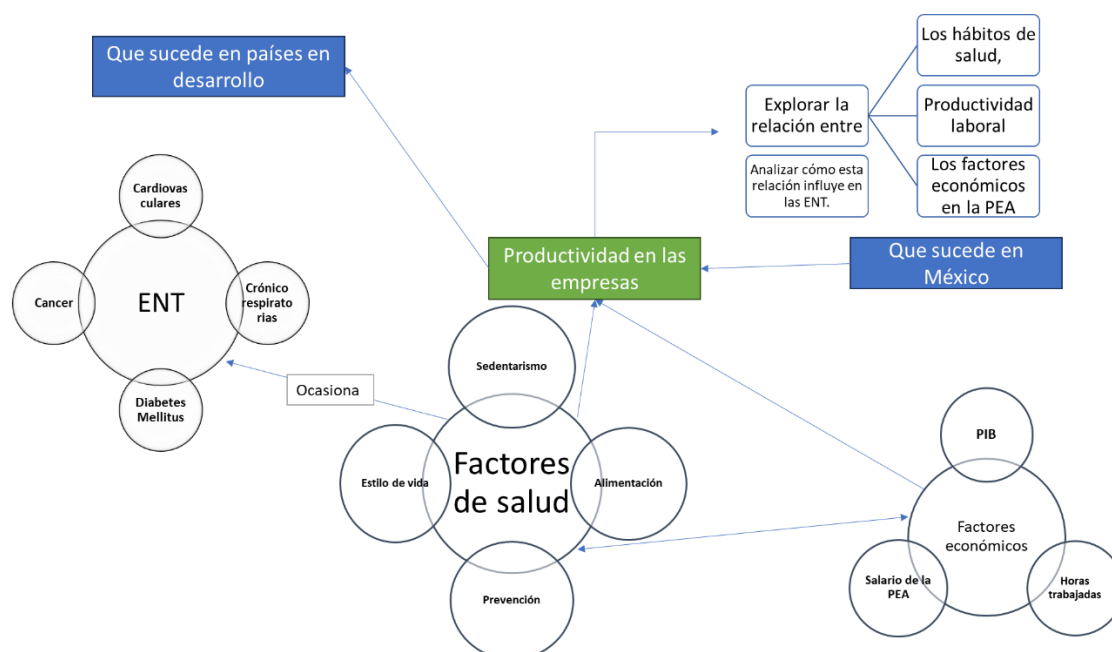
El marco conceptual de esta investigación se construye a partir de la integración de diversas dimensiones que vinculan la salud, la economía y la productividad laboral. El objetivo de este esquema es representar gráficamente la interacción entre los determinantes de la salud con especial énfasis en las ENT y su efecto sobre el desempeño económico de la PEA, así como los factores contextuales que inciden en esta relación tanto a nivel nacional como internacional. Dicho esquema puede ser observado en la Figura 8.

En el centro del diagrama se encuentra la productividad laboral en las empresas, entendida como el resultado económico del desempeño laboral individual y colectivo, influenciada por

factores de salud y económicos. Esta productividad se ve afectada directamente por los factores de salud, entre los que se incluyen el sedentarismo, la alimentación, los estilos de vida y la prevención. Estos factores, a su vez, inciden en la aparición y desarrollo de las ENT, tales como las enfermedades cardiovasculares, el cáncer, las enfermedades respiratorias crónicas y la diabetes mellitus.

Figura 8

Esquema conceptual para el análisis del impacto de las ENT en la productividad laboral



Fuente. Elaboración propia

La sección izquierda del diagrama aborda el contexto de los países en desarrollo, con base en información proveniente de organismos internacionales como la OCDE, la OMS y la FAO. Este bloque destaca el enfoque analítico utilizado para posicionar a los países según índices de impacto de las ENT, así como los procedimientos de análisis, que incluyen la selección de variables, la creación de variables latentes y la generación de índices. También se sugieren modelos estadísticos adecuados para esta evaluación.

La parte derecha del diagrama se enfoca en el caso de México, utilizando como fuente principal la Encuesta Nacional de Salud y Nutrición (ENSANUT). Aquí se plantea un análisis de la causalidad entre las ENT, los hábitos de salud y la productividad laboral, considerando como

variables relevantes el salario de la PEA, las horas trabajadas y su contribución al Producto Interno Bruto. La propuesta metodológica incorpora técnicas de análisis multivariado y aprendizaje automático con enfoque causal, incluyendo ecuaciones estructurales y modelos de inferencia causal.

La estructura conceptual sugiere que el análisis debe considerar simultáneamente dos dimensiones: por un lado, la exploración de la relación entre los hábitos de salud, la productividad y los factores económicos; y por otro, cómo esta interacción influye en la prevalencia y carga de las ENT. Esta doble perspectiva permite avanzar hacia una comprensión más robusta del problema y sienta las bases para el diseño de estrategias de intervención más eficaces en el ámbito empresarial y de salud pública.

CAPÍTULO III. Diseño de la investigación

La presente investigación es de corte cuantitativo y de tipo causal-explicativo. Este enfoque, de acuerdo con Schenker y Rumrill (2004), se orienta a la identificación y explicación de los factores subyacentes al fenómeno de estudio mediante el análisis sistemático de registros observacionales con el fin de formular inferencias de carácter causal. En particular, la investigación pone énfasis en el desarrollo de una metodología para analizar las relaciones entre indicadores alimentarios, económicos y de estilos de vida, en comparación con el comportamiento de las enfermedades no transmisibles, tanto a nivel internacional como nacional, incorporando técnicas econométricas y de aprendizaje automático con enfoque causal, como el *Double Machine Learning*, para estimar el efecto de las ENT sobre la productividad laboral.

La estrategia metodológica se estructura en tres niveles de análisis complementarios, diferenciados tanto por su alcance analítico como por su dimensión temporal. En el primer nivel se desarrolla un análisis exploratorio y comparativo de las principales dimensiones asociadas a las ENT en una muestra de países en desarrollo, empleando información correspondiente al periodo previo a la pandemia por COVID-19. Este análisis internacional tiene un carácter estructural y busca identificar patrones comunes en la relación entre factores económicos, alimentarios, ambientales y de salud pública, mediante técnicas de reducción de dimensionalidad, análisis multivariado y posicionamiento relativo. La utilización de datos pre-pandemia responde a la necesidad de evitar distorsiones derivadas de choques exógenos excepcionales, permitiendo una comparación más estable entre países y la identificación de determinantes estructurales de las ENT.

En un segundo nivel se aplica un modelo de ecuaciones estructurales para evaluar las relaciones de causalidad entre variables latentes, tales como hábitos de salud, factores socioeconómicos y condiciones laborales, y la prevalencia de ENT. Este análisis se apoya en información proveniente de encuestas representativas y permite cuantificar la magnitud y dirección de los efectos de cada constructo, contribuyendo a una comprensión integrada de los mecanismos subyacentes que vinculan comportamientos de riesgo y condiciones estructurales con la aparición de ENT.

En el tercer nivel se estima el efecto de las ENT sobre la productividad laboral en México mediante un corte transversal correspondiente al año 2022, utilizando métodos de aprendizaje automático con enfoque causal. Esta etapa se orienta a medir el impacto diferencial de las ENT sobre variables económicas clave, como el ingreso laboral y las horas trabajadas de la población económicamente activa, capturando el contexto empírico específico posterior a la pandemia. La articulación de este análisis nacional con el marco internacional pre-pandemia permite combinar una perspectiva estructural de largo plazo con una estimación puntual y contextualizada del impacto de las ENT en la productividad laboral en México.

En conjunto, esta estrategia metodológica busca ofrecer una visión integral, coherente y basada en evidencia, que sustente el diseño de propuestas de intervención orientadas a la prevención de las ENT y a la mejora de la productividad en el ámbito de la salud ocupacional.

3.1 Variables y Diseño Metodológico

El objeto es analizar patrones, relaciones y tendencias de las variables seleccionadas y posteriormente modelar algoritmos matemáticos que puedan mostrar la relación entre las variables presentadas y diversas variables latentes. Debido a que tanto los efectos como las supuestas causas ya han ocurrido, se estudian *ex post facto* (Azalea, 2022). Para que esta investigación tenga razonamiento estadístico, se considerará una muestra grande, disponibilidad de la información en cada una de las variables, buscar la precisión y validez en cada momento y considerar cuidadosamente los estándares de investigación cuantitativa.

3.1.1 Operacionalización de Variables

La operacionalización de variables consiste en definir y estructurar los conceptos teóricos en indicadores observables y medibles que permitan ser analizados empíricamente. En esta investigación se integró un enfoque multinivel y multidimensional para analizar el impacto de las ENT sobre la productividad laboral, utilizando tres bases de datos complementarias.

Se emplearon diversas fuentes de organismos internacionales (OCDE, Banco Mundial, FAO, WHO e IHME) para construir un análisis comparado entre países en desarrollo a través de indicadores económicos, nutricionales, ambientales y sanitarios relevantes. Posteriormente, se utilizó la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2022 para caracterizar los hábitos de salud, el estado clínico y la prevalencia de ENT en la PEA de México, en el contexto posterior

a la pandemia por COVID-19. Se incorporaron datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022 para estimar el efecto económico de dichas enfermedades, mediante el análisis de ingresos, condiciones laborales, gasto en salud y consumo alimentario a nivel individual. La operacionalización de las variables se organizó según categorías teóricas y empíricas, permitiendo establecer relaciones observables entre determinantes estructurales, condiciones clínicas y resultados económicos.

Factores relacionados con las ENT en países en desarrollo durante la edad laboral

Para el análisis comparativo de los factores relacionados con las ENT en países en desarrollo se utilizaron diversas fuentes internacionales reconocidas por su solidez metodológica, cobertura global, consistencia en la recolección y actualización de datos. Las bases de datos seleccionadas permiten integrar variables de carácter sanitario, económico, demográfico y alimentario con una estructura comparable entre países.

La información contenida en la recopilación de datos por parte de la Organización para la Cooperación y el Desarrollo Económicos (OCDE, 2023) proporcionó indicadores clave sobre inversión en salud, gasto público y privado, presión arterial sistólica promedio, carga de enfermedad medida en años de vida ajustados por discapacidad y estimaciones demográficas como población total y PEA. Estos datos permiten contextualizar el entorno estructural de los sistemas de salud y su impacto en la incidencia y el control de las ENT.

La base del Banco Mundial (WBOD, 2023) fue utilizada para obtener variables macroeconómicas, como el Producto Interno Bruto, niveles de inversión pública, emisiones de metano como proxy ambiental y el índice de Gini para desigualdad económica. Estos indicadores permiten evaluar los determinantes sociales y económicos asociados a la salud en contextos nacionales diferenciados.

El *Institute for Health Metrics and Evaluation* (IHME, 2020) aportó información detallada sobre la carga específica de las ENT, incluyendo datos de incidencia, prevalencia, mortalidad y causas específicas de muerte atribuibles a factores de riesgo como consumo de alcohol, tabaco, carne procesada, azúcares añadidos, sodio, grasas trans, actividad física insuficiente y exposición a contaminantes. Esta fuente se distingue por ofrecer estimaciones consistentes y comparables a nivel global, basadas en el estudio de la Carga Global de Enfermedad.

Complementariamente, se incorporaron datos de la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, 2023), los cuales permitieron incluir variables de consumo alimentario per cápita, expresadas en kilocalorías o litros por persona/día. Entre ellas destacan el consumo de vegetales, carnes, frutas, legumbres, azúcares, bebidas alcohólicas, aceites, cereales, y productos de origen animal, factores estrechamente vinculados con los hábitos de salud de las poblaciones.

El indicador de productividad por hora trabajada se obtuvo a partir de los datos estandarizados publicados en *The American Economic Review* por Feenstra *et al.* (2022), los cuales ofrecen comparabilidad internacional y una medición robusta del rendimiento económico ajustado por tiempo laboral.

Se recopiló información de las bases de datos obtenidas centradas en indicadores de salud de ENT, información económica y alimentaria como se muestra en la Tabla 4, correspondientes al periodo de 1990 a 2019, para una muestra de 14 países en desarrollo: Polonia, Turquía, Ucrania, Egipto, Marruecos, Sudáfrica, México, Brasil, Chile, Colombia, Argentina, India, Malasia y Tailandia. Esta selección de países obedece a su clasificación como países en desarrollo según los criterios establecidos por la Organización para la Cooperación y el Desarrollo Económicos (OCDE, 2023), donde se muestran los primeros 14 países. Esta categorización toma en cuenta indicadores como el ingreso nacional bruto per cápita, la estructura económica, el nivel de industrialización y la transición demográfica, permitiendo delimitar un grupo de países con características comparables en términos de desafíos estructurales para el desarrollo, condiciones sanitarias y capacidades institucionales.

Tabla 4

Bases de datos globales

Variables	Unidad de medida	Fuente
Incidencia	Número de personas	IHME, 2020
Prevalencia	Número de personas	IHME, 2020
Mortalidad	Número de personas	IHME, 2020
DALYs	Años por cada 100,000 personas	OCDE, 2023
Presión sanguínea sistólica promedio	mmHg	OCDE, 2023

Metano en el medio ambiente	kT _{CO2} equivalente	WBOD, 2023
Probabilidad de muerte entre 30 y 70 años por ENT	Probabilidad en %	WHO, 2023
Inversión en salud	Millones de dólares	OCDE, 2023
Inversión gubernamental en salud	Millones de dólares	OCDE, 2023
Producto Interno Bruto	Millones de dólares	WBOD, 2023
Indicador de GiNi	Millones de dólares	WBOD, 2023
Gasto privado en salud	Millones de dólares	OCDE, 2023
Gasto de bolsillo	Millones de dólares	OCDE, 2023
Población total	Millones de personas	OCDE, 2023
Población laboral	Millones de personas	OCDE, 2023
Productividad por horas trabajadas	Tarifa por hora	Feenstra <i>et al.</i> , 2022
IMC promedio	kg/m2	WHO, 2023
Prevalencia de sobrepeso superior al índice IMC 25	% De población	WHO, 2023
Consumo total de alimentos	kcal/persona/día	FAO, 2023
Consumo total de vegetales	kcal/persona/día	FAO, 2023
Consumo total de carne	kcal/persona/día	FAO, 2023
Cereales	kcal/persona/día	FAO, 2023
Almidón	kcal/persona/día	FAO, 2023
Azúcar	Litros per cápita	FAO, 2023
Legumbres	kcal/persona/día	FAO, 2023
Nueces	Kilogramos per cápita	FAO, 2023
Cultivos oleaginosos	kcal/persona/día	FAO, 2023
Aceite	kcal/persona/día	FAO, 2023
Vegetales	kcal/persona/día	FAO, 2023
Frutas	kcal/persona/día	FAO, 2023
Bebidas alcohólicas	Litros/persona/día	FAO, 2023
Carne	kcal/persona/día	FAO, 2023
Viseras	kcal/persona/día	FAO, 2023

Leche	kcal/persona/día	FAO, 2023
Huevo	kcal/persona/día	FAO, 2023
Pescado	kcal/persona/día	FAO, 2023
Mortalidad por 2 o más causas de ENT	% de casos	WHO, 2023
Muertes atribuidas a la contaminación	Número de personas	IHME, 2020
Muertes atribuidas al consumo de alcohol	Número de personas	IHME, 2020
Muertes causadas por el consumo de carne procesada	Número de personas	IHME, 2020
Muertes atribuidas al consumo de carne	Número de personas	IHME, 2020
Muertes atribuidas a la ingesta excesiva de sodio	Número de personas	IHME, 2020
Muertes atribuidas al consumo de bebidas azucaradas	Número de personas	IHME, 2020
Muertes atribuidas al consumo de grasas trans	Número de personas	IHME, 2020
Muertes atribuidas a la falta de control por LDL	Número de personas	IHME, 2020
Muertes atribuidas a la exposición a carcinógenos	Número de personas	IHME, 2020
Muertes atribuidas a la exposición a elementos causantes de asma	Número de personas	IHME, 2020
Muertes atribuidas al consumo de tabaco	Número de personas	IHME, 2020
Muertes atribuidas a la exposición a partículas nocivas en gases	Número de personas	IHME, 2020
Muertes atribuidas a la baja actividad física	Número de personas	IHME, 2020

Fuente: Elaboración propia

La información fue filtrada para incluir únicamente datos asociados a personas entre 20 y 64 años, lo que corresponde con la definición estándar de población en edad laboral utilizada por organismos como la OCDE y la OIT. Esta restricción busca garantizar que los indicadores analizados reflejen condiciones de salud y productividad directamente vinculadas con la fuerza laboral, evitando sesgos derivados de edades extremas o de dependencia.

La elección del periodo comprendido entre 1990 y 2019 coincide con el inicio y consolidación de esfuerzos sistemáticos internacionales para la medición de las ENT y sus factores de riesgo, como es el caso del *Global Burden of Disease Study* coordinado por el Instituto para la evaluación y métricas de salud (*Institute for Health Metrics and Evaluation*), cuyos datos han sido

recopilados, estandarizados y actualizados de forma continua desde 1990 (IHME, 2020). A partir de ese año, también se observó una mayor disponibilidad de indicadores internacionales comparables en bases como la OCDE, el Banco Mundial y la FAO, particularmente en lo referente a inversión en salud, variables alimentarias y condiciones ambientales (Ritchie *et al.*, 2023). El rango de 30 años resulta metodológicamente adecuado ya que permite capturar tendencias estructurales de largo plazo, observar las transiciones epidemiológicas y nutricionales en países en desarrollo e identificar los efectos acumulativos de políticas públicas sobre la salud de la PEA. A diferencia de periodos más cortos, esta ventana temporal permite evaluar la persistencia y evolución de los factores de riesgo y su relación con la productividad laboral sin incurrir en sesgos por fluctuaciones coyunturales o eventos extremos. Asimismo, se evita el impacto exógeno provocado por la pandemia de COVID-19, ya que el año 2019 representa el último año con datos no distorsionados por esta contingencia sanitaria global. Adicionalmente, estudios epidemiológicos de referencia como el *Framingham Heart Study* han estructurado sus evaluaciones en ciclos de 20 a 30 años para permitir el seguimiento efectivo de indicadores de salud poblacional, validando empíricamente este horizonte temporal como una unidad de análisis robusta en salud pública (Mahmood *et al.*, 2014).

Las variables empleadas en esta investigación han sido seleccionadas con base en su relevancia empírica y teórica para explicar la relación entre las ENT, los determinantes estructurales de salud y el desempeño económico en países en desarrollo. Estas variables se han agrupado en cinco grandes categorías: salud poblacional, condiciones económicas, indicadores laborales, factores alimentarios y determinantes ambientales.

Variables de salud poblacional

Estas variables permiten medir directamente la carga de las ENT en las poblaciones de estudio. Incluyen indicadores como la incidencia, prevalencia y mortalidad por ENT, así como métricas agregadas como los Años de Vida Ajustados por Discapacidad (DALYs). Asimismo, se incorporan variables de riesgo como el índice de masa corporal promedio (IMC), la presión arterial sistólica promedio y la probabilidad de muerte por ENT, las cuales ofrecen una visión integral del estado de salud y del riesgo epidemiológico de las poblaciones laborales. Las cuales son esenciales para cuantificar el impacto directo e indirecto de las ENT en el bienestar y la capacidad laboral.

Variables económicas y laborales

Esta categoría agrupa los indicadores que permiten contextualizar el entorno económico general y medir el desempeño de la fuerza laboral. Se incluyen variables macroeconómicas como el Producto Interno Bruto (PIB), el índice de Gini (como medida de desigualdad), la inversión gubernamental en salud, el gasto privado y de bolsillo en salud, y la población total, que permiten identificar diferencias estructurales en la capacidad institucional para atender los desafíos sanitarios. En cuanto al desempeño laboral, se consideran variables como la productividad por hora trabajada y la PEA (20–64 años), que sirven para establecer el vínculo cuantificable entre el estado de salud y el rendimiento económico.

Variables alimentarias y de estilo de vida

Dado que la dieta es uno de los principales determinantes de las ENT, se incorporaron variables detalladas sobre el consumo de alimentos expresados en kilocalorías/persona/día, incluyendo cereales, vegetales, frutas, carne, leche, huevo, pescado, azúcar, legumbres, nueces, aceites, bebidas alcohólicas y alimentos procesados. Estas variables obtenidas de la FAO permiten modelar patrones nutricionales y establecer asociaciones con la carga de enfermedad, ya que la ingesta calórica y la calidad de los alimentos están directamente relacionadas con el riesgo de obesidad, hipertensión, diabetes tipo 2 y enfermedades cardiovasculares.

Variables ambientales y de riesgo asociado

Como factores contextuales que afectan la salud, se consideraron variables como la presencia de metano en el medio ambiente (ktCO₂ equivalente) y un conjunto de variables de muertes atribuibles a factores de riesgo específicos tales como el consumo de alcohol, tabaco, grasas trans, sodio, bebidas azucaradas y la exposición a contaminantes y carcinógenos atmosféricos. Estas variables permiten vincular las condiciones ambientales y conductuales con los desenlaces en salud, ofreciendo una perspectiva integral sobre la multicausalidad de las ENT.

Fueron utilizados los valores consolidados de la Tabla 4 y en específico para el año 2019. Fue utilizada una variable latente llamada factores H que se refiere al efecto de la variable observada que impacta en las ENT derivadas de una alta incidencia de la variable, por ejemplo, el alto consumo de tabaco. Por otro lado, se tiene la variable latente de factores L que indica el efecto de una baja utilización o la falta de esa variable en los individuos, por ejemplo, el bajo consumo

de leguminosas. En el caso de las enfermedades crónico-respiratorias solo se encontraron variables de corte alto, etiquetadas como factores H, que afectan la enfermedad. En la Tabla 5 se identifican las variables seleccionadas para cada factor H y L de las diferentes ENT.

Tabla 5

Variables seleccionadas en Factores H y L de las ENT

Enfermedad	Tipo de factor	Variables
Cardiovasculares	Factores H	Contaminación ambiental, consumo de alcohol, carnes procesadas, carne roja, sodio, bebidas azucaradas, grasas trans, IMC alto, glucosa elevada, presión sistólica.
	Factores L	Baja actividad física, bajo consumo de tabaco, baja fibra, bajo consumo de frutas, legumbres, nueces y semillas, grasas poliinsaturadas, <i>omega</i> 3, vegetales, granos integrales, LDL bajo.
Diabetes <i>Mellitus</i>	Factores H	Contaminación, carnes procesadas, carne roja, bebidas azucaradas, IMC alto, glucosa elevada.
	Factores L	Bajo consumo de tabaco, fibra, frutas, nueces y semillas, granos integrales, baja actividad física.
Neoplasias	Factores H	Contaminación, alcohol, carnes procesadas, carne roja, sodio, IMC alto, glucosa elevada, exposición a carcinógenos laborales.
	Factores L	Bajo consumo de tabaco, calcio, fibra, frutas, lácteos, vegetales, granos integrales, baja actividad física.
Enfermedades respiratorias	Factores H	Contaminación, IMC alto, exposición ocupacional a asma, exposición ocupacional a carcinógenos, tabaquismo.
	Factores L	No se identifican factores L en esta categoría.

Fuente: Elaboración propia

Hábitos de salud en la PEA y sus efectos en las ENT en México

Para el análisis correspondiente a México, se utilizó la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2022, la cual constituye una de las principales fuentes estadísticas de información en salud en el país. ENSANUT es una encuesta representativa a nivel nacional y

estatal, que proporciona datos detallados sobre las condiciones de salud, los determinantes sociales, el estado nutricional y los estilos de vida de la población.

La edición 2022 es particularmente relevante ya que representa el primer levantamiento amplio de datos postpandemia con mediciones objetivas y estandarizadas tras el impacto del llamado “*health debt*” (rezago sanitario acumulado) provocado por la COVID-19. Esto la convierte en una base idónea para analizar el estado actual de los hábitos de salud y las ENT en la PEA en un contexto de recuperación y adaptación estructural en salud pública.

El análisis se enfocó exclusivamente en la población de 20 a 64 años correspondiente a la edad laboral estándar a fin de identificar patrones de comportamiento y condiciones de salud que pudieran estar afectando la productividad. La selección de variables se realizó con base en cuatro categorías analíticas: características sociodemográficas, indicadores antropométricos y bioquímicos, estilos de vida y diagnóstico de ENT. Estas dimensiones permiten evaluar de manera integral los determinantes individuales que influyen en la aparición y evolución de enfermedades crónicas, así como su asociación con factores económicos y ocupacionales. Se utilizaron las variables que contienen información de los hábitos de salud en la PEA en México conforme a lo mostrado en la Tabla 6.

Tabla 6

Variables seleccionadas de la ENSANUT

Categoría	Variable	Descripción	Relevancia
Generales	Entidad	Región geográfica	Importante para el análisis regional y la comprensión de las disparidades geográficas en salud
	Sexo	Hombre o mujer	Diferencias de género en los comportamientos y resultados de salud
	Edad	Edad en años	La edad es un predictor significativo del estado de salud y del riesgo de ENT
	Escolaridad	Nivel educativo	La educación influye en la alfabetización sanitaria y en el acceso a los recursos sanitarios
	Salario	Nivel de ingresos	El nivel socioeconómico es un determinante de los resultados de salud

	Peso	Peso corporal en kg	Esencial para calcular el IMC y comprender los riesgos para la salud relacionados con el peso
	Altura	Altura en cm	Se utiliza en el cálculo del IMC
	Circunferencia	Circunferencia de la cintura en cm	Indicador de obesidad central, que es un factor de riesgo para las ENT
Bioquímico	TRIG	Triglicéridos	Los niveles altos son un factor de riesgo para las enfermedades cardiovasculares
	COL_LDL	Colesterol LDL	Los niveles altos son un factor de riesgo para las enfermedades cardiovasculares
Estilo de vida	Sístole	Presión arterial sistólica	Los niveles altos son un factor de riesgo para la hipertensión y las enfermedades cardiovasculares
	kg_win	Aumento de peso	Refleja los cambios en el peso corporal, que pueden afectar la salud en general
	kg_lose	Pérdida de peso	Refleja los cambios en el peso corporal, que pueden afectar la salud en general
	Cansado	Frecuencia de sensación de cansancio	Indicador de bienestar general y posibles problemas de sueño
	Dormir bien	Calidad del sueño	La mala calidad del sueño está relacionada con varios problemas de salud
	Tabaquismo	Estado de tabaquismo	El tabaquismo es un factor de riesgo importante para numerosas enfermedades
	AF_v	Actividad física vigorosa	La actividad física protege contra muchas enfermedades crónicas
	AF_m	Actividad física moderada	La actividad física protege contra muchas enfermedades crónicas
	Hrs_c	Caminar como actividad física	La actividad física protege contra muchas enfermedades crónicas
	Hrs_s	Sedentario	La actividad física protege contra muchas enfermedades crónicas
Diagnosis	Diabetes	Diagnóstico de la diabetes	Principales ENT con importantes repercusiones en la salud

	Hipertensión	Diagnóstico de la hipertensión	Principales ENT con importantes repercusiones en la salud
	Enfermedad oncológica	Diagnóstico de la enfermedad oncológica	Principales ENT con importantes repercusiones en la salud

Fuente: ENSANUT, 2022

Se realizó una transformación de la información de peso y altura para homologar el resultado con los estándares internacionales y convirtiéndola a índice de masa corporal, le fue asignado el nombre de BMI (Véase ecuación 1).

$$BMI = \frac{Peso}{Altura^2} \quad (1)$$

Se utilizaron las variables DM (Diabetes *Mellitus*), HT (Hipertensión) y Evento_CAD (Enfermedad coronaria), que son indicadores de individuos que han sido diagnosticados con alguna ENT, y se procedió a generar una variable llamada ENT en donde indica si el individuo cuenta con alguna de las tres ENT.

Impacto de los hábitos de salud y las ENT en la productividad laboral en México

Para complementar el análisis de salud con una dimensión económica y social más amplia, se utilizó la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022 elaborada por el Instituto Nacional de Estadística y Geografía (INEGI). Esta encuesta es representativa a nivel nacional y estatal y proporciona información detallada sobre los ingresos, los gastos, el empleo, las condiciones de vida y el acceso a servicios de los hogares mexicanos. Las variables seleccionadas se muestran en la Tabla 7.

Tabla 7

Variables seleccionadas de la ENIGH

Población	Gasto	Ingreso	Trabajo	Hogar
• Sexo • Edad • Discapacidad	• Tipo de gasto (Salud) • Frecuencia • Monto	• Tipo de ingreso • Ingreso mensual	• Ingresos laborales • Horas trabajadas	• Ubicación geográfica • Estrato socioeconómico • Perceptores de ingreso • Ingreso trabajo

• Causa de discapacidad • Años trabajando • Horas de trabajo • Años con problemas de salud • Problemas de salud • Pago por atención en salud • Diabetes • Presión alta • Control de peso y talla • Entidad federativa • Factor de expansión			• Servicio médico	• Sueldos • Alimentos • Alimentos dentro del hogar • Cereales • Carnes • Pescado y mariscos • Leche y derivados • Huevo • Aceites y grasas • Tubérculos • Verduras • Frutas • Azúcares • Bebidas • Tabaco • Cuidados de la salud • Atención primaria • Atención hospitalaria • Medicinas • Percepciones totales
---	--	--	---	--

Fuente: ENIGH, 2022

El análisis se centró en la PEA aplicando un filtro en la base de datos para seleccionar únicamente a las personas que se encontraban laborando. La estructura del análisis se basó en variables agrupadas en cinco dimensiones: características individuales, condiciones de salud, ingreso, gasto y alimentación.

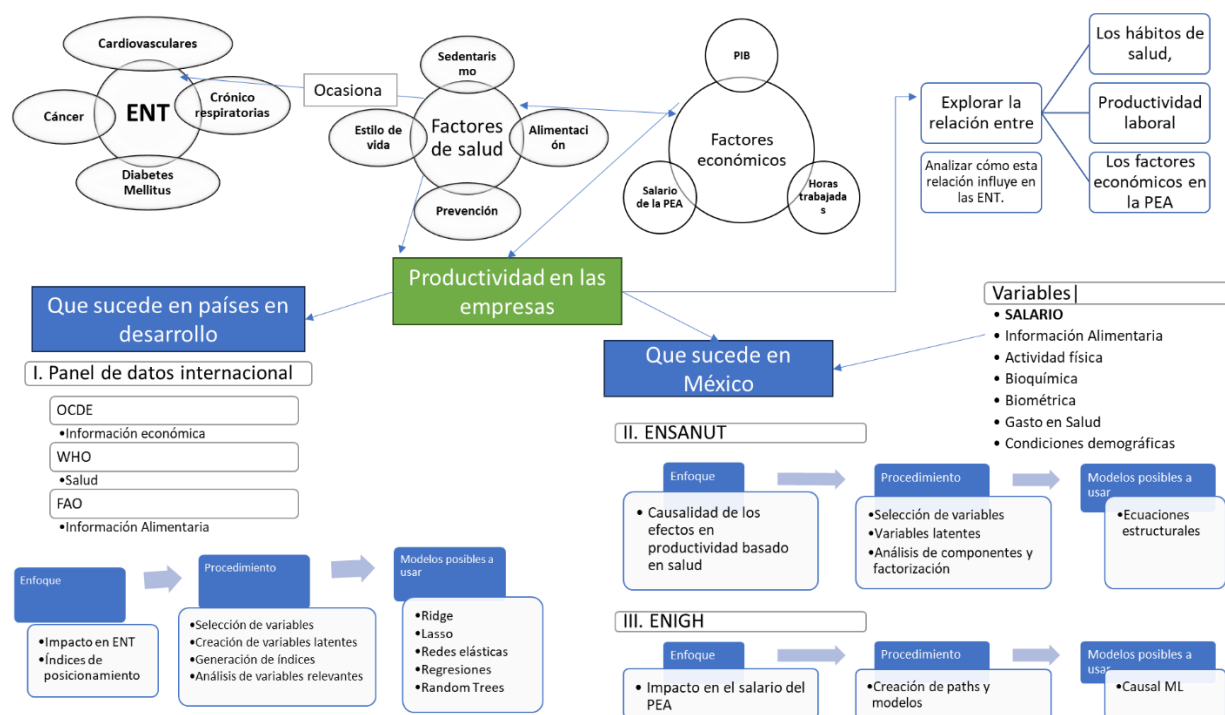
Si bien la unidad de análisis fue el individuo, es importante señalar que a cada persona se le imputaron ciertas características del hogar, como nivel de gasto y tipo de alimentación, ya que estas variables, aunque recolectadas a nivel del hogar, influyen directamente en el contexto individual. Se trabajó con una base de datos depurada, en la que se identificaron y trataron adecuadamente los valores perdidos y los casos atípicos.

3.1.2 Diseño metodológico

Para comprender con mayor profundidad el modo en que las ENT impactan en la productividad laboral, se diseñó un esquema secuencial de análisis que permite abordar los factores de salud y económicos de forma integrada. Esta secuencia metodológica parte de una perspectiva comparativa internacional y culmina en la estimación del efecto causal de las ENT sobre la productividad laboral en México. El propósito general es explorar la interacción entre los hábitos de salud, la prevalencia de ENT y sus consecuencias económicas, en particular en el segmento de la PEA. En la Figura 9 se observa el diagrama del diseño metodológico utilizado.

Figura 9

Esquema del diseño de la investigación



Fuente: Elaboración propia

El presente estudio se desarrolló en tres etapas complementarias articuladas con base en el análisis del impacto de las ENT sobre la salud y la productividad laboral. Estas etapas se describen brevemente en la siguiente sección.

I. Factores relacionados con las ENT en países en desarrollo durante la edad laboral.

Se recurrió a bases de datos provenientes de organismos internacionales tales como la OCDE, el Banco Mundial (WBOD, 2023), el *Institute for Health Metrics and Evaluation* (IHME, 2020), la Organización Mundial de la Salud (WHO, 2023), *American Economic Review* (Fenestra *et al.*, 2022) y la FAO (2023). Se seleccionaron 14 países en desarrollo seleccionados como se menciona en la sección de obtención de muestras que fueron: Polonia, Turquía, Ucrania, Egipto, Marruecos, Sudáfrica, México, Brasil, Chile, Colombia, Argentina, Malasia y Tailandia. Se consideró un filtro para la edad entre 20 y 64 años porque se considera como la edad laboral.

A partir de variables económicas, alimentarias y sanitarias, se construyeron índices compuestos mediante análisis de componentes principales (PCA), lo que permitió sintetizar patrones de riesgo y ubicar a México en un contexto comparado.

Posteriormente, se aplicaron técnicas de regresión penalizada para identificar en orden y magnitud los principales determinantes de mortalidad por ENT en enfermedades específicas (cardiovasculares, respiratorias, cáncer y diabetes).

II. Efecto de factores relacionados con la salud y sus efectos en las ENT en México.

Con el fin de profundizar en la caracterización individual del vínculo entre hábitos de salud y ENT en México. Se empleó la base de datos de la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2022. Se utilizaron modelos de regresión logística y algoritmos de aprendizaje automático, específicamente bosques aleatorios, para seleccionar las variables predictoras relevantes.

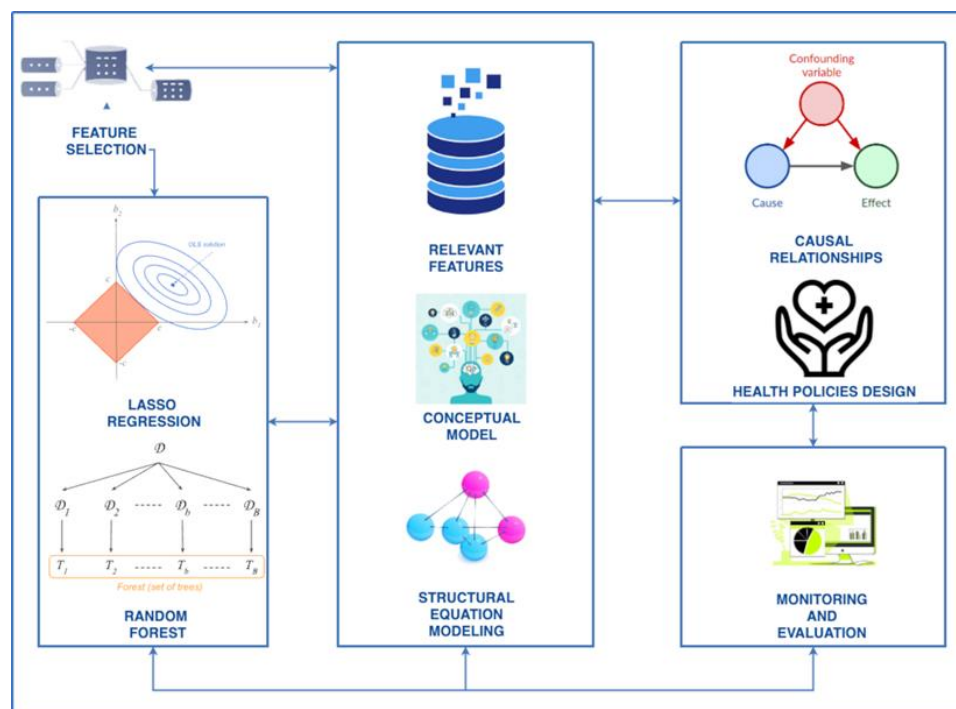
Se implementó un modelo de ecuaciones estructurales (SEM) con el objetivo de construir variables latentes (como hábitos de salud) y estimar sus efectos sobre las ENT permitiendo integrar múltiples indicadores observados en una misma dimensión analítica (Domínguez-Miranda *et al.*, 2023b).

A menudo ocurre que el marco conceptual aceptado no se refleja necesariamente en las variables medidas, por lo que es necesario ajustar dicho marco conceptual para adaptarlo a la realidad observada. En este caso se propone un enfoque inverso, al utilizar modelos de ML es posible identificar de forma exhaustiva grandes conjuntos de variables relacionadas con el fenómeno a analizar y seleccionar un conjunto de variables relevantes. La Figura 10 muestra el

enfoque propuesto del marco analítico híbrido que combina ML y SEM. Se presentan los resultados del mejor modelo incluyendo los coeficientes estimados para cada variable, sus intervalos de confianza y las estadísticas de ajuste del modelo.

Figura 10

Enfoque de análisis híbrido utilizando aprendizaje automático (ML) y modelos de ecuaciones estructurales (SEM)



Fuente: Dominguez-Miranda et al. (2025)

III. Impacto de los hábitos de salud de la PEA en la productividad laboral en México.

Finalmente, en la tercera etapa se recurrió a la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022 para estimar el impacto de las ENT sobre la productividad laboral. Dado que la ENSANUT no recoge información directa sobre productividad laboral ni condiciones económicas individuales, se incorporó en una tercera etapa la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022.

Con el objetivo de generar una base de datos integrada que permitiera análisis comparativos, se identificaron variables equivalentes entre la Encuesta Nacional de Salud y Nutrición (ENSANUT) y la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). Las

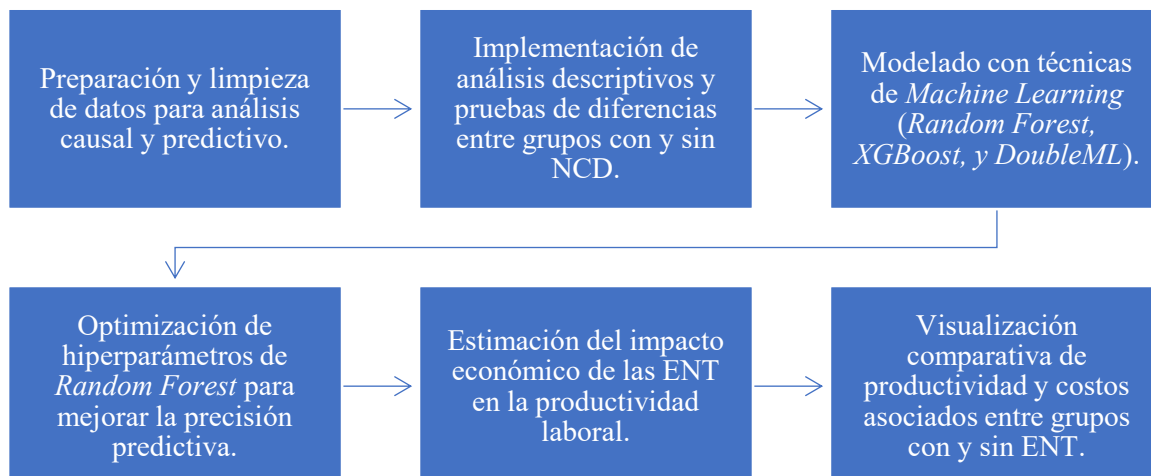
variables seleccionadas incluyeron estado de salud autodeclarado, presencia de enfermedades crónicas, acceso a servicios médicos, gasto en salud y patrones de alimentación. El establecimiento de equivalencias se realizó mediante un mapeo temático que vinculó categorías comunes tales como consumo de alimentos (legumbres, carnes, frutas, verduras), diagnóstico de enfermedades crónicas (diabetes e hipertensión) y características socioeconómicas (nivel educativo, ingresos y región de residencia). Este procedimiento fue validado conforme a los códigos y definiciones oficiales publicados por el Instituto Nacional de Salud Pública y el INEGI.

Se aplicaron modelos de aprendizaje automático y técnicas de DML para estimar el efecto promedio del tratamiento, definido como la presencia de ENT, sobre la variable de respuesta, que corresponde a la productividad laboral individual. Esta etapa permitió cuantificar la penalización en productividad atribuible a las ENT, controlando por múltiples variables confusoras, y estimar su costo económico agregado.

Los pasos realizados incluyeron la preparación de los datos, el ajuste de variables relevantes, la implementación de modelos para estimar el impacto de las ENT y la validación de los resultados conforme a la Figura 11.

Figura 11

Proceso Analítico para Evaluar el Impacto de las ENT mediante Técnicas Estadísticas y de Machine Learning



Fuente: Elaboración propia

Finalmente, se calculó el impacto económico total de las ENT estimando la diferencia en productividad ajustada y su equivalente monetario, a partir de la tarifa horaria promedio. Los resultados fueron visualizados gráficamente destacando la brecha productiva entre grupos.

El entorno experimental para este estudio se configuró en un sistema operativo *Windows* 11 utilizando un procesador *Intel Core i7*, 64 GB de RAM y almacenamiento de 1 *TB HDD*, con tareas computacionales respaldadas por una *GPU NVIDIA RTX 3080*. El análisis se realizó en R 4.4.3. Las bibliotecas clave incluyeron *randomForest*, *ranger* y *randomForestSRC* para los métodos de *machine learning*, *caret* para la validación cruzada de modelos y *tree* para el análisis de árboles de decisión. Para la explicabilidad e interpretabilidad de los modelos de *ML* se utilizaron *randomForestExplainer* y *DALEX*. La implementación del modelo de ecuaciones estructurales (SEM) se llevó a cabo utilizando el paquete *lavaan*.

En conjunto, estas tres fases metodológicas permitieron vincular evidencia internacional, datos nacionales y enfoques estadísticos avanzados para comprender de forma integral cómo los hábitos de salud, las condiciones estructurales y la presencia de ENT impactan en el desempeño laboral de la población activa en México. El resumen de los pasos se observa en la Tabla 8.

Tabla 8

Resumen del diseño metodológico

Pasos	Objetivo	Nombre del estudio	Modelos utilizados
1	Explorar la relación entre las dimensiones económicas, hábitos de salud y laborales con las ENT.	Factores relacionados con las ENT en países en desarrollo durante la edad laboral.	Regresiones lineales <i>Ridge</i> <i>Lasso</i> Redes elásticas PCA Análisis de conglomerados
2	Evaluar el impacto de los hábitos de salud de la PEA en la productividad laboral en México.	Efecto de factores relacionados con la salud y sus efectos en las ENT en México.	Regresiones penalizadas <i>Random Forest</i> Modelos de Ecuaciones Estructurales

3	Determinar el impacto de los hábitos de salud en la productividad laboral en México.	Impacto de los hábitos de salud de la PEA en la productividad laboral en México.	<i>Double Machine Learning</i>
---	--	--	--------------------------------

Fuente: Elaboración propia

A través de esta secuencia metodológica se busca ofrecer un análisis integral de los factores que determinan la salud de la PEA en México, su impacto en la productividad laboral y la magnitud de las desigualdades en contextos emergentes, brindando así un marco empírico sólido para el diseño de intervenciones en salud ocupacional y políticas públicas.

3.2 Muestra y recolección de datos

3.2.2 Muestra

La tarea más importante aquí es definir cuidadosamente la característica a estudiar y luego seleccionar grupos que difieran en esta característica. Como lo explica Azalea (2022), la gran medida en que se definan cuidadosamente los grupos de comparación será el éxito de encontrar patrones de causalidad. Es muy importante seleccionar grupos que sean homogéneos con respecto a algunas variables importantes. Una vez seleccionados los grupos definidos, se pueden emparejar en una o más variables. Este proceso del emparejamiento controla ciertas variables, eliminando así cualquier diferencia de grupo en estas variables.

Países en vías de desarrollo seleccionados

Con el objetivo de realizar un análisis comparativo internacional sobre la desigualdad económica y diversos indicadores de desarrollo, se seleccionó un conjunto de 14 países en desarrollo mediante un proceso de filtrado sistemático. La fuente principal utilizada para identificar estos países fue la base de datos del sitio *World Data Info* (2023), la cual consolida clasificaciones internacionales basadas en niveles de ingreso, desarrollo humano y características estructurales de las economías. Este portal integra información proveniente de instituciones reconocidas como el Banco Mundial, el Fondo Monetario Internacional (FMI) y el Programa de las Naciones Unidas para el Desarrollo (PNUD), proporcionando una lista accesible de países clasificados como "en desarrollo".

A partir de esta lista se recuperó un conjunto inicial de países que fue evaluado según tres indicadores principales: el Ingreso Nacional Bruto (INB) per cápita, el Índice de Desarrollo Humano (IDH) y el Índice de Gini, este último como medida de desigualdad en la distribución del ingreso. Se dio prioridad a aquellos países cuyo INB per cápita se encuentra, de acuerdo con la clasificación del Banco Mundial, en los rangos de ingreso medio-bajo a medio-alto (aproximadamente entre 1,000 y 13,000 USD). Simultáneamente, el IDH permitió identificar países con avances medibles en salud, educación y calidad de vida, mientras que el Índice de Gini sirvió para valorar y comparar las desigualdades internas de cada nación.

Como parte del procedimiento, se calculó el PIB estimado de cada país multiplicando el INB per cápita por la población total, lo cual permitió estandarizar la magnitud económica de las naciones analizadas. Posteriormente se ordenó el listado completo de países en desarrollo de mayor a menor según su PIB estimado con el fin de visualizar y confirmar la posición relativa de los países seleccionados dentro del conjunto. Esta ordenación permitió validar que los 14 países seleccionados: Polonia, Turquía, Ucrania, Egipto, Marruecos, Sudáfrica, México, Brasil, Chile, Colombia, Argentina, India, Malasia y Tailandia, no solo formaban parte del grupo de países en desarrollo, sino que además se encontraban entre las economías con mayor peso dentro de ese grupo y con niveles de desigualdad relevantes al posicionarse también entre los 30 países con mayor índice de Gini a nivel global.

De estos países seleccionados se hizo un filtrado adicional basado en confirmar los países que contaran con disponibilidad sistemática de datos públicos entre 1990 y 2019 en fuentes como la OCDE, el Banco Mundial, el *Institute for Health Metrics and Evaluation* (IHME), la Organización Mundial de la Salud (OMS), el Banco Mundial y la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO).

El periodo 1990–2019 fue delimitado con el fin de capturar una trayectoria evolutiva de tres décadas antes de los efectos disruptivos de la pandemia por COVID-19, permitiendo así analizar tendencias estables en indicadores de salud y productividad sin distorsiones recientes. Asimismo, se filtraron las bases de datos considerando únicamente a personas entre los 20 y 64 años, grupo definido como población en edad de trabajar (PEA), de acuerdo con los estándares de la Organización Internacional del Trabajo (OIT).

Esta primera etapa permitió construir un conjunto equilibrado de observaciones comparables con el objetivo de estimar modelos explicativos robustos que relacionan las ENT, sus factores de riesgo y variables económicas clave como el producto interno bruto, la inversión en salud y la productividad laboral.

Encuesta Nacional de Salud y Nutrición (ENSANUT)

De acuerdo con ENSANUT (2022), el tamaño de muestra fue calculado mediante la ecuación (2):

$$n = Z^2 \frac{p(1-p)}{\delta^2 RK} deff \quad (2)$$

Donde:

n: Tamaño de la muestra en viviendas.

p: Proporción a estimar.

Z: Cuantil 97.5 de una distribución normal unitaria (Z=1.96).

δ : Semi amplitud del intervalo de confianza, $\Pr(P - P \leq \delta) = 0.95$.

R: Tasa de respuesta esperada.

K: Porcentaje de viviendas donde se espera obtener a alguna persona del grupo de interés.

deff = Efecto de diseño de muestra (1.83).

De acuerdo con la metodología de la obtención de los datos de ENSANUT, se consolidó una muestra nacional representativa de los ámbitos urbano/rural, regional y por estados, en todos los grupos de población, obteniéndose 59,604 encuestas respondidas. Dichas encuestas representan estadísticamente a los 23,849,760 individuos en México y cada observación tiene un factor de expansión, el cual fue calculado con las observaciones finales (Romero-Martínez *et al.*, 2022). Las observaciones utilizadas de la ENSANUT (2022) fueron todos los individuos que respondieron a todas las encuestas; cabe destacar que este análisis es a nivel de individuo. Si bien la base de datos contiene muestras de una gran diversidad de individuos, solo serán utilizadas las pertinentes a la edad laboral entre los 20 y los 65 años, obteniéndose 550 observaciones. Esta encuesta tiene representatividad estatal, por estratos nacionales urbano y rural.

Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)

Respecto a la muestra de la ENIGH (2022) los resultados obtenidos de la encuesta se generalizan a toda la población de dominio estudio, estratificado y por conglomerados, donde la unidad última de selección es la vivienda y la unidad de observación es el hogar y a nivel individual, las unidades primarias de muestreo fueron seleccionadas por medio de un muestreo con probabilidad proporcional al tamaño mediante la ecuación (3):

$$P\{U_{ehi} \in S_m\} = \frac{k_{eh}m_{ehi}}{m_{eh}} \quad (3)$$

Donde:

U_{ehi} : la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

S_m : la muestra maestra.

k_{eh} : número de UPM en la muestra maestra, en el h-ésimo estrato, en la e-ésima entidad.

m_{ehi} : número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad en el Censo de Población y Vivienda 2010.

m_{eh} : número de viviendas en el h-ésimo estrato, en la e-ésima entidad en el Censo de Población y Vivienda 2010 (Véase ecuación 4).

$$F_D'' = F_D' \frac{PEST_D}{PEXP_D} \quad (4)$$

Donde:

F_D'' : factor de expansión corregido por proyección en el dominio D.

F_D' : factor de expansión corregido por no respuesta en el dominio D.

$PEST_D$: población en el dominio D, según estimación de población de INEGI.

$PEXP_D$: población total a la que se expande la encuesta en el dominio D.

D : es el nivel de desagregación de la población al que se realiza el ajuste por estimación de población.

Como lo señala el INEGI (2023), “la muestra fue diseñada para proporcionar estimaciones con precisión adecuada para los principales indicadores a nivel nacional, por entidad federativa y tipo de localidad”. La información recolectada se organiza en módulos que abarcan ingresos laborales y no laborales, transferencias, remesas, gastos monetarios y no monetarios, así como características sociodemográficas y de salud de los hogares. Esta estructura convierte a la ENIGH

en una fuente confiable y robusta, ampliamente utilizada para evaluar condiciones de vida y desigualdades en México, además de facilitar comparaciones entre distintos grupos poblacionales y monitorear cambios a lo largo del tiempo.

La ENIGH se estructura en varios grupos de tablas temáticas que abarcan aspectos clave del entorno familiar y económico de los hogares mexicanos. Entre estas se incluyen las tablas de Hogares y Viviendas que describen las características generales del hogar y su infraestructura; Población que contienen información sociodemográfica de los integrantes; Trabajo que detallan las condiciones laborales y el tiempo dedicado a actividades económicas; Ingresos que registran percepciones monetarias y no monetarias provenientes de diversas fuentes y Gastos que documentan tanto el consumo monetario como no monetario en rubros como alimentación, salud, educación y transporte. Adicionalmente, se incluyen tablas sobre bienes durables y condiciones socioeconómicas del hogar. Para este estudio se seleccionaron únicamente aquellas tablas que contenían variables comparables con las disponibles en la ENSANUT, priorizando las relacionadas con población, ingresos, gastos, trabajo y estructura del hogar, a fin de facilitar un análisis integrado entre condiciones económicas y de salud.

Para el presente estudio se seleccionaron aquellas tablas que permitieran establecer equivalencias con las variables clave disponibles en la Encuesta Nacional de Salud y Nutrición (ENSANUT), específicamente en los temas de población, gasto, ingreso, trabajo y hogar. En particular se integraron las tablas de los hogares y viviendas que ofrecen información sobre la estructura del hogar y sus características generales, la tabla de trabajo que detalla las horas laborales y ocupación de la PEA, ingresos corrientes que permite identificar los flujos de recursos monetarios y no monetarios percibidos y la tabla de gastos monetarios y no monetarios que proporciona un desglose detallado de las erogaciones en salud, alimentación y otros rubros esenciales a nivel individual.

La base de datos utilizada en este estudio proviene de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) correspondiente al año 2022, la cual fue descargada y trabajada a nivel de individuo. Esta base contenía inicialmente 530,603 observaciones y 39 variables relacionadas con características sociodemográficas, ingresos, gastos, hábitos alimenticios y algunas condiciones de salud autorreportadas por los jefes de hogar.

Se realizó una depuración de las variables categóricas, se llevó a cabo un proceso de filtrado con el objetivo de excluir observaciones con valores atípicos o extremadamente elevados con el fin de preservar la calidad y la consistencia de los datos. Se identificaron variables con valores faltantes, y en los casos donde los datos ausentes representaban conceptos susceptibles de asumir un valor nulo (como ciertos gastos no realizados), se sustituyeron los valores faltantes por ceros.

Como resultado de este proceso de depuración, transformación y limpieza, se obtuvo una base de datos final compuesta por 57,182 registros y 35 variables a analizar que sirvió como insumo para el desarrollo de los análisis estadísticos y econométricos presentados en este trabajo.

3.3 Pruebas paramétricas y no paramétricas

En esta sección se contemplan las pruebas y algoritmos matemáticos a realizar con base en modelos estadísticos recomendados en la revisión de la literatura y los modelos utilizados de *machine learning* seleccionados para encontrar patrones de los indicadores enfocados al impacto en la productividad de los empleados.

3.3.1 Pruebas paramétricas

Regresión logística

En un modelo de regresión logística binomial se utilizan ciertos fundamentos teóricos y cálculos matemáticos para estimar cómo las variables predictoras influyen en la probabilidad de que ocurra un resultado binario (Hosmer Jr. *et al.*, 2013).

La regresión logística modela la probabilidad de que ocurra un evento utilizando la función logística, que es una curva sigmoideal que toma valores entre 0 y 1. Esta función se define mediante la ecuación (5).

$$P(Y = 1|X) = \frac{1}{1 + e^{-\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \dots - \beta_p X_p}} \quad (5)$$

Donde:

$\beta_0, \beta_1, \dots, \beta_p$: son los coeficientes de regresión.

$X_1, X_2, \dots, X_p$: son las variables independientes.

La función de verosimilitud se utiliza para estimar los parámetros del modelo. Busca maximizar la probabilidad de observar los datos bajo el modelo. La función de verosimilitud se muestra en la ecuación (6).

$$L(y|x; \beta) = \left(\frac{1}{1 + e^{-x^T \beta}} \right)^y \left(\frac{e^{-x^T \beta}}{1 + e^{-x^T \beta}} \right)^{1-y} \quad (6)$$

Se calculan intervalos de confianza para los coeficientes de regresión. Se utilizan las estimaciones de los coeficientes para hacer predicciones sobre las probabilidades de ocurrencia del evento. Dados los valores de las variables predictoras, se calcula la probabilidad de que Y sea 1.

Se puede evaluar la calidad del modelo utilizando diversas métricas tomando en cuenta el estadístico de prueba de Wald, el estadístico de razón de verosimilitudes, el AIC y el BIC. Para esto, los coeficientes $\beta_0, \beta_1, \dots, \beta_p$ se estiman ajustando el modelo a los datos.

Utilizando métodos como la máxima verosimilitud se pueden encontrar los valores óptimos de los coeficientes, así como calcular los coeficientes de Wald para evaluar la significancia de los coeficientes de regresión. El coeficiente de Wald para β_i se muestra en la ecuación (7).

$$Wald_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (7)$$

Donde:

$\hat{\beta}_i$: Coeficiente estimado.

$SE(\hat{\beta}_i)$: Error estándar del coeficiente.

La técnica de criterio de selección de información de Akaike (AIC) para la selección de modelos en el contexto de la estimación de máxima verosimilitud, el AIC proporciona una medida de la capacidad predictiva de un modelo ajustado en función de la complejidad del modelo. El AIC penaliza la complejidad del modelo, lo que significa que premia los modelos que logran un buen ajuste a los datos con el menor número posible de parámetros (Ver ecuación 8).

$$AIC = N * \ln \frac{SS_E}{N} + 2K \quad (8)$$

Donde:

- N : Número de observaciones.
- SS_E : El cuadrado medio del error.
- K : Número de parámetros.

Para realizar una selección adecuada del mejor modelo, se puede realizar una comparativa por medio de una prueba de análisis de varianza (ANOVA), que es una técnica estadística utilizada para comparar la media de dos o más grupos y determinar si hay diferencias significativas entre ellos. La ecuación general para el ANOVA depende del tipo específico de ANOVA que se esté realizando, siendo la de un factor el tipo más común utilizado cuando se compara un conjunto de datos entre tres o más grupos (Ver ecuación 9).

La ecuación para el ANOVA de un factor es:

$$F = \frac{MS_{entre\ grupos}}{MS_{dentro\ de\ grupos}} \quad (9)$$

Donde:

- F es el estadístico de prueba de la ANOVA, conocido como la razón de varianzas.
- $MS_{entre\ grupos}$ Es la media de las sumas de cuadrados entre los grupos, también conocida como varianza entre grupos, y se calcula como la suma de cuadrados entre los grupos dividida por los grados de libertad entre grupos.
- $MS_{dentro\ de\ grupos}$ Es la media de las sumas de cuadrados dentro de los grupos, también conocida como varianza dentro de los grupos o error residual, y se calcula como la suma de cuadrados dentro de los grupos dividida por los grados de libertad dentro de los grupos.

Ecuaciones estructurales

Las ecuaciones estructurales (SEM, por sus siglas en inglés: *Structural Equation Modeling*) constituyen una metodología estadística multivariada que permite analizar simultáneamente relaciones de dependencia entre variables observadas y latentes dentro de un sistema teórico. Este enfoque combina aspectos de análisis factorial y regresión múltiple, permitiendo modelar constructos no observables mediante indicadores empíricos, así como evaluar efectos directos e

indirectos en un mismo marco analítico. Su uso es especialmente pertinente cuando se requiere validar modelos teóricos complejos que involucran mediadores, moderadores o estructuras jerárquicas, bajo supuestos de causalidad basada en correlaciones (Kline, 2016).

La complejidad de los sistemas de salud y la naturaleza multifacética de los determinantes de la salud requieren técnicas analíticas avanzadas. Las SEM, como enfoque estadístico multivariado, permiten analizar simultáneamente múltiples variables dependientes e independientes, lo que las hace adecuadas para examinar relaciones intrincadas (Hoyle, 2012; Bollen, 1989; Kline, 2023).

La metodología puede dividirse en dos componentes: el modelo de medición, que captura la relación entre variables latentes y sus indicadores observados, y el modelo estructural, que examina las relaciones entre variables latentes. La aplicación de las SEM requiere hipótesis claramente definidas sobre las relaciones causales que se representan gráficamente mediante diagramas de ruta. Estos diagramas facilitan la interpretación de efectos directos, indirectos y totales entre las variables (Bentler & Chou, 1987).

De acuerdo con Kline (2023), una vez que haya seleccionado el mejor modelo, se enfoca en la utilización de las variables más significativas para la generación de un modelo basado en ecuaciones estructurales. Los pasos para la realización de un modelaje estructural son:

1. Especificación del modelo: Se define un modelo teórico que describe las relaciones entre las variables observadas y no observadas. Esto incluye la especificación de las relaciones causales entre las variables latentes y las observadas.
2. Identificación del modelo: Se determina si el modelo es identificable, es decir, si las relaciones entre las variables pueden ser estimadas de manera única a partir de los datos observados.
3. Estimación del modelo: Se utilizan métodos estadísticos para estimar los parámetros del modelo, como el método de máxima verosimilitud o mínimos cuadrados parciales.
4. Evaluación del modelo: Se evalúa la bondad de ajuste del modelo utilizando diferentes medidas, como el índice de ajuste comparativo (CFI), el índice de *Tucker-Lewis* (TLI), el error cuadrático medio de aproximación (RMSEA) y el índice de ajuste del error cuadrático medio (SRMR).

5. Modificación del modelo: Se realizan ajustes al modelo según sea necesario para mejorar su ajuste a los datos observados. Esto puede incluir la adición o eliminación de variables, la modificación de las relaciones entre las variables y la incorporación de correlaciones entre los errores.
6. Validación del modelo: Se valida el modelo utilizando técnicas como el análisis de invarianza factorial para verificar si el modelo es válido en diferentes subgrupos de la muestra.
7. Interpretación de los resultados: Se interpretan los parámetros del modelo estimado y se discuten las implicaciones teóricas de los hallazgos.

La formulación general de las SEM consta de dos conjuntos de ecuaciones. El modelo de medición se muestra en las ecuaciones (10) y (11).

$$X = \Lambda_x \xi + \delta \quad (10)$$

$$Y = \Lambda_y \eta + \varepsilon \quad (11)$$

Donde:

X y Y : vectores de variables observadas.

ξ y η : vectores de variables latentes.

Λ_x y Λ_y : matrices de cargas factoriales.

δ y ε : términos de error de medición.

El modelo estructural se representa mediante la ecuación (12):

$$\eta = B_n + \Gamma_\xi + \zeta \quad (12)$$

Donde:

B : matriz de relaciones entre variables latentes endógenas,

Γ : matriz de relaciones entre variables latentes exógenas y endógenas

ζ : vector de errores residuales (Schumacker & Lomax, 2016).

3.3.2 Pruebas no paramétricas

Análisis de componentes principales (PCA)

El análisis de componentes principales (PCA, por sus siglas en inglés: *Principal Component Analysis*) es una técnica estadística de reducción de dimensionalidad que transforma un conjunto de variables correlacionadas en un nuevo conjunto de variables no correlacionadas denominadas componentes principales. Estos componentes son combinaciones lineales de las variables originales y se ordenan de forma que el primero explica la mayor varianza posible seguido por los siguientes en orden descendente de varianza explicada. El PCA permite simplificar conjuntos de datos complejos conservando la mayor parte de la información, lo cual resulta útil para identificar patrones subyacentes, eliminar redundancias y mejorar la interpretación de los datos en contextos exploratorios y predictivos (Jolliffe & Cadima, 2016).

Como inicio de un PCA, es relevante realizar una exploración descriptiva de los datos mediante estadísticas básicas (media y varianza), análisis de correlación y visualización con diagramas de dispersión, y en caso de existir sesgos por escalas heterogéneas, las variables deben ser estandarizadas.

Se inicia con la evaluación de la varianza explicada (PVE) por cada componente, así como su acumulado, a fin de determinar el número óptimo de componentes a conservar, seguida de la descomposición de la matriz de covarianza, obteniendo los vectores y valores propios, lo que permite identificar los componentes principales que capturan la mayor variabilidad del sistema.

Posteriormente, se deben calcular las puntuaciones de cada componente para todas las observaciones, representando así sus coordenadas en el nuevo espacio reducido. La interpretación de los componentes se basa en el análisis de las cargas factoriales, donde se identifican las variables con mayor contribución y se interpreta la dirección de su influencia mediante el signo de las cargas.

Una vez que se determinan las ponderaciones de los componentes principales, se calcula un índice compuesto agregando las puntuaciones de los componentes principales para cada observación, ponderadas por las ponderaciones correspondientes, como se refleja en la ecuación (13).

$$Y_i = \frac{x_{1i}w_1 + x_{2i}w_2 + \dots + x_{ki}w_k}{w_1 + w_2 + \dots + w_k} \quad (13)$$

Donde Y_i es el valor del índice para cada conjunto de variables i , x_{ki} es el valor de la coordenada del factor k seleccionado y w_i es la proporción de la inercia total explicada por esa coordenada. El índice compuesto proporciona una medida resumida derivada de los componentes principales y sus ponderaciones que refleja los patrones y relaciones subyacentes capturados por el análisis PCA (Shlens, 2014). Como elemento final, es importante la construcción de índices estandarizados donde el valor de Y_i se ajusta a un rango entre $[0,100]$, donde un valor más alto indica un mayor equilibrio o carga, según el tipo de dimensión analizada.

Regresiones penalizadas (Ridge, Lasso y Redes elásticas)

Las regresiones penalizadas son técnicas estadísticas utilizadas para mejorar la estimación de modelos de regresión cuando existen problemas de multicolinealidad o cuando el número de predictores es grande en relación con el número de observaciones. Estas técnicas introducen un término de penalización en la función de pérdida del modelo, lo que permite controlar la complejidad del modelo y evitar el sobreajuste. Entre las más utilizadas se encuentran *Ridge*, *Lasso* y *Elastic Net* que resultan especialmente útiles en contextos de análisis predictivo con alta dimensionalidad y son ampliamente empleadas en epidemiología, bioestadística y aprendizaje automático (Hastie *et al.*, 2015).

La regresión de *Ridge* es una técnica de regularización que agrega un término de penalización a la regresión de mínimos cuadrados ordinarios (Schreiber-Gregory, 2018). Este término de penalización es proporcional a la suma de los coeficientes al cuadrado multiplicado por un hiperparámetro λ . La regresión de *Ridge* fomenta valores de coeficiente más pequeños y reduce el impacto de predictores irrelevantes mientras mantiene todas las variables en el modelo (Ver ecuación 14). En la práctica, el valor óptimo del hiperparámetro λ a menudo se determina mediante validación cruzada u otras técnicas de selección de modelos para lograr el equilibrio adecuado entre la complejidad del modelo y la generalización (Hesterberg *et al.*, 2008).

Para incorporar la penalización L_1 en la regresión de *Lasso* la función objetivo se modifica de la siguiente manera (Ver ecuación 14):

$$\sum_{i=1}^N \sum_{t=1}^T (y_{it} - x_{it}\beta)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (14)$$

Donde p es el número total de variables independientes y λ es el parámetro de regularización que controla la fuerza de la penalización.

La regresión *Lasso* u operador de contracción y selección mínima absoluta es otra técnica de regularización muy utilizada. Introduce un término de penalización λ_1 a la función de pérdida y que promueve la dispersión en el vector coeficiente al alentar a que algunos coeficientes sean exactamente cero. Para resolver el problema de regresión de *Lasso* se pueden emplear varios algoritmos de optimización como el descenso de coordenadas o la regresión de ángulo mínimo (Friedman *et al.*, 2010). Estos algoritmos actualizan iterativamente los coeficientes para minimizar la función objetivo (Hesterberg *et al.*, 2008). La solución de forma cerrada para la regresión de *Lasso* no existe debido a la no diferenciabilidad de la norma λ_1 . Sin embargo, los algoritmos de optimización eficientes pueden encontrar la solución de manera efectiva (Hastie *et al.*, 2015).

Para incorporar la penalización L_2 en la regresión de *Lasso* la función objetivo se modifica de la siguiente manera (Ver ecuación 15):

$$\sum_{i=1}^N \sum_{t=1}^T (y_{it} - x_{it}\beta)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (15)$$

Donde p es el número total de variables independientes y λ es el parámetro de regularización que controla la fuerza de la penalización.

La primera parte de la función objetivo minimiza la suma de residuos al cuadrado representando el ajuste del modelo a los datos. La segunda parte introduce la penalización de regresión λ_1 para *Lasso* y λ_2 para *Ridge* que promueve la escasez y fomenta que algunos coeficientes sean exactamente cero, facilitando la selección de variables tanto en países como en períodos de tiempo.

La regresión de redes elásticas combina las fortalezas de la regresión *Ridge* y *Lasso* agregando términos de penalización λ_1 y λ_2 a la función de pérdida. Este enfoque híbrido aborda las limitaciones de cada método y logra un equilibrio entre la regularización y la selección de variables. Para resolver el problema de regresión de red elástica, se pueden utilizar varios

algoritmos de optimización, como el descenso de coordenadas o el descenso de gradiente proximal. Estos algoritmos actualizan iterativamente los coeficientes para minimizar la función objetivo. La solución de forma cerrada para la regresión de red elástica tiene dificultades debido a la no diferenciabilidad del estándar λ_1 . Sin embargo, los algoritmos de optimización eficientes pueden encontrar efectivamente la solución (Zou & Hastie, 2005). La regresión de red elástica proporciona un marco flexible que se adapta tanto a predictores dispersos como correlacionados, lo que la hace particularmente adecuada para aplicaciones de atención médica donde puede haber una combinación de predictores fuertes y débiles.

Para incorporar penalizaciones λ_1 y λ_2 en la regresión de red elástica, la función objetivo se plantea de la siguiente manera (Ver ecuación 16).

$$\min_{\hat{\beta}} \left\{ \sum_{i=1}^n (y_{it} - x_{it}^T x_{it} \hat{\beta})^2 + \lambda_1 \sum_{j=1}^m |\hat{\beta}_j| + \lambda_2 \sum_{j=1}^m |\beta_j|^2 \hat{\beta}_j^2 \right\} \quad (16)$$

Donde p representa el número total de variables independientes, y λ_1 y λ_2 son los parámetros de regularización que controlan la intensidad de las penalizaciones L_1 (absoluta) y L_2 (cuadrática), respectivamente.

En esta etapa, se verifica la evaluación de los métodos, la integridad y exactitud de los datos obtenidos y evaluados. Se sugiere realizar una comparación del rendimiento de los modelos utilizando medidas de rendimiento basadas en los valores de precisión con error cuadrático medio (MSE), R-cuadrado (R^2) y puntuación de varianza explicada (EVS) (Ver ecuaciones 17-19) (Kline, 2023).

MSE se calcula utilizando:

$$MSE = \frac{1}{2} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (17)$$

R^2 viene dada por:

$$R^2 = 1 - \left(\frac{\sum (y_i - \hat{y}_i)^2}{\sum (y - \hat{y})^2} \right) \quad (18)$$

Y EVS viene dado por:

$$EVS = 1 - \frac{Var(y - \hat{y})}{Var(y)} \quad (19)$$

Donde y representa el valor estimado, $\hat{y}$ es el valor real y n es el número de observaciones.

Random Forest

El algoritmo *Random Forest* es una técnica de aprendizaje automático basada en conjuntos (*ensemble learning*) que combina múltiples árboles de decisión para mejorar la precisión predictiva y la estabilidad del modelo. Esta metodología consiste en construir un número determinado de árboles sobre diferentes subconjuntos aleatorios del conjunto de datos y variables, y luego combinar sus resultados mediante votación (para clasificación) o promedio (para regresión). Al reducir la varianza y mitigar el sobreajuste típico de los árboles individuales, *Random Forest* se ha consolidado como una herramienta robusta tanto en contextos supervisados como en exploración de datos de alta dimensionalidad (Breiman, 2001). Su capacidad para manejar grandes volúmenes de datos y su eficiencia en la selección de variables lo hacen ampliamente aplicable en ciencias de la salud, economía y bioinformática.

El modelo de *Random Forest* es utilizado para capturar relaciones no lineales y posibles interacciones entre variables. Este modelo es robusto y adecuado para manejar datos complejos como los de productividad o salud. Los pasos para el modelaje por medio de este método se enumeran a continuación (Louppe, 2014):

1. Estructura del árbol de decisiones: se determinará la estructura y los parámetros de su árbol de decisión como el criterio de división, la profundidad máxima del árbol y la cantidad mínima de muestras necesarias para dividir un nodo. La forma de declarar la decisión de la división de los nodos es realizada por la ecuación (20):

$$\Delta i(s, t) = i(t) - p_L i(t_L) - p_R i(t_R) \quad (20)$$

Donde:

p_L : proporción $\frac{N_{t_L}}{N_t}$

Siendo t_L y t_R la división de los nodos a la derecha y la izquierda.

2. Entrenamiento del árbol de decisiones: Utilizado para el conjunto de entrenamiento para entrenar el árbol de decisiones al porcentaje evaluado por el investigador, como sugerencia del 30% al 50% de la muestra. El algoritmo del árbol de decisiones dividirá recursivamente

los datos según los criterios de división seleccionados, creando nodos y ramas para formar la estructura de árbol. Posteriormente se evalúan diferentes estructuras de árbol y parámetros utilizando el conjunto de validación, finalizando con el ajuste de los parámetros para mejorar el rendimiento del modelo y evitar el sobreajuste.

3. Interpretación del modelo: Se interpreta el árbol de decisiones para obtener información sobre las relaciones y las reglas de decisión aprendidas por el modelo, así como analizar las divisiones y la importancia de las características al hacer predicciones.
4. Se sugiere una visualización del árbol de decisiones para proporcionar una representación clara del proceso de toma de decisiones.

En el modelo de *Random Forest* basado en un conjunto de árboles de decisión ($T_1, T_2, \dots, T_B$) cada árbol es entrenado en una muestra aleatoria del conjunto de datos teniendo como enfoque a la predicción final ($\hat{Y}$) que representa el promedio (regresión) o el voto mayoritario (clasificación) de todos los árboles. Se puede calcular mediante la ecuación (21):

$$\hat{Y} = \frac{1}{B} \sum_{b=1}^B f_b(X) \quad (21)$$

Donde:

- $\hat{Y}$ es la predicción final.
- B es el número de árboles en el bosque.
- $f_b(X)$ es la predicción del árbol b para la variable dependiente.

Gradient Boosting (XGBoost)

El método de *Gradient Boosting*, y particularmente su implementación optimizada *XGBoost (Extreme Gradient Boosting)*, es una técnica de aprendizaje automático basada en el principio de ensamblaje secuencial de modelos débiles, como los árboles de decisión, con el objetivo de corregir iterativamente los errores cometidos por modelos anteriores. A diferencia de *Random Forest*, que construye los árboles en paralelo, *XGBoost* construye los árboles de forma secuencial, mejorando el ajuste en cada iteración mediante la minimización de una función de pérdida diferenciable. *XGBoost* ha demostrado alta eficiencia computacional, manejo efectivo de datos faltantes y control del sobreajuste mediante regularización, lo que lo convierte en una

herramienta ampliamente utilizada en competencias de ciencia de datos y aplicaciones en salud, finanzas y análisis de riesgos (Chen & Guestrin, 2016).

El algoritmo *XGBoost* es conocido por su eficiencia y alta capacidad predictiva y se deben contemplar los siguientes pasos:

- Preparación de datos: Los datos se transforman en una matriz de diseño compatible con el formato que requiere *XGBoost*.
- Entrenamiento: Se entrena al modelo con un objetivo de error cuadrático medio y diversas rondas de iteración.
- Evaluación:
 - La importancia de las variables se calcula y visualiza mediante gráficos.
 - Se implementa validación cruzada con cinco particiones para evaluar la estabilidad del modelo.

Gradient Boosting combina predicciones de múltiples árboles de decisión secuenciales, ajustando cada árbol para minimizar la función de pérdida (Chen & Guestrin, 2016) (Ver ecuación 22):

$$L(y, \hat{y}) = \sum_{i=1}^N \ell(y_i, \hat{y}_i) + \Omega(f_k) \quad (22)$$

Donde:

- $L(y, \hat{y})$: Función de pérdida total.
- $\ell(y_i, \hat{y}_i)$: Función de pérdida para una observación.
- $\Omega(f_k)$: Penalización de regularización.
- (f_k) : Función de predicción de cada árbol.

XGBoost optimiza el modelo ajustando cada árbol en función de los errores residuales.

El algoritmo *XGBoost* es particularmente relevante para el análisis en salud debido a su capacidad para manejar grandes volúmenes de datos heterogéneos, detectar interacciones complejas entre variables clínicas y no clínicas, y ofrecer predicciones robustas aun en presencia de ruido o datos faltantes. En contextos sanitarios, donde los factores que influyen en los resultados de salud suelen ser no lineales y altamente interdependientes (por ejemplo, edad, comorbilidades, hábitos de vida y determinantes sociales), *XGBoost* permite modelar estas relaciones de manera

eficaz gracias a su enfoque secuencial de aprendizaje y a la incorporación de técnicas de regularización que mitigan el sobreajuste. Además, su alta precisión predictiva y capacidad de interpretación a través de la importancia de variables lo hacen útil tanto para la investigación epidemiológica como para aplicaciones clínicas, como la predicción de enfermedades crónicas, la clasificación de riesgo o la priorización de intervenciones médicas personalizadas.

Double Machine Learning (DML)

El objetivo principal del DML es aislar el efecto neto de las enfermedades no transmisibles (ENT) sobre la productividad laboral, eliminando el sesgo generado por múltiples variables de confusión. En estudios observacionales, como el presente, la productividad no depende únicamente de la presencia de una enfermedad, sino también de factores como la edad, el nivel educativo, el tipo de empleo, los ingresos, los hábitos de salud y las condiciones laborales. Si estos factores no se controlan adecuadamente, el efecto atribuido a la enfermedad puede estar distorsionado. El DML se utiliza precisamente para “limpiar” este sesgo, separando la influencia de las variables de confusión y permitiendo identificar el impacto causal específico de las ENT sobre la productividad laboral.

Desde un punto de vista técnico, el DML es un enfoque estadístico diseñado para estimar efectos causales en contextos donde existen múltiples covariables y posibles relaciones no lineales entre variables. Este método combina técnicas de aprendizaje automático con herramientas de inferencia causal, dividiendo el proceso de estimación en dos etapas. En la primera, se estiman de manera flexible las funciones condicionales del resultado y del tratamiento utilizando algoritmos de aprendizaje automático, como *Random Forest* o *XGBoost*. En la segunda etapa, se utiliza la parte de la variabilidad no explicada por estas funciones para estimar el efecto causal de interés.

El DML permite controlar de manera eficiente los sesgos causados por variables confusoras incluso en escenarios de alta dimensionalidad, multicolinealidad o relaciones no lineales, condiciones frecuentes en bases de datos de salud y economía. Su estructura de estimación doble y el uso del principio de orthogonalización hacen que el estimador del efecto causal sea menos sensible a errores en la especificación de los modelos auxiliares, lo que lo convierte en una herramienta especialmente adecuada para estudios observacionales en los que la asignación aleatoria no es factible (Chernozhukov et al., 2018).

El enfoque DML se implementa siguiendo una secuencia metodológica estructurada que comprende los siguientes pasos:

- Preparación de los datos: Los datos se transforman en un formato compatible con los algoritmos de aprendizaje automático, convirtiendo variables categóricas en representaciones numéricas y definiendo la variable de tratamiento (ENT) como una variable binaria. Asimismo, se especifican la variable dependiente (productividad laboral ajustada), la variable de tratamiento y el conjunto de variables explicativas.
- Creación del objeto DML: Se utiliza la clase *DoubleMLData* para estructurar formalmente los datos, permitiendo su procesamiento dentro del marco metodológico de *Double Machine Learning*.
- Configuración de los modelos de aprendizaje: Se definen modelos de aprendizaje automático para las dos ecuaciones auxiliares: una para modelar la variable dependiente y otra para modelar la probabilidad de tratamiento, empleando algoritmos como *Random Forest* o *XGBoost*.
- Ajuste del modelo y estimación causal: El modelo DML se ajusta combinando los resultados de ambas ecuaciones auxiliares, lo que permite obtener una estimación robusta del efecto causal de las ENT sobre la productividad laboral.

De manera formal, el DML estima efectos causales al dividir el análisis en dos etapas: modelado de la variable dependiente (Y) y modelado de la variable de tratamiento (D). (Ver ecuación 24).

$$Y = D\theta + g(X) + \varepsilon \quad (24)$$

Donde:

- Y : Variable dependiente (productividad ajustada).
- D : Variable de tratamiento (ENT).
- θ : Efecto causal estimado.
- $g(X)$: Función flexible que captura otras variables explicativas (X).

La técnica DML es robusta frente a confusores y garantiza estimaciones válidas de efectos causales.

El uso de *Double Machine Learning* en esta investigación permite combinar el poder predictivo del aprendizaje automático con los principios de la inferencia causal, proporcionando estimaciones sólidas y robustas del impacto de las ENT en la productividad laboral, lo que resulta fundamental para generar evidencia empírica confiable que apoye la toma de decisiones en materia de salud ocupacional y política pública.

CAPÍTULO IV. Resultados

Esta sección presenta los principales hallazgos derivados del análisis cuantitativo, estructurados conforme a los objetivos específicos del estudio y al orden metodológico previamente establecido. El propósito general fue explorar la relación entre dimensiones económicas, hábitos de salud y condiciones laborales con la prevalencia de ENT, así como su impacto en la productividad laboral durante la etapa de vida económicamente activa. Para ello, se empleó una estrategia analítica que combinó técnicas estadísticas clásicas con métodos avanzados de aprendizaje automático y análisis multivariado.

Los resultados que se presentan a continuación están organizados de acuerdo con esta secuencia analítica, permitiendo observar de forma integrada cómo se vinculan los factores de salud, económicos y laborales con la prevalencia de ENT y la productividad individual en el contexto mexicano.

4.1 Factores relacionados con las ENT en países en desarrollo durante la edad laboral

Con el objetivo de identificar los factores asociados a la prevalencia de ENT en contextos laborales internacionales, se realizó un análisis comparativo centrado en países en desarrollo. Este ejercicio se fundamentó en bases de datos públicas y oficiales como el Banco Mundial, la Organización Mundial de la Salud (OMS) y otras fuentes internacionales a partir de las cuales se seleccionaron variables representativas de las dimensiones económica (PIB, gasto en salud, inversión pública), de salud (mortalidad por ENT, esperanza de vida, acceso a servicios) y de nutrición (índices de alimentación saludable y balanceada)².

Mediante un análisis de componentes principales (PCA), se construyeron índices sintéticos que permitieron reducir la dimensionalidad de la información, manteniendo las principales estructuras de variabilidad. Estos índices fueron utilizados para llevar a cabo un análisis de conglomerados a fin de identificar agrupamientos naturales entre los países y ubicar la posición relativa de México respecto a sus homólogos.

La ecuación (25) del índice para cada observación i , basada en los componentes seleccionados y sus respectivas varianzas explicadas, se expresa como:

² Véase Diagrama de Inferencia 1 en Anexo 2, Tabla 27.

$$I_i = \frac{\sum_{k=1}^K x_{ki} w_k}{\sum_{k=1}^K w_k} \quad (25)$$

donde:

- I_i es el valor del índice compuesto para los países en desarrollo seleccionados (i).
- x_{ki} es la coordenada (score) del país i en el componente principal k .
- w_k representa la proporción de varianza explicada por el componente k .
- K es el número total de componentes seleccionados para esa dimensión (económica, alimentaria, ENT).

Posteriormente, el valor de I_i fue rescatado linealmente al rango $[0,100]$ para facilitar su interpretación, siguiendo la transformación observada en la ecuación (26):

$$I_i^{(esc)} = \frac{I_i - \min(I)}{\max(I) - \min(I)} * 100 \quad (26)$$

El índice económico construido a partir de dos componentes principales que explican el 86.84% de la varianza total, refleja la eficiencia en el uso de recursos y la solidez del entorno económico relativo de cada país. Los valores altos (ceranos a 100) indican contextos económicos más favorables.

Por su parte, el índice de balance alimentario, derivado de cuatro componentes principales que explican el 78.06% de la varianza, representa la disponibilidad y el equilibrio de alimentos saludables en cada nación de modo que puntuaciones elevadas reflejan una mejor calidad y diversidad alimentaria. En contraste, el índice de factores de riesgo de ENT fue diseñado para capturar la carga negativa asociada a condiciones como obesidad, consumo nocivo de sustancias y baja actividad física. Este índice se construyó con dos componentes principales y explica entre el 88% y el 94% de la varianza dependiendo del grupo específico de enfermedades consideradas (cardiovasculares, diabetes, neoplasias y enfermedades respiratorias). Los resultados se observan en la Tabla 9.

Tabla 9

Varianza explicada acumulada por las dimensiones seleccionadas en el análisis de componentes principales (PCA)

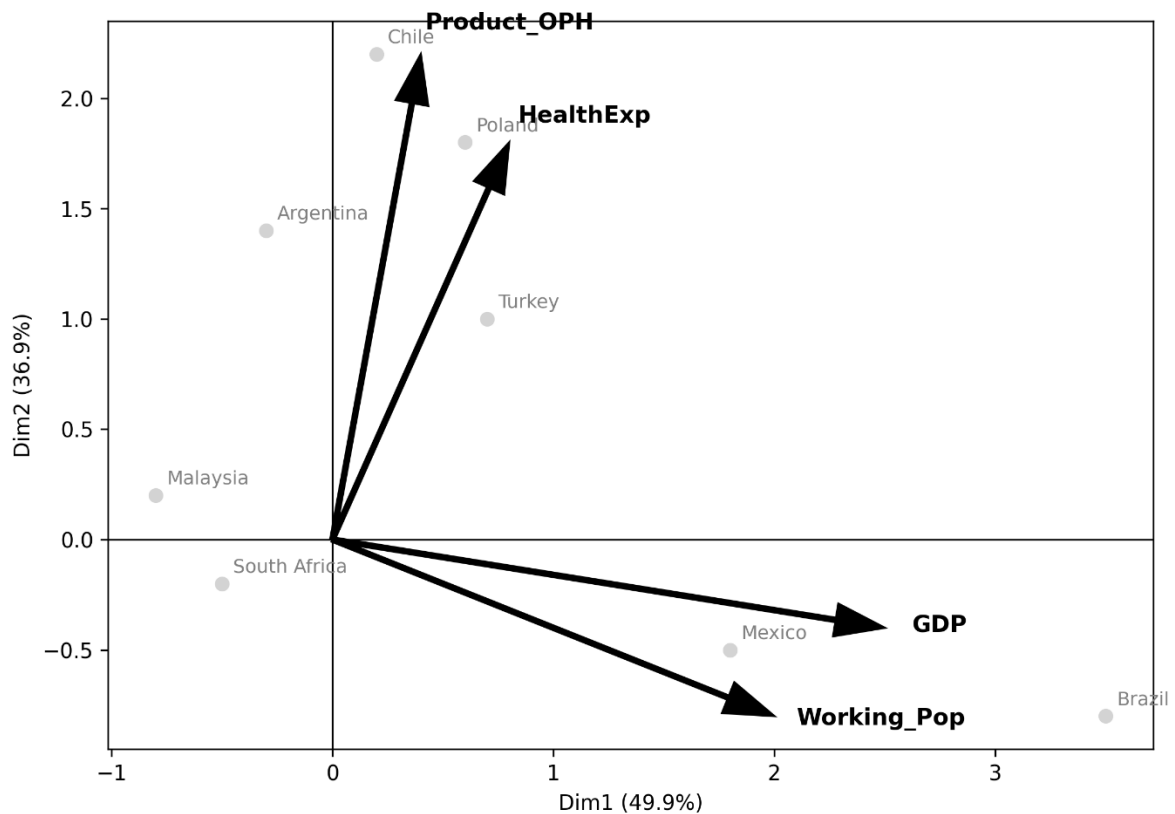
Nombre	Índice 1 Económico	Índice 2 Balance alimentario	Índice 3 Factores de riesgo			
			CAD	Resp	DM	Neo
Dimensiones seleccionadas	2	4	2	2	2	2
Proporción variada acumulada	86.84%	78.06%	90%	94%	91%	88%

Fuente: Elaboración propia

Variables económicas

Figura 12

Distribución de países y carga de variables económicas en las dos primeras dimensiones



Fuente: Elaboración propia

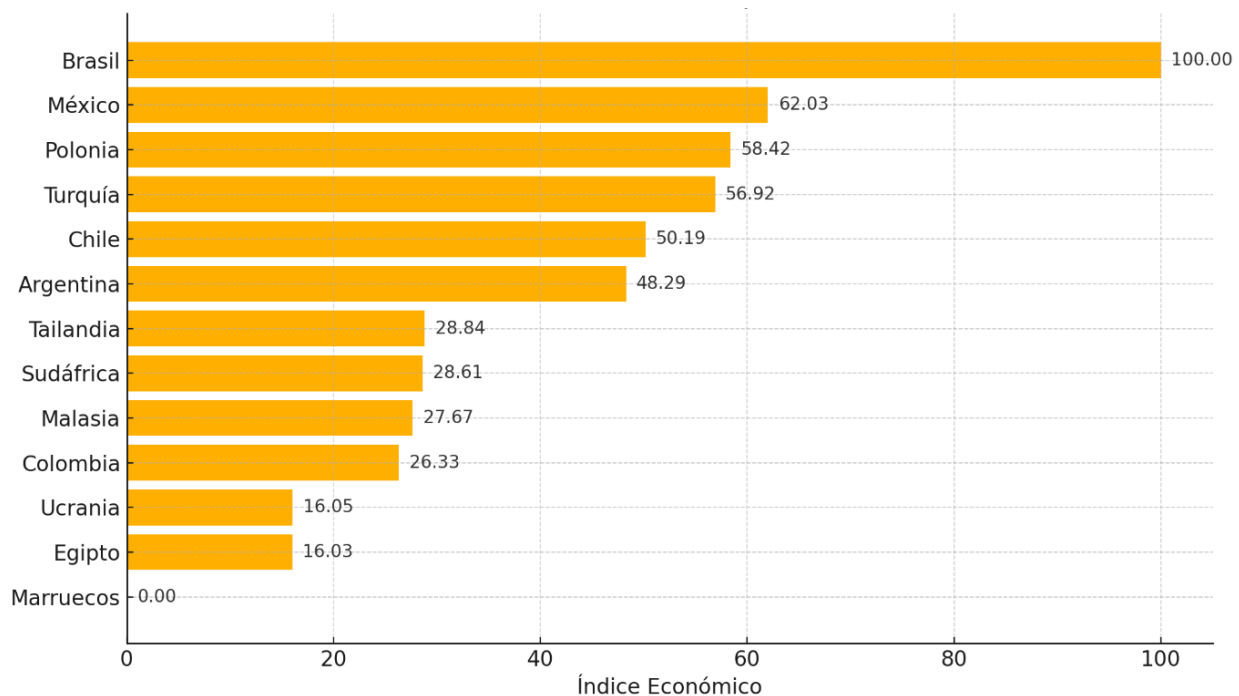
Se observa en la Figura 12 que consolida el 86.84% de la varianza total de las 2 dimensiones, que el comportamiento de los elementos económicos considerados se agrupa en dos sectores, en el primero, PIB y toda la población activa, se comportan de manera similar. Por otro lado, se puede encontrar en un comportamiento similar la inversión en salud (HealthExp) junto con las horas trabajadas (Product_OPH) localizadas en el mismo sector; esta información indica cómo se correlacionan las variables y como hallazgo relevante, la influencia de las horas trabajadas y las inversiones en salud se encuentran en el mismo cuadrante.

Por otro lado, se observa que los países de Malasia, Turquía, Tailandia, Colombia, Egipto, Ucrania, Sudáfrica y Marruecos tienen una similitud con respecto a los indicadores económicos seleccionados. Polonia, Chile, Turquía y Argentina tienen un comportamiento similar. México y Brasil se encuentran en el mismo cuadrante con un desempeño cercano con más tendencia hacia la dimensión 2.

En la Figura 13 se puede observar el índice económico generado por PCA que permite sintetizar las variables económicas seleccionadas vinculadas a las ENT.

Figura 13

Índice de factores económicos



Fuente: Elaboración propia

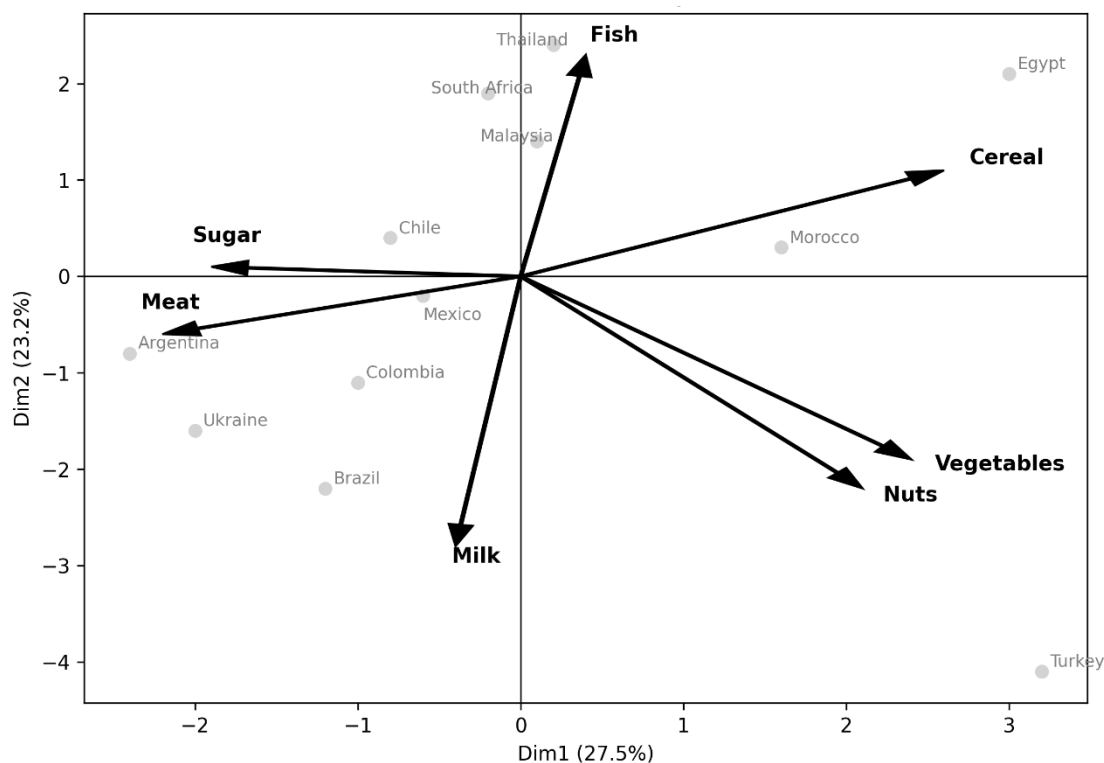
Este índice nos permite entender qué países tienen un mayor valor en variables económicas y cuál de ellos tiene menos inversión. Adicionalmente, se observa que Brasil está mejor en términos económicos que México según las variables del índice; en la lista inferior, Ucrania, Egipto y Marruecos se mantienen con menos del 20% del índice.

Consumo alimentario

El consumo de alimentos de los países seleccionados se evaluó en función de las calorías. El objetivo de este análisis es encontrar patrones de agrupación y cómo es el comportamiento de los diferentes países con respecto al consumo alimentario³. En primera instancia, se observa en la Figura 14 el gráfico *Biplot* que considera la información del 47% de la variabilidad de las primeras 2 dimensiones; las variables se agrupan por componentes alimentarios y se agrupan por países.

Figura 14

Distribución de países y carga de variables alimentarias en las dos primeras dimensiones



Fuente: Elaboración Propia

³ Véase Diagrama de Inferencia 2 en Anexo 2, Tabla 28.

Desde la perspectiva de las variables alimentarias, se identifica una heterogeneidad clara en los patrones de consumo, con una agrupación coherente de alimentos por similitud nutricional y epidemiológica. En particular, se observa que el azúcar, la carne y la leche se concentran en un mismo cuadrante, lo que refleja un patrón dietario asociado en la literatura con una mayor prevalencia de enfermedades no transmisibles. En contraste, frutas, vegetales, legumbres y nueces se agrupan en un cuadrante diferenciado, representando un patrón de consumo predominantemente basado en plantas, mientras que cereales, cultivos oleaginosos y pescado conforman un tercer eje dietario asociado a fuentes de energía y de grasas saludables.

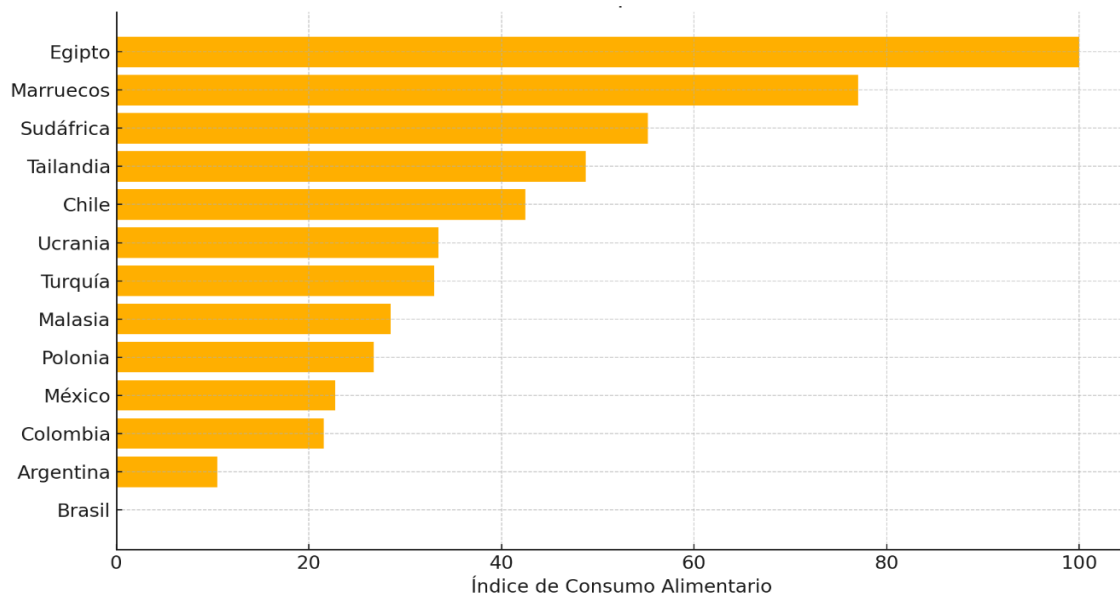
La interpretación del *biplot* se apoya, además, en las variables con mayor peso vectorial, las cuales concentran la mayor capacidad explicativa del análisis. En la primera dimensión (Dim1, 27.5%) destacan principalmente el consumo de cereales, vegetales y nueces, mientras que la segunda dimensión (Dim2, 23.2%) está dominada por el consumo de pescado, carne, leche y azúcar. Estas variables sintetizan los principales patrones dietarios estructurales observados en los países analizados y facilitan la comprensión de las diferencias en la transición nutricional.

En cuanto al posicionamiento de los países, se observa que México, Chile, Polonia, Argentina, Colombia, Brasil y Ucrania se concentran en la región asociada al primer grupo de alimentos, caracterizado por un mayor consumo relativo de azúcar y carne. Por su parte, Sudáfrica, Malasia, Tailandia, Egipto y Marruecos se posicionan en la región vinculada al tercer grupo, lo que es consistente con una tendencia en la dieta principalmente basada en alimentos de origen vegetal. Finalmente, Turquía presenta un comportamiento diferenciado, ubicándose de manera aislada y con mayor influencia del grupo asociado a cereales y pescado.

Gracias al análisis de los componentes principales, se generó la Figura 15 con el índice de alimentos de variables seleccionadas de los diversos países seleccionados; este índice nos permite entender qué países están consumiendo más calorías que otros. Se observa que Egipto es el que consume más calorías de todos los productos en su conjunto y Brasil el que menos calorías consume de todos los productos. Se puede observar que diversos países ubicados en América Latina tienen un menor balance alimentario junto con Polonia.

Figura 15

Índice de consumo alimentario



Fuente: Elaboración propia

Para entender mejor el comportamiento del consumo de los países, se seccionó el consumo de alimentos considerando los grupos encontrados en la Figura 16. Azúcar, huevos, carne, grasas animales y leche se agruparon en un segmento denominado Sección 1 y el resto de las variables en otro segmento denominado Sección 2. Con la intención de comprender mejor el consumo de alimentos, pero con base en aquellos alimentos que tienen un mayor aporte calórico y que se corrobora con la Figura 16, podrían haber tenido beneficios en la dieta y que se encuentran en el lado positivo de la dimensión 1. Así es como se observan los resultados en la Figura 16.

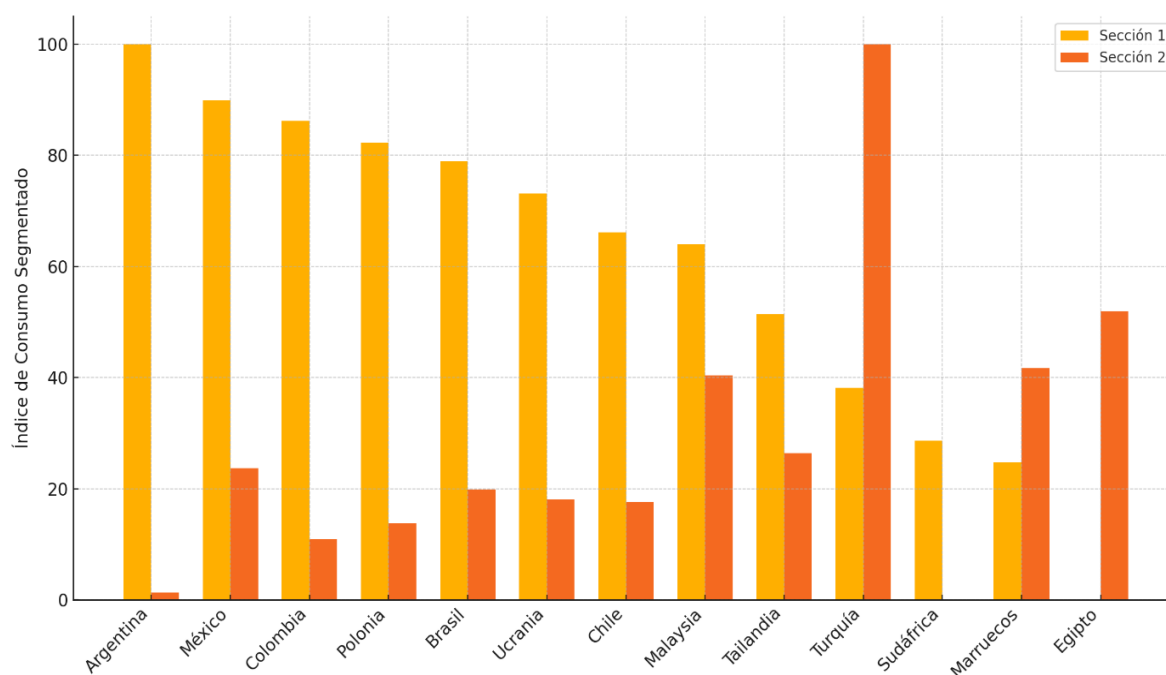
El análisis del índice de consumo segmentado permite visualizar con mayor claridad los patrones diferenciados de alimentación entre los países analizados. Se dividió el consumo en dos secciones representativas construidas a partir de la disponibilidad relativa de grupos de alimentos en cada país. La Sección 1 agrupa alimentos generalmente considerados como más básicos o tradicionales dentro de la dieta de los países latinoamericanos, mientras que la Sección 2 recoge una combinación de alimentos procesados, industriales o con mayor valor agregado.

En la Figura 16 puede observarse que los países de América Latina (Argentina, Brasil, Chile, Colombia y México), junto con Polonia, presentan un consumo más elevado en la Sección 1. Este patrón podría estar asociado con una dieta más basada en productos tradicionales y de

consumo interno, lo que sugiere una estructura de consumo menos influenciada por la globalización alimentaria. No obstante, pese a liderar en la Sección 1, algunos de estos países, como Argentina y Brasil, se encuentran en posiciones bajas dentro del índice total de consumo alimentario, lo que sugiere que su consumo no es equilibrado o que la cantidad relativa en la Sección 2 es muy reducida.

Figura 16

Índice de consumo segmentado



Fuente: Elaboración Propia

Por otro lado, países como Egipto y Marruecos presentan un comportamiento contrastante. Egipto, por ejemplo, encabeza el índice de consumo total según la Figura 16, pero su consumo se concentra notablemente en la Sección 2, lo que sugiere una dieta más diversificada o distinta a la latinoamericana, posiblemente relacionada con patrones culturales, disponibilidad regional o subsidios alimentarios. Marruecos, aunque con una participación más equilibrada entre ambas secciones, muestra también una preferencia ligeramente mayor por la Sección 2.

En conjunto, este análisis evidencia la heterogeneidad en los patrones de consumo alimentario entre países y cómo distintas estructuras de dieta pueden influir en los resultados de los índices compuestos de alimentación.

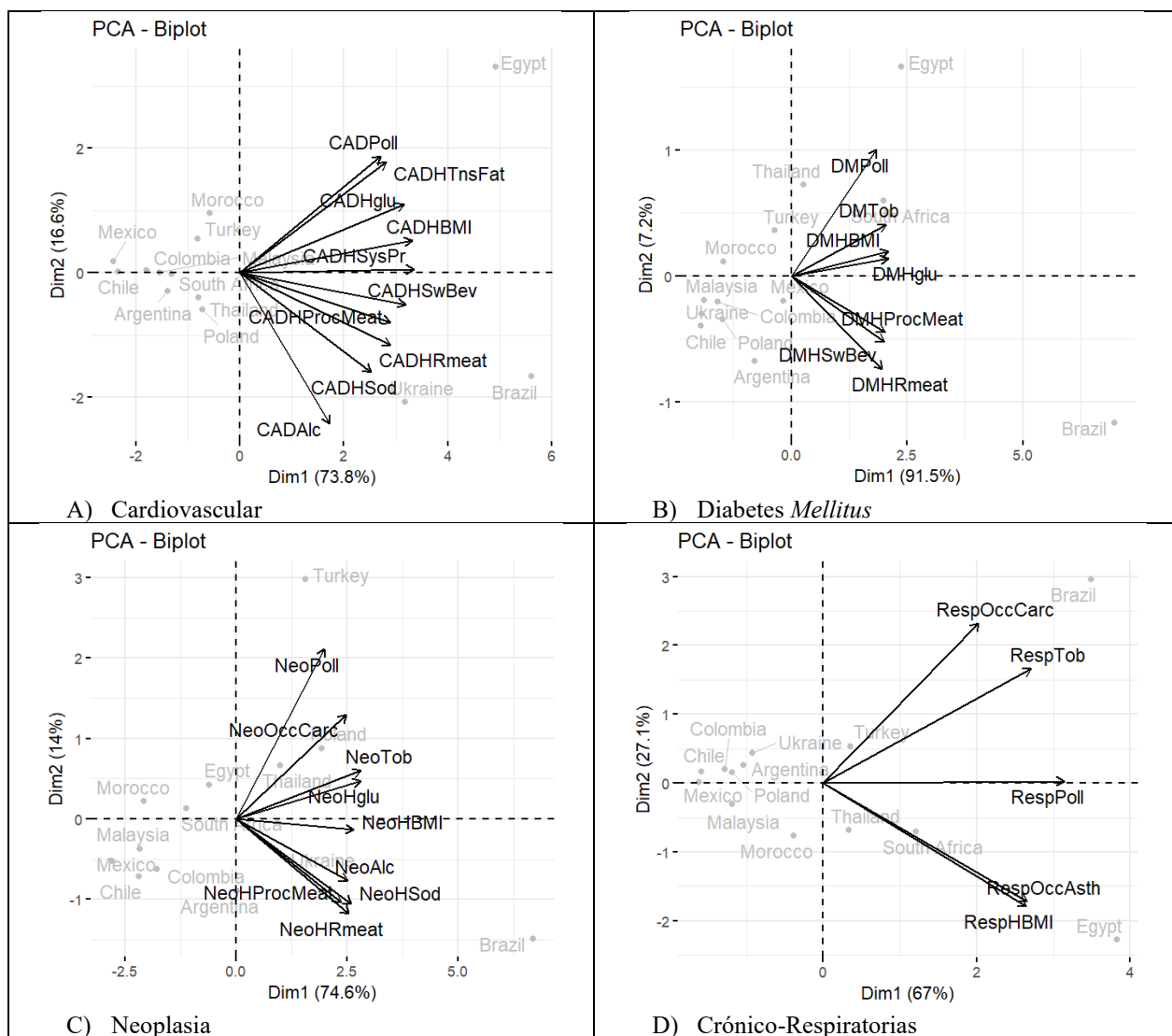
Factores de riesgo

Factores-H

En primer lugar, aquellas variables que aparecen como de alta influencia que afectan a la mortalidad y que, por su característica, aparecen con mayor frecuencia en las estadísticas de las bases de datos se denominan Factor H⁴.

Figura 17

Distribución de países y carga de variables entre los factores H en las dos primeras dimensiones



Fuente: Elaboración Propia

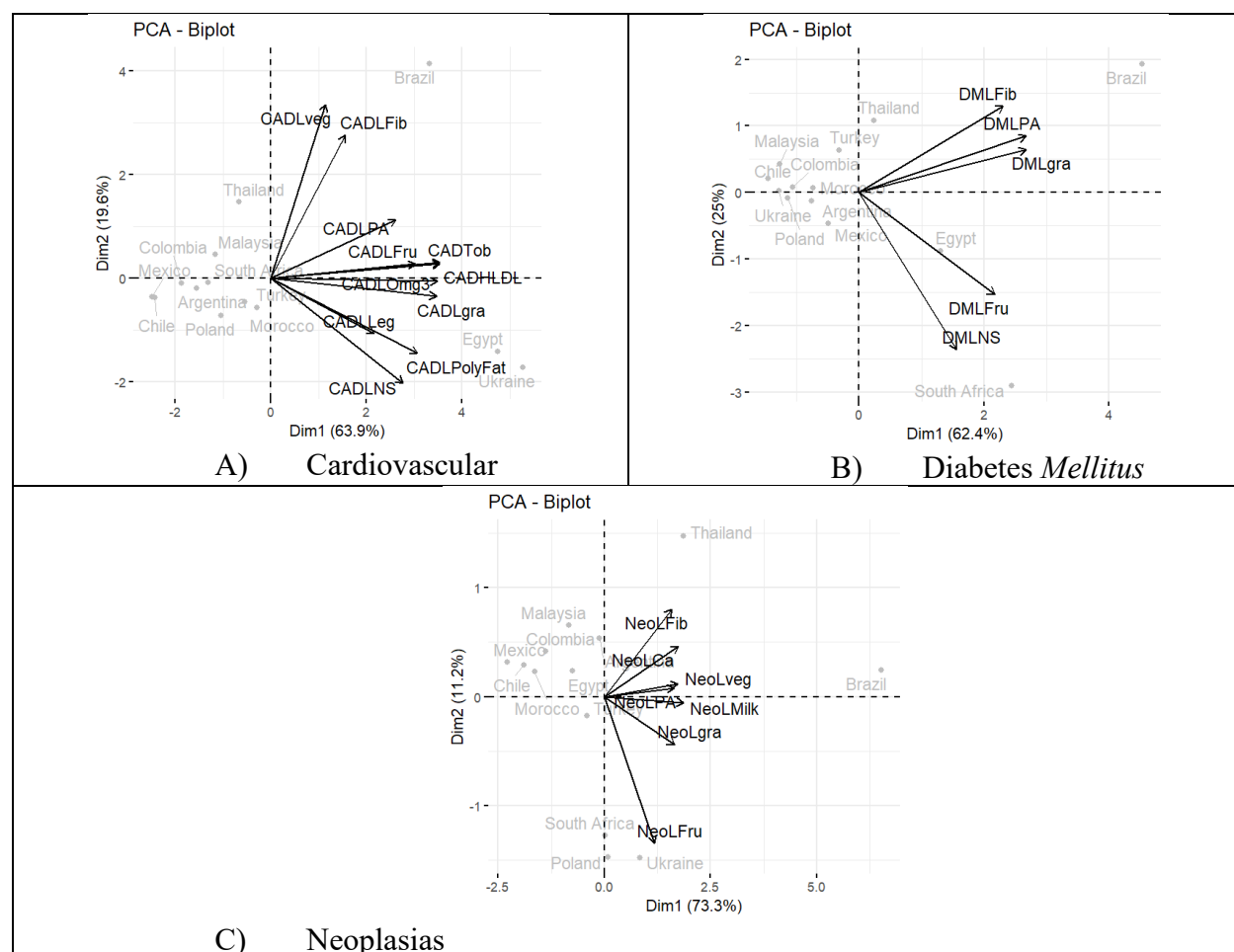
⁴ Véase Diagrama de Inferencia 3 en Anexo 2, Tabla 29.

La Figura 17 muestra cómo se agrupan las causas de muerte para cada ENT, así como la agrupación de países por su similitud obtenida por sus componentes principales.

Se observa a partir del análisis de los factores H que influyen en la mortalidad que existe una alta concentración del comportamiento de diversas variables en la mayoría de los países seleccionados, excepto en Brasil, que se enmarca en otro comportamiento. En el caso de las enfermedades cardiovasculares, se observó que el IMC se comporta de manera similar con los altos niveles de glucosa en grasas trans o hipertensión, que son indicadores de salud.

Figura 18

Distribución de países y carga de variables entre los factores L en las dos primeras dimensiones



Fuente: Elaboración propia

En el caso de la diabetes *mellitus* y enfermedades neoplásicas, el estilo de salud se comporta de manera similar al IMC, niveles altos de glucosa y, en este caso, el consumo de tabaco y la

influencia de la contaminación. Se observa que, en enfermedades respiratorias crónicas, el tabaco y la exposición a carcinógenos se agrupan de manera similar. En todos los casos se agrupan alimentos como procesados, bebidas dulces o consumo de alcohol.

A partir de los resultados obtenidos, se realizó una evaluación de las cuatro ENT, en primera instancia, de aquellas variables que aparecen como un bajo porcentaje de su uso que afectan la mortalidad y que, por su característica, aparecen con menor frecuencia en las estadísticas de las bases de datos, se denominaron Factores L mostrados en la Figura 18.

Factores-L

En cuanto a las variables etiquetadas como Factor L, al observar cómo se posicionan los países con respecto a la diversidad de variables económicas, alimentarias y de factores de riesgo, se observa un comportamiento similar en casi todos los países, excepto una vez más en Brasil. En esta evaluación, el país de Egipto también se desempeña por separado del grupo principal de países en desarrollo. La selección de países se puede considerar como óptima para el análisis incluso cuando hay países aislados que podrían modificar la información.

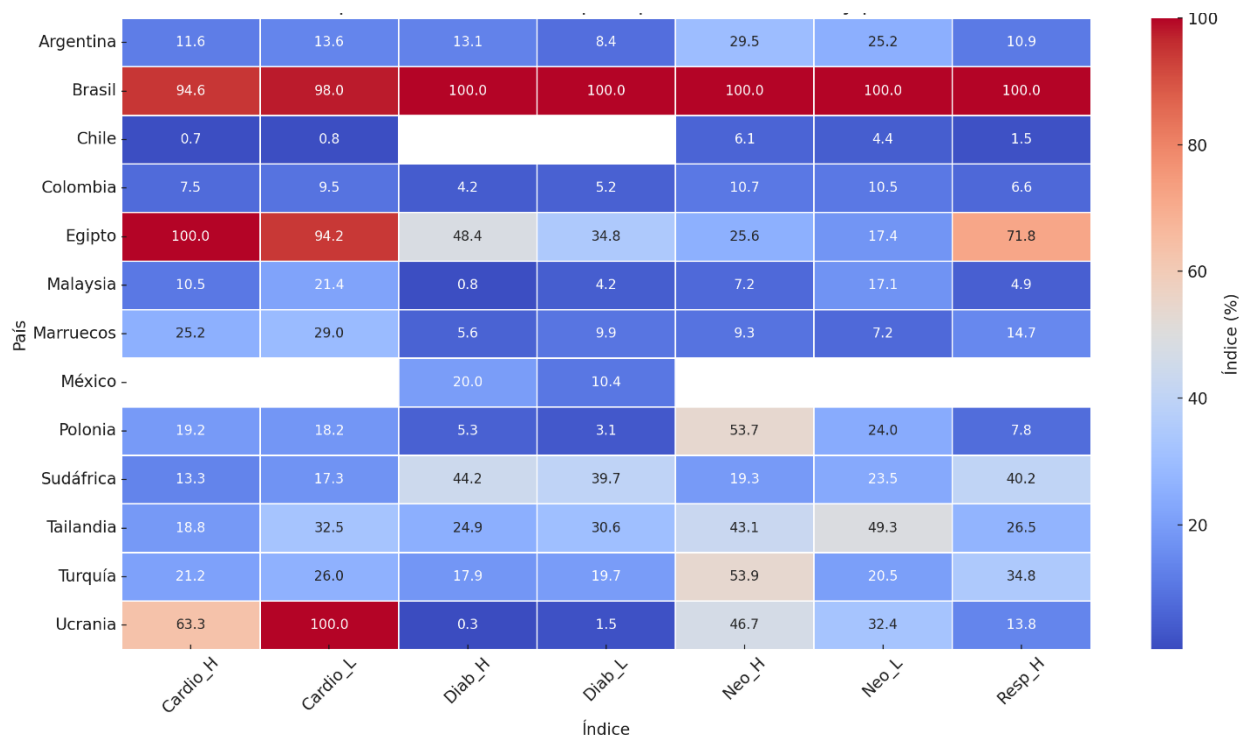
En la Figura 19 se presentan los índices resumidos de carga por ENT, diferenciando entre factores tipo H (alta presencia de riesgo) y tipo L (baja presencia de factores protectores), agrupados por tipo de ENT: enfermedades cardiovasculares, diabetes *mellitus*, neoplasias y enfermedades respiratorias crónicas. Los países están ordenados alfabéticamente para facilitar la comparación general y los valores representan puntuaciones relativas estandarizadas entre 0 y 100.

Los índices H reflejan el nivel de afectación negativa asociada a variables como alta presión arterial, hiperglucemia, tabaquismo o sedentarismo, es decir, una mayor carga de riesgo para cada grupo de enfermedades. En contraste, los índices L capturan la ausencia o insuficiencia de factores protectores como actividad física regular, consumo de grasas saludables (*omega 3*) o acceso a diagnóstico temprano, lo que puede indicar vulnerabilidad en políticas preventivas.

Los resultados muestran claramente que Brasil lidera de manera consistente en la mayoría de los índices H y L posicionándose como el país con mayor carga relativa en las ENT analizadas, lo que puede reflejar tanto una alta prevalencia de factores de riesgo como deficiencias en medidas preventivas. En el caso de Egipto, también se observa una carga significativa, especialmente en enfermedades respiratorias.

Figura 19

Mapa de calor de los Índices de factores de riesgo por enfermedad y país



Fuente: Elaboración propia

Por otro lado, México presenta una situación heterogénea. Mientras muestra un desempeño favorable en enfermedades cardiovasculares, neoplasias y respiratorias (al no figurar entre los países con alta carga H ni baja protección L), su posición en diabetes *mellitus* es intermedia, lo que refleja una vulnerabilidad específica en esta área. Este contraste refuerza la necesidad de enfoques diferenciados por tipo de enfermedad y grupo de riesgo.

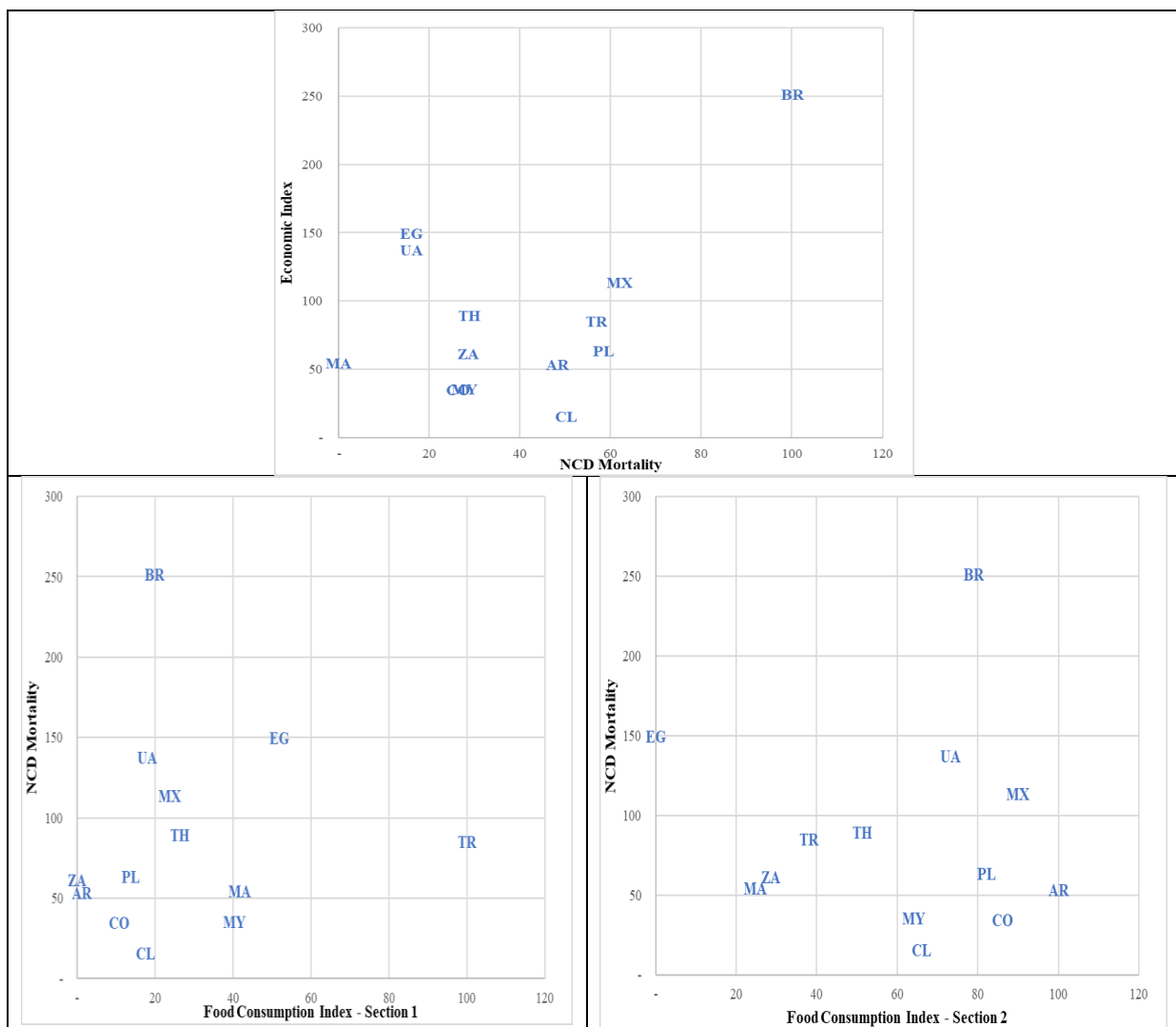
Análisis de conglomerados

Para identificar patrones comunes entre los países en desarrollo analizados, se llevó a cabo un análisis de conglomerados con base en los índices construidos a partir de PCA. Previamente se exploró la relación directa entre cada uno de los índices económicos, de consumo de alimentos y de factores de riesgo con los niveles de mortalidad por las ENT, a fin de identificar posibles asociaciones entre dimensiones estructurales y resultados sanitarios. Esta etapa permitió visualizar cómo varían las tasas de mortalidad en función de las características socioeconómicas, alimentarias y sanitarias sintetizadas por los índices (ver Figura 20).

Posteriormente, se aplicó un algoritmo de clasificación no supervisada agrupando a los países mediante k-medias sobre el espacio definido por los índices seleccionados. El objetivo fue detectar agrupamientos naturales entre países con perfiles similares, facilitando así la caracterización de clústeres con comportamientos homogéneos en cuanto a sus condiciones estructurales y riesgo asociado a ENT. Esta segmentación no solo permite comparar el posicionamiento relativo de cada país, sino que también ofrece una herramienta útil para diseñar intervenciones diferenciadas a nivel regional.

Figura 20

Índices vs mortalidad por ENT



Fuente: Elaboración propia

En este contexto, se contrastaron las posiciones de los países según el índice económico (Figura 13) y los dos índices de consumo alimentario (Figura 15) con los niveles de mortalidad por ENT reportados, como se sintetiza en la Figura 20.

Derivado del comportamiento económico, se observa que países como Ucrania o Egipto, que están por debajo del índice económico, presentan una mayor mortalidad. Otros países muestran que, al estar en valores de índice intermedios como es el caso de Chile, Argentina, Polonia o Turquía, el indicador de mortalidad es menor que el de otros países. En el caso de México o Brasil, aunque muestran un índice económico más alto, los resultados no han sido favorables con respecto a la mortalidad por ENT.

Por otro lado, en cuanto al consumo de alimentos se observa que países como México, Ucrania o Brasil que muestran un mayor número dentro del índice de consumo de alimentos que se consideraron dentro de la sección 1 que son azúcar, huevos, carne, grasas animales y leche, presentan una mortalidad moderada comparado con países como Malasia, Sudáfrica o Turquía que están por debajo del índice muestran una menor mortalidad en las ENT. Sin embargo, se observa un comportamiento diferente en países como Argentina, que, siendo el número 1 del índice, no muestra mortalidad significativa, o Egipto, que, aunque se encuentra en la parte inferior del índice, muestra una modalidad moderadamente alta. Continuando con el consumo de alimentos en la sección 2, Turquía, que se encuentra en el número 1 en el índice de mayor consumo de estos alimentos, muestra una baja tasa de mortalidad; en contraste, países como Brasil o Ucrania, que estaban por debajo de este índice, tienen una mayor mortalidad.

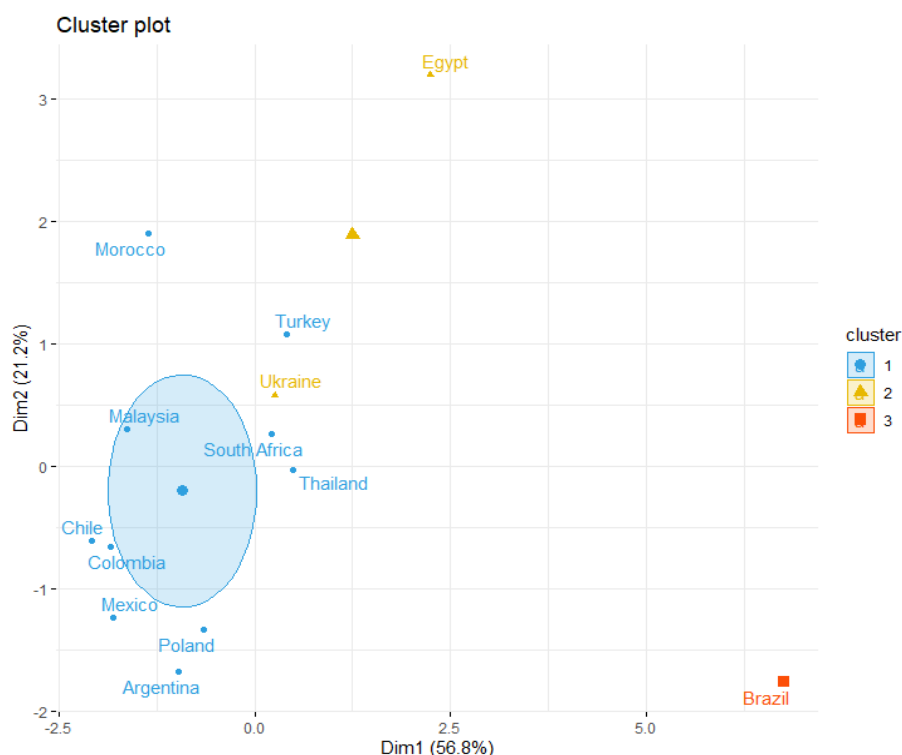
El resultado de la agrupación de los índices mediante k-medias se puede observar en la Figura 21, donde se consideraron 3 grupos diferentes, en los que la mayoría de los países se comportan de manera similar. Los países de Egipto y Ucrania muestran una agrupación diferente y Brasil es el único país considerado único; esto puede explicarse por el comportamiento de los índices.

En el caso de Ucrania y Egipto, índices como el consumo de alimentos y diversos indicadores de salud son los que se observan en los niveles más bajos, a diferencia de los otros países que se mantienen en el promedio o por debajo. Además, Brasil es el país que se encuentra en los primeros lugares en la mayoría de los índices, por lo que al realizar el método de cluster, identifica los pesos de todos los indicadores y es visto como una entidad independiente.

Este análisis nos ayuda a entender cómo México se posiciona frente a otros países en desarrollo y puede ser de utilidad al momento de gestar políticas públicas o para la implementación de buenas prácticas tomando en cuenta similitudes en el comportamiento económico, alimentario y en la comparativa de riesgos respecto a las ENT enfocadas en la PEA.

Figura 21

Agrupación de índices K-medias



Fuente: Elaboración propia

Efectos de los factores relacionados con las ENT de países en desarrollo durante la edad laboral

Como parte del análisis global, se busca verificar las variables más significativas que impactan de forma negativa a los trabajadores por la carga relevante que significan las ENT. El conjunto de datos emanado de las distintas bases de datos mundiales se estudia y se aplican diferentes métodos para seleccionar subconjuntos de atributos, como regresiones penalizadas (regresión de *Lasso*, regresión de *Ridge*)⁵.

Aunque se implementaron diversos procedimientos econométricos para evaluar la estructura de los datos de panel, incluyendo pruebas de normalidad, homocedasticidad,

⁵ Véase Diagrama de Inferencia 4 en Anexo 2, Tabla 30.

autocorrelación, endogeneidad y especificación de modelo. Los resultados obtenidos a partir de estos modelos no serán presentados en la sección de resultados, dado que el enfoque metodológico definitivo se basó exclusivamente en la aplicación de técnicas de regularización como *Ridge*, *Lasso* y redes elásticas⁶. Estas últimas permitieron una selección más robusta y automatizada de variables predictoras, adecuándose mejor a la naturaleza de alta dimensionalidad y colinealidad de los datos analizados. No obstante, las pruebas estadísticas mencionadas anteriormente fueron útiles en la validación inicial del enfoque empírico y sirvieron como insumo exploratorio para delimitar la estrategia metodológica final.

Se realizó un análisis de las cuatro ENT bajo varios métodos de análisis incluyendo regresiones penalizadas por el método de *Ridge*, *Lasso* y redes elásticas. El modelo *Ridge* reduce la varianza al penalizar la suma de cuadrados de los coeficientes, *Lasso* permite la selección automática de variables al forzar a cero algunos coeficientes y la Red Elástica combina ambos enfoques. El uso de estas técnicas facilitó la reducción del número de predictores sin comprometer la precisión del modelo y permitió establecer patrones de asociación robustos entre el contexto socioeconómico, los comportamientos individuales y la carga de enfermedad cardiovascular observada en cada país. A continuación, se presentan los resultados de cada uno de los análisis realizados y sus hallazgos.

Cardiovascular

Con el objetivo de identificar los factores con mayor asociación a la prevalencia de enfermedades cardiovasculares en países en desarrollo, se aplicaron modelos de regresión penalizada. Estas técnicas permitieron controlar la multicolinealidad y seleccionar las variables más relevantes en contextos con alta dimensionalidad de datos, sobre la ecuación (27):

$$Cardio_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \varepsilon_i \quad (27)$$

Donde $Cardio_i$ representa la mortalidad por enfermedades cardiovasculares para el país i , X_{ji} corresponde al conjunto de variables explicativas correspondientes a ese país. Entre estas variables se incluyeron dimensiones económicas (porcentaje del gasto en salud, inversión pública

⁶ Puede observarse los resultados completos en Arturo & Rodríguez-Aguilar, 2024.

en prevención), factores alimentarios (consumo de sodio, frutas, verduras, aceites saludables), hábitos de vida (nivel de actividad física, patrones de sueño) y factores de riesgo (tabaquismo, hipertensión, índice de masa corporal promedio) de riesgo, así como subcomponentes asociados a estilos de vida y condiciones estructurales.

Con el propósito de validar la solidez de los modelos utilizados y garantizar la confiabilidad de los resultados obtenidos, se evaluó la precisión de las predicciones mediante métricas de desempeño estadístico. Se comparó el rendimiento de los modelos aplicados mediante el cálculo del error cuadrático medio (MSE) como se presenta en la Tabla 10.

Tabla 10

MSE en enfermedad cardiovascular

Modelo	MSE
<i>Ridge</i>	17,207,063
<i>Lasso</i>	16,154,102
Redes elásticas con el mejor λ	15,701,764

Fuente: Elaboración propia

La Tabla 11 presenta los resultados de los modelos *Lasso*, redes elásticas con ponderaciones proporcionales entre *Ridge* y *Lasso*, así como redes elásticas optimizadas con el valor de *lambda* que minimiza el error. Esta comparación se enfocó exclusivamente en la mortalidad por enfermedades cardiovasculares, dado que los valores de error cuadrático medio (MSE) entre los modelos fueron muy similares. La selección de estos enfoques busca resaltar las diferencias en la selección de variables y el ajuste del modelo más allá del desempeño predictivo marginal.

Tabla 11

Regresiones penalizadas – enfermedades cardiovasculares

	<i>Lasso</i>	<i>Ridge</i>	Redes elásticas con el mejor λ
PPHW	0.02239	0.0214	0.0211
Carne procesada	6.7729	1.665	1.0132
Bebidas azucaradas	0.0869		5.6644
Bajo consumo de granos	0.1666	0.0021	0.3497

Colesterol LDL	0.0018	0.1731	
Alta presión sistólica	0.0039	0.0787	
Tabaquismo	0.0267	0.0173	
Dieta baja en ácidos grasos poliinsaturados			0.7196
Bajo consumo de pescado y <i>omega 3</i>			0.3933
λ	9.5		39
α		1	0.45

Fuente: Elaboración propia

Cabe destacar que las bebidas azucaradas, la carne procesada y un mayor número de horas trabajadas son las que pueden causar un mayor impacto en la mortalidad cardiovascular dentro de la etapa de trabajo; por otro lado, se observa que el consumo de cultivos oleaginosos, leguminosas, semillas y pescado podría reducir este efecto.

Enfermedades Crónico-respiratorias

Para identificar los factores asociados a las enfermedades respiratorias crónicas en países en desarrollo, se estimaron modelos de regresión penalizada mediante *Ridge*, *Lasso* y redes elásticas. La variable dependiente fue la mortalidad de enfermedades respiratorias crónicas por país. La forma funcional del modelo se representa en la ecuación (28):

$$Resp_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \varepsilon_i \quad (28)$$

Donde $Resp_e$ representa el índice de carga por enfermedades crónico-respiratorias para el país i , J_i corresponde al conjunto de variables explicativas correspondientes a ese país. Entre estas variables se incluyeron dimensiones económicas (porcentaje del gasto en salud, inversión pública en prevención), factores alimentarios (consumo de sodio, frutas, verduras, aceites saludables), hábitos de vida (nivel de actividad física, patrones de sueño) y factores de riesgo (tabaquismo, hipertensión, índice de masa corporal promedio) de riesgo, así como subcomponentes asociados a estilos de vida y condiciones estructurales. Se evaluó la integridad y exactitud de los datos obtenidos y evaluados comparando el rendimiento de los modelos utilizando las medidas de rendimiento basadas en los valores de precisión con el error cuadrático medio de las predicciones de acuerdo con la Tabla 12.

Tabla 12

MSE en enfermedades crónico-respiratorias

Modelo	MSE
<i>Ridge</i>	1,756,756
<i>Lasso</i>	1,788,221
Redes elásticas con el mejor λ	1,629,811

Fuente: Elaboración propia

Se observa que las redes elásticas nuevamente tienen un mejor rendimiento en el modelo, aunque la técnica de *Lasso* es muy efectiva, esto se deriva del hecho de que el cálculo de la mejor α está muy cerca del valor utilizado para la técnica de *Lasso*; sin embargo, el modelo sigue siendo mayormente efectivo cuando se enfoca en el rendimiento de los mejores λ . La Tabla 13 muestra los resultados obtenidos en los modelos con respecto a la mortalidad por enfermedades crónico-respiratorias.

Tabla 13

Regresiones penalizadas – enfermedades crónico-respiratorias

	<i>Lasso</i>	<i>Ridge</i>	Redes elásticas con el mejor λ
DALYS	0.01639	0.0168	0.0144
Contaminación	0.4228	0.4381	0.4831
Asma ocupacional		0.3658	1.8981
Carcinógenos ocupacionales	1.4214		0.0088
Tabaquismo	0.0017	0.001	0.002
λ	7.5		39
α		1	0.09

Fuente: Elaboración propia

Cabe señalar que un elemento esencial son los carcinógenos profesionales consistentes en sustancias químicas, agentes físicos o biológicos y procesos de trabajo que, por su naturaleza y niveles de exposición, plantean un riesgo considerable para el trabajador.

Los elementos mostrados en este análisis concuerdan con la literatura, ya que el efecto del tabaquismo ha sido evaluado anteriormente; sin embargo, a nivel estadístico no se había encontrado evidencia. Adicionalmente, los efectos de la contaminación y los carcinógenos

ocupacionales tienen un efecto mayor; más aún, el asma ocupacional afecta como elemento primario a la enfermedad crónica respiratoria.

Enfermedades Neoplásicas

El análisis de los determinantes de enfermedades oncológicas se abordó mediante la estimación de modelos de regresión penalizados. La variable dependiente fue la prevalencia ajustada de neoplasias malignas en cada país de la muestra. La forma funcional del modelo se observa en la ecuación (29):

$$Neo_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \varepsilon_i \quad (29)$$

Donde Neo_i representa el índice de carga por enfermedades neoplásicas para el país i , X_{ji} corresponde al conjunto de variables explicativas correspondientes a ese país. Entre estas variables se incluyeron dimensiones económicas (porcentaje del gasto en salud, inversión pública en prevención), factores alimentarios (consumo de sodio, frutas, verduras, aceites saludables), hábitos de vida (nivel de actividad física, patrones de sueño) y factores de riesgo (tabaquismo, hipertensión, índice de masa corporal promedio) de riesgo, así como subcomponentes asociados a estilos de vida y condiciones estructurales.

Fue evaluada la exactitud de los datos obtenidos y evaluados comparando el rendimiento de los modelos utilizando las medidas de rendimiento basadas en los valores de precisión con el error cuadrático medio de las predicciones obteniendo los valores de la Tabla 14.

Tabla 14

MSE en enfermedades neoplásicas

Modelo	MSE
<i>Ridge</i>	2,536,641
<i>Lasso</i>	2,623,714
Redes elásticas con el mejor λ	2,548,993

Fuente: Elaboración propia

Se observa que en esta evaluación es el modelo de *Ridge* seguido muy de cerca por el análisis de redes elásticas con los mejores hiperparámetros como los mejores modelos de análisis para la enfermedad neoplásica.

La Tabla 15 muestra los resultados obtenidos en los modelos con respecto a la mortalidad por neoplasias.

Tabla 15

Regresiones penalizadas – Enfermedades neoplásicas

	<i>Lasso</i>	<i>Rigde</i>	Redes elásticas con el mejor λ
DALYS	0.0229	0.0231	0.0219
Consumo de alcohol	0.1185	0.0022	
Alto consumo de sodio	0.0019		
Bajo consumo de leche	0.0002		3.9397
Baja actividad física	3.1483	0.0183	3.9145
Niveles altos de glucosa	1.1579	0.0459	1.2035
Tabaquismo	0.0054	0.0262	2.0396
λ	8		39
α		1	0.51

Fuente: Elaboración propia

Cabe destacar que el alto consumo de glucosa, el consumo de alcohol, el tabaquismo y la baja actividad física pueden causar un mayor impacto en la mortalidad por neoplasias según las estadísticas dentro de la etapa de trabajo; por otro lado, se observa que el consumo de leche puede reducir este efecto.

Diabetes Mellitus

La diabetes *mellitus* fue evaluada mediante técnicas de regresión penalizadas con el objetivo de determinar los principales factores asociados a su prevalencia en países en desarrollo. La variable dependiente fue la tasa ajustada de diabetes *mellitus* tipo 2. La forma funcional adoptada se observa en la ecuación (30):

$$DB_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \varepsilon_i \quad (30)$$

Donde DB_i representa el índice de carga por la diabetes *mellitus* para el país i , X_{ji} corresponde al conjunto de variables explicativas correspondientes a ese país.

Entre estas variables se incluyeron dimensiones económicas (porcentaje del gasto en salud, inversión pública en prevención), factores alimentarios (consumo de sodio, frutas, verduras, aceites saludables), hábitos de vida (nivel de actividad física, patrones de sueño) y factores de riesgo (tabaquismo, hipertensión, índice de masa corporal promedio) de riesgo, así como subcomponentes asociados a estilos de vida y condiciones estructurales.

Fue evaluada la integridad y exactitud de los datos obtenidos y evaluados comparando el rendimiento de los modelos utilizando las medidas de rendimiento basadas en los valores de precisión con el error cuadrático medio de las predicciones, manteniendo los siguientes valores en la Tabla 16.

Tabla 16

MSE en Diabetes Mellitus

Modelo	MSE
<i>Ridge</i>	71,734.30
<i>Lasso</i>	65,247.34
Redes elásticas con el mejor λ	71,723.98

Fuente: Elaboración propia

Se observa que esta vez es el modelo *Lasso* el que tiene un mejor rendimiento, aunque cercano a la red elástica con mejor λ . La Tabla 17 muestra los resultados obtenidos en los modelos de *Lasso*, redes elásticas con valores proporcionales de *Ridge* y *Lasso* y redes elásticas con la mejor *lambda* con respecto a la mortalidad derivada de la diabetes *mellitus*.

Tabla 17

Regresiones penalizadas – diabetes Mellitus

	<i>Lasso</i>	<i>Ridge</i>	Redes elásticas con el mejor λ
PPHW	0.0002		0.002
Contaminación	0.0265		
Bajo consumo de fibra	0.0477		
Bajo consumo de grasas	0.316	0.0021	0.1854
Glucosa elevada	0.8233	0.915	0.84
Tabaquismo	0.0116	0.0174	0.2093

λ	6		39
α		1	0.21

Fuente: Elaboración propia

En este resultado se observa que este modelo muestra un mayor impacto derivado de la dieta alta en azúcar y el tabaquismo, aunque también aparece el alto número de horas trabajadas en los empleados. Por otro lado, se observa que es probable que un mayor consumo de grasas pudiera minimizar el efecto de la diabetes *mellitus*.

4.2 Hábitos de salud en la PEA y sus efectos en las ENT en México

A partir de la caracterización internacional y con el interés de profundizar en el contexto nacional, el análisis se centró en la información proporcionada por la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2022 con el propósito de examinar la forma en que los hábitos de salud de la PEA en México influyen en la prevalencia de ENT. Esta base de datos ofrece una amplia gama de variables relevantes para este estudio, incluyendo condiciones de salud auto percibidas, diagnósticos previos, antecedentes familiares, consumo alimentario, actividad física, consumo de alcohol, calidad del sueño y variables sociodemográficas⁷.

Inicialmente se realiza un análisis de la distribución de la muestra y sus elementos estadísticos. Se observa una edad promedio de 47.9 años (IC 95%: 46.89–48.84) con un índice de masa corporal (IMC) promedio de 30.5 kg/m² (IC 95%: 30.00–30.91), lo que indica una población predominantemente con sobrepeso u obesidad y con baja actividad física (ver Tabla 18).

Tabla 18

Estadísticas descriptivas de las variables seleccionada

Variable	Media	DE	Min	Max	Mediana	CI 95%
Edad	47.9	11.7	22.0	65.0	50.0	46.89–48.84
BMI	30.5	5.4	16.7	49.6	29.8	30.00–30.91
Cintura	100.3	13.1	61.2	137.5	99.3	99.16–101.34
Sístole	123.6	18.9	81.3	200.0	122.0	122.01–125.16
Kg (ganados)	1.5	3.0	-	20.0	-	1.22–1.72
Kg (perdidos)	2.6	4.4	-	21.0	-	2.22–2.96
Consumo de alcohol	13.3	9.6	5.0	34.0	6.0	12.49–14.09

⁷ Véase Diagrama de Inferencia 5 en Anexo 2, Tabla 31.

(litros por año)						
Actividad física (vigorosa)	1.0	2.1	-	7.0	-	0.86–1.22
Actividad física (moderada)	3.6	3.0	-	7.0	3.0	3.32–3.83
Caminar como actividad física	1.6	2.0	-	12.0	0.7	1.40–1.74
Sedentario	3.0	2.5	-	12.0	2.0	2.81–3.24
Muestra (n): 550						

Fuente: Elaboración propia con base en Ensanut 2022

En cuanto a la distribución de las variables categóricas, el 48.7% eran mujeres, mientras que el 51.3% eran hombres. Con respecto a la presencia de ENT, el 51.8% de los participantes reportaron al menos una condición (diabetes, hipertensión o eventos cardiovasculares) mientras que el 48.2% no reportaron dichas condiciones. En lo que respecta a los factores sociodemográficos, se observa una amplia variedad de niveles salariales similar a los niveles educativos de la población que también presentan una gran diversidad (ver Tabla 19).

Tabla 19

Distribución proporcional de las variables categóricas seleccionadas - ENSANUT 2022

Variable	Categoría	%
NCD	No	50.36
	Si	49.64
Sexo	Hombre	52.36
	Mujer	47.63
LDL (Alto)	No	75.27
	Si	24.73
Percepción de cansancio	Nunca	53.27
	Un Poco	24.00
	Mucho	13.82
	Siempre	8.91
Durmió bien	Nunca	39.82
	Un Poco	24.36
	Mucho	19.27
	Siempre	16.55
Tabaquismo	No fuma	83.27
	Todos los días	10.00
	Algunos días	6.73
Educación	Secundaria	32.00

	Primaria	24.00
	Preparatoria	20.36
	Licenciatura	12.55
	Ninguna	5.27
	Técnico	4.91
	Posgrado	0.91
Salario (Rango de pesos en miles)	6–9.9	45.09
	10–13.9	32.00
	14–21.9	14.36
	>22	8.55
Horas de sueño	8 h	33.09
	7 h	23.27
	6 h	18.55
	<9 h	12.73
	<5 h	12.36
Triglicéridos (Alto)	No	73.09
	Yes	26.91

Fuente: Elaboración propia con base en Ensanut 2022

Con respecto a la variable de respuesta, los datos no presentaron problemas de desbalance, por lo que no se aplicó ningún método de submuestreo o sobre muestreo para abordar esta condición. La proporción de datos faltantes fue mínima, por lo que se decidió eliminar estos registros incompletos, asegurando la consistencia entre los indicadores seleccionados. No se aplicaron operaciones de normalización ni escalamiento, ya que todas las variables eran categóricas o estaban medidas en escalas comparables, y dado que esto no impactaba en los métodos de aprendizaje automático utilizados. Además, esta decisión permitió una interpretación directa de los resultados del modelo SEM.

Selección de variables mediante ML

Con base en los resultados obtenidos de los modelos de *Random Forest* y regresión penalizada *Lasso*, se procedió a realizar una integración metodológica para la selección de variables relevantes a ser incluidas en el modelo de ecuaciones estructurales (SEM). La combinación de ambos enfoques permitió aprovechar la capacidad del *Random Forest* para capturar relaciones no lineales y detectar interacciones complejas entre variables junto con la propiedad de regularización del modelo *Lasso* que elimina automáticamente aquellas variables con

escasa relevancia predictiva al reducir sus coeficientes a cero. Se seleccionaron las variables que coincidieron en ambos métodos como robustas y significativas, asegurando una parsimonia en el modelo estructural sin perder capacidad explicativa. Esta estrategia garantizó que el SEM estuviera construido sobre un conjunto depurado de predictores derivados empíricamente mediante algoritmos de aprendizaje automático, lo que fortalece la validez del modelo en términos estadísticos y conceptuales.

El modelo *Random Forest* alcanzó una precisión del 78.66% en el conjunto de prueba. El modelo obtuvo una sensibilidad del 62.67%, una especificidad del 93.29%, un VPP del 89.52% y un VPN del 73.21%. Estos resultados indican que, si bien el modelo identifica eficazmente los verdaderos negativos (alta especificidad) y mantiene un fuerte VPP, presenta una sensibilidad moderada, lo que sugiere que hay margen de mejora en la detección de casos positivos verdaderos de ENT. No obstante, para fines de selección de variables, el modelo muestra resultados consistentes dentro del marco conceptual.

Por su parte, la regresión penalizada *Lasso* utiliza la penalización con la norma L_1 de modo que los coeficientes de aquellas variables que no son muy relevantes en la predicción del modelo se reducen a cero. En este caso, el desempeño del modelo *Lasso* es muy cercano al generado por *Random Forest*, arrojando las siguientes métricas de desempeño: una precisión del 71.27%, una sensibilidad del 64.23%, una especificidad del 78.26%, un VPP del 74.58% y un VPN del 68.79%. Los resultados pueden observarse en la Tabla 20.

Tabla 20

Comparación de desempeño de modelos de selección de variables

Métrica	<i>Random Forest</i>	<i>Lasso</i>
Precisión	78.66%	71.27%
Sensibilidad	62.67%	64.23%
Especificidad	93.29%	78.26%
VPP	89.52%	74.58%
VPN	73.21%	68.79%

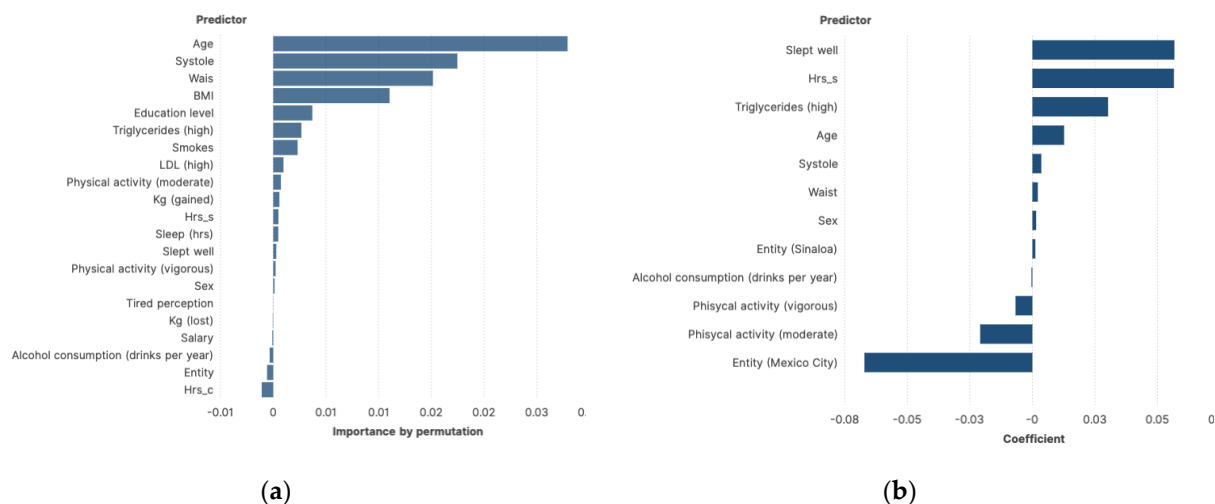
Fuente: Elaboración propia basado en ENSANUT 2022

Los resultados indican que la edad es el factor más influyente, hallazgo consistente con estudios epidemiológicos que muestran que el envejecimiento es uno de los principales impulsores

de la prevalencia de enfermedades crónicas. Otras variables relevantes incluyen la presión arterial sistólica, la circunferencia de cintura, el IMC y el nivel educativo de los individuos (véase Figura 22).

Figura 22

Características seleccionadas mediante aprendizaje automático: (a) Random Forest; (b) Regresión Lasso



Fuente: Elaboración propia con base en Ensanut 2022

Las variables seleccionadas por el modelo *Lasso* son consistentes con las reportadas por el modelo *Random Forest*. Las principales variables que contribuyen a la presencia de ENT son la percepción de dormir bien, las horas de sueño (hrs_s), los niveles altos de triglicéridos, la edad y la presión arterial sistólica. Por otro lado, las variables que disminuyen la probabilidad de desarrollar ENT son la actividad física moderada y vigorosa, así como el estado de residencia (Ciudad de México) (véase Figura 22b).

El análisis demuestra la naturaleza complementaria de los métodos de selección de variables. Mientras que la regresión *Lasso* proporciona coeficientes fácilmente interpretables, puede pasar por alto relaciones no lineales entre las variables predictoras y la variable de respuesta. En contraste, el modelo *Random Forest* permite capturar patrones complejos e interacciones entre variables. Su capacidad para manejar un gran número de predictores y detectar relaciones no lineales lo hace particularmente eficaz para conjuntos de datos complejos.

La combinación de ambos enfoques de selección de variables permite identificar variables relevantes para la determinación de la variable de respuesta. Combinado con el marco conceptual de la relación entre los factores de riesgo o hábitos de salud negativos y la presencia de ENT, esto permitió la definición de un modelo SEM coherente basado en la evidencia empírica identificada.

Este enfoque estructural es particularmente útil para captar la multidimensionalidad del fenómeno en estudio, ya que considera simultáneamente los efectos individuales de cada variable y sus relaciones dentro de un sistema teórico coherente. Así, el modelo no solo estima el impacto de cada hábito específico, sino que también proporciona una visión más holística sobre cómo la configuración general del estilo de vida incide en la salud de la PEA mexicana.

El modelo de ecuaciones estructurales propuesto contempla el uso de un modelo con variables latentes. Para ello, se definieron seis variables latentes: Y_1 (Perfil), Y_2 (Biometría), Y_3 (Estrés), Y_4 (Bioquímica), Y_5 (Consumo) y Y_6 (Hábitos de salud negativos). Estas se basaron en conjuntos relacionados de variables observadas y consideraron los resultados de los modelos de regresión *Lasso* y *Random Forest* como mecanismo de selección de variables.

- Y_1 (Perfil) se conformó a partir de las variables observadas X_1 (edad) y X_2 (escolaridad).
- Y_2 (Biometría) incluyó X_3 (IMC), X_4 (circunferencia de cintura), X_5 (presión arterial sistólica), X_6 (kg ganados/perdidos) y X_7 (kg ganados/perdidos).
- Y_3 (Estrés) se construyó con las variables X_8 (cansancio), X_9 (horas de sueño) y X_{10} (dormir bien).
- Y_4 (Bioquímica) se compuso de X_{11} (triglicéridos altos) y X_{12} (LDL alto).
- Y_5 (Consumo) incluyó X_{13} (consumo de alcohol) y X_{14} (consumo de tabaco).

La actividad física vigorosa fue considerada como una variable externa X_{15} y posteriormente, la variable latente Y_6 (hábitos de salud negativos) se construyó integrando las variables latentes Y_1 a Y_5 junto con dos variables exógenas adicionales: X_{15} (actividad física) y X_{16} (salario). El modelo de medición se define de acuerdo con las siguientes ecuaciones (Ecuaciones 37-42).

$$Y_1 = \sim X_1 + X_2 \quad (37)$$

$$Y_2 = \sim X_3 + X_4 + X_5 + X_6 + X_7 \quad (38)$$

$$Y_3 = \sim X_8 + X_9 + X_{10} \quad (39)$$

$$Y_4 = \sim X_{11} + X_{12} \quad (40)$$

$$Y_5 = \sim X_{13} + X_{14} \quad (41)$$

$$Y_6 = \sim Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + X_{15} + X_{16} \quad (42)$$

La variable de resultado W_1 (ENT) se modeló utilizando Y_6 , lo que nos permitió comprender cómo las variables latentes y exógenas influyen en el estado de salud de los individuos. El modelo estructural se presenta en la ecuación (43).

$$W_1 \sim Y_6 \quad (43)$$

En la siguiente sección se presentan los resultados del modelo SEM identificado y validado, el cual permite medir los efectos causales directos e indirectos entre las variables sobre las ENT. Dadas las características del modelo y las variables utilizadas, las estimaciones se realizaron utilizando el método de mínimos cuadrados ponderados (WLS, por sus siglas en inglés). Se evaluaron diferentes especificaciones y estructuras y se seleccionó el modelo de medición y estructural en el que todas las variables consideradas resultaron estadísticamente significativas. Además, se tomaron en cuenta aquellos modelos que presentaron adecuados indicadores de ajuste y coherencia en los coeficientes esperados de acuerdo con el modelo conceptual (véase Tabla 21).

Se observa un buen ajuste con un valor de chi-cuadrado de 449.186 (p-valor = 0.000), lo que indica que el modelo se ajusta bien a los datos. Las medidas de ajuste global como el TLI y el CFI superan el umbral de 0.9, mientras que el RMSEA y el SRMR están por debajo de 0.06, lo que sugiere un buen ajuste del modelo.

Tabla 21

Bondad de ajuste del modelo SEM

Índice	Valor
<i>Chi-cuadrada</i>	449.186
<i>Tucker-Lewis Index</i>	111.000
<i>Comparative Fit Index</i>	0.000
<i>Root Mean Squared Error Approximation</i>	0.074
<i>Standardized Root Mean Square Residual</i>	0.077

Fuente: Elaboración propia con base en Ensanut 2022.

El modelo de medición (análisis factorial confirmatorio) ofrece una visión detallada de cómo se relacionan las variables latentes con las observadas. Se observa que las variables latentes están significativamente influenciadas por sus variables observadas y que los signos corresponden con lo esperado. Por ejemplo, la presión arterial sistólica (X5) tiene un efecto positivo sobre la variable latente biomarcadores (Y2), mientras que la variable Kg perdidos (X6) tiene un impacto negativo sobre esta misma variable latente.

Los modelos estructurales muestran el efecto causal de las variables exógenas salario (X16) y actividad física vigorosa (X15), las cuales impactan negativamente en los hábitos de salud negativos. La variable latente hábitos de salud negativos (Y6) tiene un efecto positivo sobre las ENT (W1) (véase Figura 23).

El Modelo SEM presentado en la Figura 23 permite analizar de manera integrada cómo los determinantes sociodemográficos, biométricos, bioquímicos y conductuales se articulan para explicar la presencia de ENT en la población económicamente activa. A diferencia de un enfoque puramente descriptivo, este modelo posibilita interpretar relaciones causales estandarizadas, tanto directas como indirectas, entre las variables latentes y observadas.

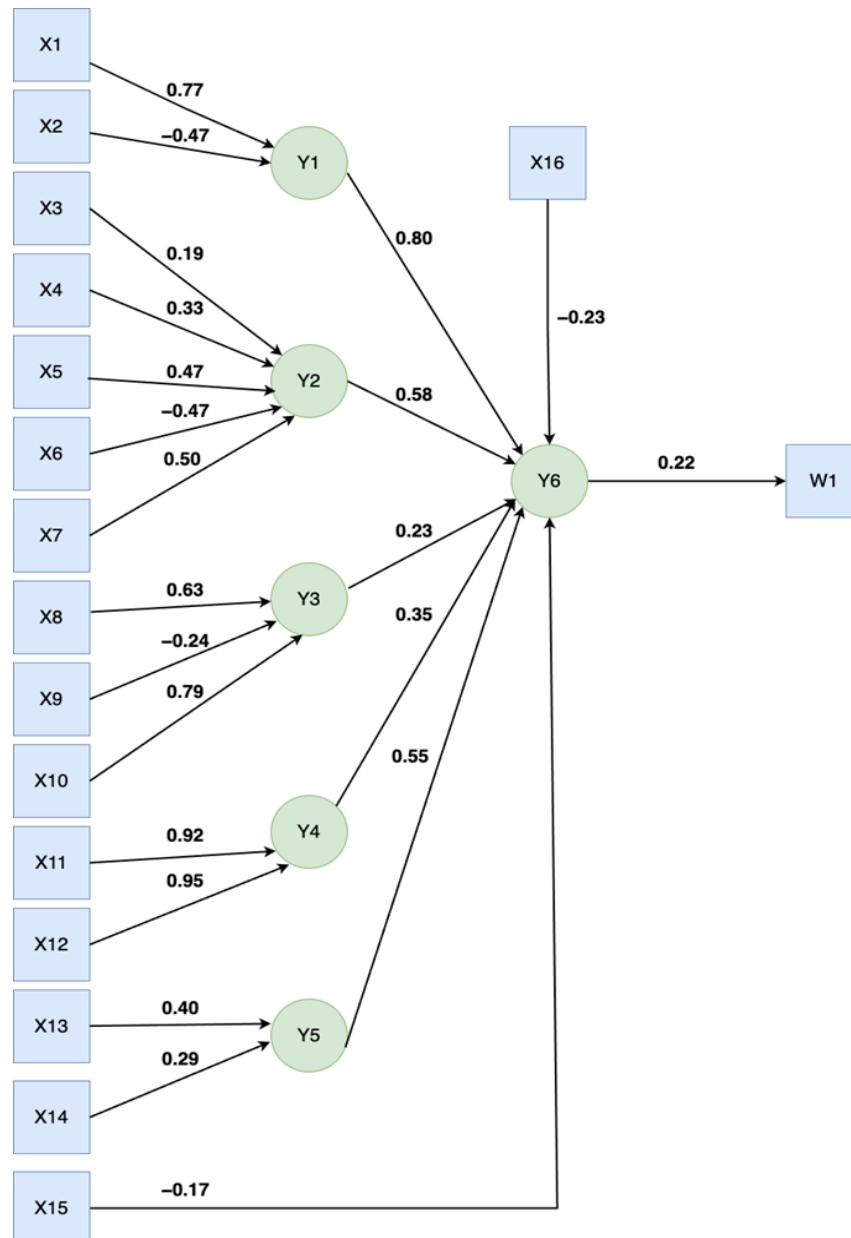
El análisis factorial confirmatorio muestra que cada variable latente está adecuadamente representada por sus indicadores observados, con cargas estandarizadas significativas y signos coherentes con la literatura. En particular, el constructo Biometría (Y_2) está fuertemente determinado por la presión arterial sistólica ($Std.lv = 8.83$) y la circunferencia de la cintura ($Std.lv = 4.33$), indicadores ampliamente reconocidos como predictores de riesgo cardiometabólico. El coeficiente negativo asociado a los kilogramos perdidos ($Estimate = -1.39$, $p < 0.01$) sugiere que la reducción de peso se asocia con una mejora en el perfil biométrico, reforzando la interpretación clínica del modelo.

De manera complementaria, la variable latente Estrés (Y_3) evidencia que la calidad del sueño desempeña un rol central en la regulación del estrés: dormir bien presenta la mayor carga estandarizada ($Std.lv = 0.79$), mientras que una menor cantidad de horas de sueño se asocia negativamente con este constructo ($Std.lv = -0.23$), lo que confirma que la insuficiencia del descanso incrementa la vulnerabilidad al estrés.

En el ámbito conductual, la variable Consumo (Y_5) está dominada principalmente por el consumo de alcohol ($Std.lv = 3.83$), superando ampliamente la contribución del tabaquismo ($Std.lv = 0.29$). Este resultado sugiere que, dentro de la PEA mexicana, el consumo de alcohol constituye un componente más relevante de los hábitos nocivos que el consumo de tabaco, al menos en términos de su contribución al riesgo global capturado por el modelo.

Figura 23

Resultados del modelo SEM estimado



Fuente: Elaboración propia con base en Ensanut 2022

La variable bioquímica (Bioq) se caracteriza por niveles elevados de triglicéridos y de colesterol LDL. Ambos indicadores presentan altas cargas estandarizadas, ligeramente mayores en el caso del LDL ($Std.lv = 0.98$ frente a 0.95), lo que indica una fuerte relación entre estos marcadores y el constructo latente bioquímico.

La variable latente de actividad física (AF) está representada por un solo indicador (AF_v) mientras que los hábitos de salud negativos (HaSn) son una variable de segundo orden que integra múltiples constructos primarios: desempeño (Perf), biomarcadores (Biom), bioquímica (Bioq), consumo (Cons), estrés (Strs) y actividad física (AF). Los indicadores que mayor impacto tienen en HaSn son el desempeño ($Std.lv = 0.80$), los biomarcadores ($Std.lv = 0.58$) y el consumo ($Std.lv = 0.55$).

Es especialmente relevante que el coeficiente asociado a la actividad física vigorosa sea negativo ($Estimate = -0.051$), lo que indica que mayores niveles de actividad física actúan como un factor protector, reduciendo la intensidad de los hábitos de salud negativos. Este resultado respalda la evidencia empírica que posiciona a la actividad física como un mecanismo clave de prevención primaria en la población trabajadora.

Además, se observa que los hábitos de salud negativos están influenciados significativamente por la edad (a través de Perf), así como por las condiciones de salud (ENT). En concreto, ENT predice positivamente a HaSn ($Estimate = 4.03, p < 0.01$) indicando que un mayor nivel de ENT se asocia con hábitos de salud más negativos. A su vez, ENT es explicada por los hábitos de salud negativos ($Estimate = 0.258, p < 0.05$), aunque con un coeficiente moderado.

De manera complementaria, se observó que la variable latente “estrés” está influenciada por la sensación de cansancio y los hábitos de sueño, mientras que los biomarcadores químicos están determinados por los niveles altos de triglicéridos y LDL. El consumo de alcohol tuvo un efecto significativo sobre la variable latente consumo, mientras que el tabaquismo mostró un efecto menor. Por otro lado, la actividad física vigorosa y el salario estuvieron directamente relacionados con los hábitos de salud negativos, impactándolos de forma negativa. Las relaciones de covarianza identificadas fueron consistentes con los resultados esperados. Una visión completa de los parámetros estimados y sus respectivos intervalos de confianza se presenta en la Tabla 22.

El modelo estructural se construyó utilizando la variable ENT (W1) como variable dependiente, mientras que los hábitos de salud negativos (Y6) se incluyeron como variable independiente. Esta última se trató como una variable latente que engloba múltiples indicadores observados relacionados con conductas perjudiciales para la salud, la actividad física, el tabaquismo y el consumo de alcohol.

Tabla 22

Resultados ecuaciones estructurales ENSANUT

Término	op	estimador	DE	Estadístico	<i>p.value</i>	<i>std.lv</i>	<i>std.all</i>	<i>std.nox</i>
Perf =~ edad	=~	1.0	0.0			9.017	0.766	0.766
Perf =~ grad_esc	=~	-0.054	0.008	-7.113	0.0	-0.486	-0.47	-0.47
Biom =~ BMI	=~	1.0	0.0			1.026	0.189	0.189
Biom =~ cintura	=~	4.224	0.742	5.696	0.0	4.332	0.332	0.332
Biom =~ sistole	=~	8.613	2.622	3.285	0.001	8.834	0.468	0.468
Biom =~ kg_ganados	=~	2.194	0.651	3.369	0.001	2.25	0.505	0.505
Biom =~ kg_perdidos	=~	-1.39	0.441	-3.154	0.002	-1.426	-0.475	-0.475
Strs =~ cansado	=~	1.0	0.0			0.637	0.634	0.634
Strs =~ sueno	=~	-0.372	0.09	-4.135	0.0	-0.237	-0.237	-0.237
Strs =~ dormido_bien	=~	1.252	0.321	3.897	0.0	0.798	0.792	0.792
Bioq =~ tiglic_alto	=~	1.0	0.0			0.948	0.924	0.924
Bioq =~ LDL_alto	=~	1.032	0.145	7.137	0.0	0.978	0.951	0.951
Cons =~ copas_ano_alcoh	=~	1.0	0.0			3.839	0.399	0.399
Cons =~ fuma	=~	0.075	0.037	2.011	0.044	0.288	0.286	0.286

Fuente: Elaboración propia con base en Ensanut 2022

Uno de los hallazgos más relevantes del modelo estructural es el coeficiente negativo y altamente significativo del salario (X_{16}) sobre los hábitos de salud negativos ($Estimate = -1.76$, $p < 0.001$). Este resultado sugiere que mejores niveles salariales están asociados con una menor

probabilidad de presentar hábitos nocivos para la salud, incluso después de controlar por edad, escolaridad, biomarcadores, estrés y consumo.

Desde una perspectiva económica y de salud pública, este efecto puede interpretarse como evidencia de que el ingreso actúa como un determinante social de la salud, facilitando el acceso a mejores condiciones de vida, mayor seguridad alimentaria, servicios de salud preventivos y entornos que favorecen conductas saludables. En consecuencia, el salario no solo funciona como una variable económica, sino como un factor protector indirecto frente al desarrollo de ENT a través de su impacto sobre los hábitos de salud.

Este resultado tiene implicaciones relevantes para el diseño de políticas públicas y estrategias empresariales, ya que sugiere que mejoras en las condiciones laborales y salariales podrían generar externalidades positivas en salud, complementando los programas tradicionales de promoción de estilos de vida saludables.

El modelo confirma que los hábitos de salud negativos tienen un efecto positivo y estadísticamente significativo sobre la presencia de ENT (*Estimate* = 0.258, $p < 0.05$). Aunque el coeficiente es moderado, su significancia indica que cambios en los hábitos de salud pueden traducirse en variaciones relevantes en el riesgo de enfermedades crónicas. En conjunto con el efecto protector del salario y de la actividad física, estos resultados refuerzan la importancia de intervenciones integrales que aborden simultáneamente factores conductuales y contextuales.

4.3 Impacto de los hábitos de salud en la productividad laboral en México

Con el objetivo de estimar el impacto de los hábitos de salud en la productividad laboral, se empleó la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022 como fuente principal de información. Esta base de datos proporciona variables clave sobre ingresos laborales, horas trabajadas y características sociodemográficas de los hogares mexicanos. A partir de estos datos, se construyó una medida aproximada de productividad individual calculando el cociente entre ingresos y horas trabajadas por cada integrante de la PEA⁸. Para abordar la naturaleza compleja y potencialmente endógena de esta relación, se aplicó el enfoque de *Double Machine Learning*, una técnica de inferencia causal que permite estimar efectos promedio del tratamiento ajustando por múltiples variables de confusión mediante algoritmos de aprendizaje automático.

⁸ Véase Diagrama de Inferencia 6 en Anexo 2, Tabla 32.

Esta metodología resultó especialmente adecuada para analizar datos de alta dimensionalidad y proporcionó estimaciones robustas y precisas sobre cómo los hábitos de salud pueden influir en el desempeño laboral.

El procesamiento de la base ENIGH 2022 implicó una serie de etapas técnicas destinadas a preparar una estructura de datos coherente y estandarizada. Se descargaron los microdatos en formato Stata (.dta) correspondientes a los cuatro trimestres del año. Posteriormente, se procedió a deflactar los valores de ingreso y gasto, ajustando los montos trimestrales mediante factores inflacionarios oficiales. Este procedimiento sigue lineamientos recomendados para la homogeneización temporal de variables económicas (INEGI, 2020; Martínez *et al.*, 2019).

Posteriormente, las diferentes bases de datos de la ENIGH fueron integradas, estandarizadas y limpiadas. Esta fusión permitió consolidar un repositorio único con variables económicas, sociodemográficas y de salud. Tras la integración, se efectuaron transformaciones iniciales, como la conversión de variables categóricas a factores.

Se realizó el cálculo de la productividad considerando solo el ingreso monetario por actividad laboral mediante la ecuación (44) (INEGI, 2015):

$$Productividad = \frac{Ingreso\ semanal}{Horas\ trabajadas\ a\ la\ semana} \quad (44)$$

A continuación, se seleccionó la submuestra correspondiente a la PEA, delimitando el análisis al nivel individual. Se eliminaron las variables no relevantes, obteniendo una base concentrada en variables clave como ingreso mensual, horas trabajadas, presencia de ENT, nivel educativo, sexo, región y gasto doméstico. Durante esta etapa se verificó la consistencia de los tipos de datos, se estandarizaron etiquetas y se aplicó el factor de expansión poblacional.

Después de aplicar criterios de filtrado, el conjunto de datos final constó de 56,088 registros. El factor de expansión agregado, que permite ponderar los resultados para representar a la población general, fue de 23,420,993 individuos.

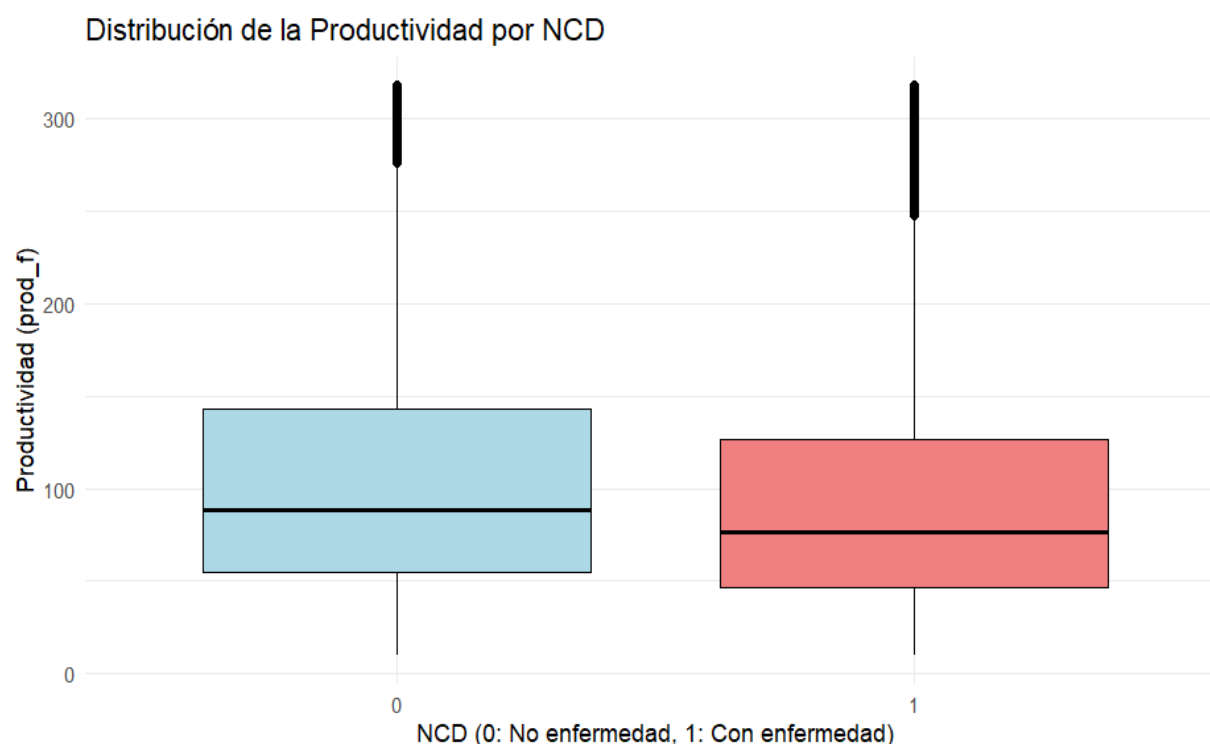
Se procedió a un análisis descriptivo inicial que incluyó resúmenes estadísticos de las principales variables, validación de etiquetas y generación de histogramas y *boxplots* para examinar la distribución de la productividad (prod_22) en función de la presencia de ENT (NCD). Se implementaron filtros para eliminar valores extremos en las variables hor_1 y prod_22 y se

definió un umbral con base en el percentil 95 de la productividad para mitigar la influencia de valores atípicos. Se calcularon estadísticas de tendencia central y dispersión para los grupos $NCD = 0$ y $NCD = 1$ y se aplicaron pruebas t para evaluar la significancia estadística de las diferencias observadas.

El análisis descriptivo de la productividad ($prod_22$) reveló diferencias significativas entre los grupos con y sin ENT. En el grupo sin ENT, la productividad promedio fue de 112 pesos por hora laborada a la semana, con una mediana de 88, desviación estándar de 78 y un rango intercuartil de 54.7 a 143. Por otro lado, en el grupo con ENT, la productividad promedio fue de 101 con una mediana de 76.2, desviación estándar de 76.6 y un rango intercuartil de 46 a 127. Estos resultados muestran que las personas con ENT tienen una productividad menor en comparación con aquellas sin estas enfermedades; esto se observa en la Figura 24.

Figura 24

Comparación de productividad entre empleados con y sin ENT



Fuente: Elaboración propia con base en la ENIGH 2022

Para evaluar si las diferencias observadas en productividad entre los grupos eran estadísticamente significativas, se realizó una prueba t de Welch como se muestra en la Tabla 23.

Los resultados confirmaron que la diferencia en las medias de productividad entre los grupos con y sin ENT es estadísticamente significativa, con una media de 112 en el grupo sin ENT y de 101 en el grupo con ENT. El intervalo de confianza del 95% para esta diferencia fue de 10.48 a 14.55 puntos, y el valor p fue menor que $2.2e-16$, lo que respalda la hipótesis de que las ENT tienen un impacto negativo en la productividad promedio.

Tabla 23

Resultados de la prueba de Welch para comparación de medias

Grupo	Media	IC 95%	Valor p
Sin ENT	112	[111, 113]	$< 2.2e-16$
Con ENT	101	[99.9, 101]	$< 2.2e-16$

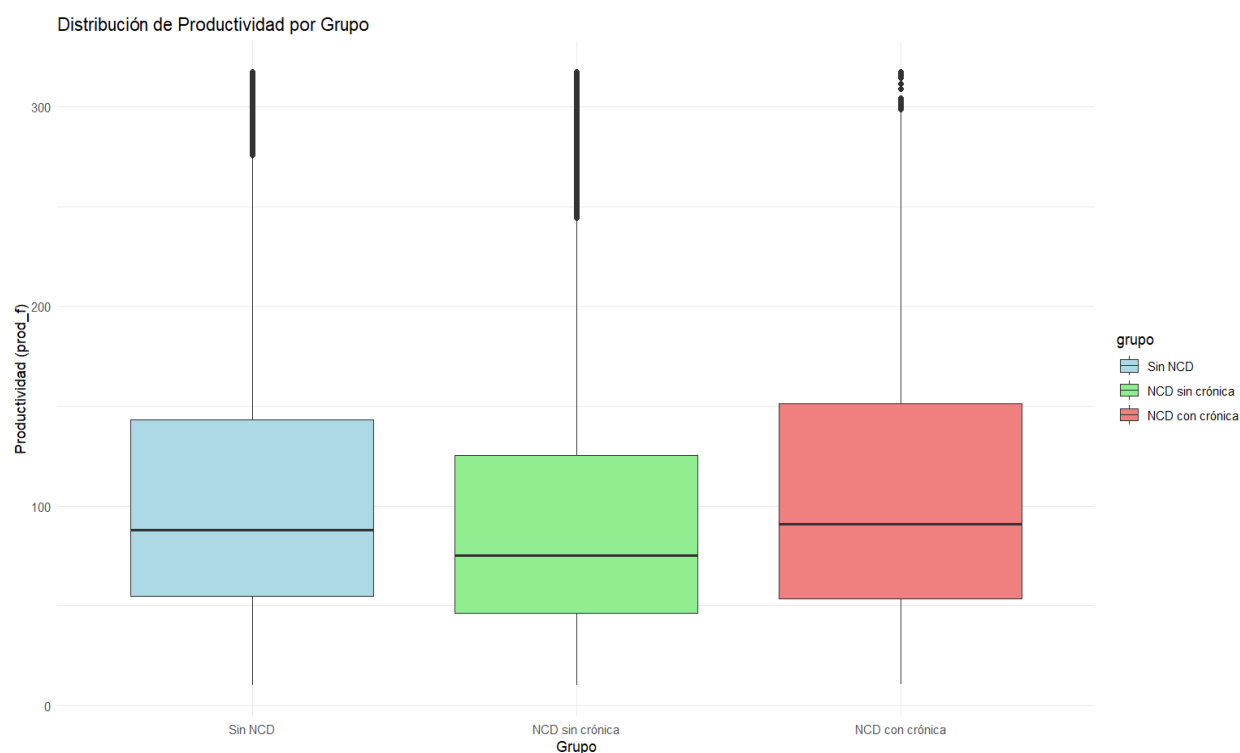
Fuente: Elaboración propia con base en la ENIGH 2022

Los hallazgos destacan que las personas con ENT presentan una productividad significativamente menor con una diferencia promedio de 11 puntos en comparación con aquellas sin ENT. Además, la presencia de valores extremos y la no normalidad en las variables clave subrayan la complejidad de los datos y la importancia de aplicar métodos robustos en los análisis futuros. Desde una perspectiva económica, estas diferencias en productividad sugieren un impacto considerable asociado a las ENT, lo que resalta la necesidad de implementar estrategias efectivas tanto a nivel empresarial como en el ámbito de la política pública.

Se definió una nueva variable categórica para distinguir entre tres subgrupos: personas sin ENT, personas con ENT sin condición crónica y personas con ENT y condición crónica. La identificación de ENT crónicas se realizó a partir de tres criterios: presencia de gasto en salud, número de visitas médicas y auto reporte de afectación laboral. Aquellos individuos que cumplieron al menos dos de estos criterios fueron clasificados con condición crónica. La distribución mostró que la mayoría de los empleados pertenecen al grupo de ENT sin crónica (18,867 empleados), seguido por el grupo sin ENT (35,123 empleados), mientras que el grupo de ENT con crónica es considerablemente menor (2,098 empleados), como se puede observar en la Figura 25.

Figura 25

Distribución de productividad por grupo



Fuente: Elaboración propia con base en la ENIGH 2022

El análisis descriptivo de la productividad ajustada (prod_22) mostró diferencias sustanciales entre los tres grupos analizados. El grupo sin NCD presentó una productividad promedio de 112, con una mediana de 88 y un rango intercuartil (RIC) de 54.7 a 143. Por su parte, el grupo con NCD sin enfermedades crónicas mostró la productividad promedio más baja, con 99.7, una mediana de 75 y un RIC de 46.3 a 125. En contraste, el grupo con NCD y enfermedades crónicas reportó la productividad promedio más alta, 117, con una mediana de 90.8 y un RIC de 53.5 a 151. Aunque podría esperarse que las condiciones crónicas disminuyeran la productividad, el grupo con NCD crónica presenta valores superiores incluso al grupo sin NCD. Este hallazgo podría deberse a una sobreestimación en la medición de la productividad ajustada, derivada de una menor cantidad de horas trabajadas, posiblemente por asistencia médica o limitaciones funcionales, que al usarse como denominador en la ecuación de productividad (44), impactarían el valor final del indicador.

Adicionalmente, se exploró la normalidad de las variables numéricas mediante pruebas como Anderson-Darling, observándose que *hor_1* no seguía una distribución normal. Se realizaron análisis visuales (*boxplots*, densidades) y estadísticos (ANOVA, prueba *post-hoc* de Tukey) para evaluar diferencias significativas entre los tres grupos. El análisis ANOVA reveló diferencias estadísticamente significativas, con un F-valor de 179.3 y un valor p de $2e-16$. Esto confirma que existen diferencias relevantes en la productividad promedio entre los grupos. Para identificar cuáles grupos difieren específicamente, se realizó una prueba *post-hoc* de Tukey. Los resultados mostraron que el grupo ENT sin crónica tiene una productividad significativamente menor que el grupo sin ENT con una diferencia promedio de -12.08. Por otro lado, el grupo ENT con crónica mostró una productividad significativamente mayor que el grupo sin ENT con una diferencia promedio de 5.14. Finalmente, el grupo ENT con crónica también superó significativamente al grupo ENT sin crónica con una diferencia promedio de 17.23. Se encuentra que el subgrupo de los pacientes con enfermedades crónicas genera un problema en la parametrización ya que, al calcular la productividad como ingresos entre horas laborales, la presencia de condiciones crónicas puede alterar los resultados por la dificultad de capturar con precisión el fenómeno del presentismo, es decir, cuando los empleados acuden a trabajar pese a su enfermedad, lo que genera una sobreestimación de su productividad aparente. Este comportamiento, lejos de reflejar un mayor rendimiento, puede deberse a la necesidad económica o a la ausencia de políticas laborales flexibles, generando una productividad “sesgada” que no se corresponde con su verdadero estado de salud ni con la calidad del trabajo realizado.

Posteriormente, se aplicaron modelos causales para estimar el impacto de las ENT en la productividad, empleando *Double Machine Learning* (DML) con el objetivo de estimar de forma robusta el impacto de las ENT sobre la productividad laboral, lo cual permite controlar múltiples variables de confusión de forma flexible. Este enfoque ha sido validado en contextos de alta dimensionalidad y es especialmente útil en entornos observacionales con múltiples covariables (Athey & Imbens, 2016).

Se excluyeron los individuos con ENT crónicas y se consideraron las siguientes variables usadas a nivel de individuo conforme a lo mostrado en la Tabla 24. Las variables seleccionadas cumplen la función de covariables explicativas o de control dentro del modelo de DML. Su incorporación permite ajustar el efecto de la variable de tratamiento sobre el resultado, reduciendo

el sesgo derivado de factores de confusión. Estas covariables recogen dimensiones sociodemográficas (sexo, edad), de comportamiento (consumo de alimentos, tabaquismo), de salud (peso, presencia de enfermedades crónicas) y económicas (gasto en salud, gasto catastrófico, incremento en el ingreso, quintil de ingreso), las cuales son fundamentales para explicar la variación observada en la población. La definición de estas variables se realizó bajo la premisa de equivalencia con la Encuesta Nacional de Salud y Nutrición (ENSANUT), lo cual garantiza comparabilidad y consistencia en la medición respecto a una de las principales fuentes de información epidemiológica y sociodemográfica en México y que corresponde con el modelo previamente estimado mediante ecuaciones estructurales, lo que asegura coherencia metodológica entre distintos enfoques analíticos.

Tabla 24

Descripción de covariables para el análisis DML

Variable	Medición	Relevancia en la medición
Consumo de alimentos	Gasto semanal en pesos destinado a cereales, carnes, pescado, leche, huevo, aceites y oleaginosas, verduras, frutas, azúcar, bebidas azucaradas.	Permite estimar la dieta y nutrición, así como la proporción del gasto destinada a alimentación, indicador del nivel socioeconómico.
Tabaquismo	Variable binaria: 1 = fumador, 0 = no fumador.	Factor de riesgo asociado a enfermedades crónicas no transmisibles (ENT) y gasto en salud.
Gasto en salud	Pesos semanales reportados en servicios médicos, medicinas, hospitalización y consultas.	Indica el esfuerzo económico destinado a la salud y se usa para calcular si hay gasto catastrófico.
Peso (categoría por BMI)	Categoría según Índice de Masa Corporal (BMI): - Normal: $BMI < 25$ - Sobrepeso: $25 \leq BMI < 30$ - Obesidad: $BMI \geq 30$	Permite clasificar riesgos de salud asociados a la nutrición y sobrepeso/obesidad.
Sexo	Masculino / Femenino	Variable sociodemográfica que permite analizar diferencias de género en consumo, salud y riesgos.

Edad	Años cumplidos.	Factor de riesgo y determinante del gasto en salud y prevalencia de enfermedades crónicas.
Condiciones crónicas de salud (ENT)	Variable binaria: 1 = tiene enfermedad crónica (ej. diabetes, hipertensión), 0 = no tiene.	Permite identificar la población vulnerable a incurrir en mayores gastos de salud.
Gasto catastrófico en salud (HGC)	Variable dicotómica: 1 = incurre en gasto catastrófico, 0 = no incurre. Criterio de la ENIGH/OMS: gasto de bolsillo en salud $\geq 30\%$ de la capacidad de pago del hogar (capacidad de pago = gasto total – gasto en alimentos).	Mide la presión financiera que los gastos de salud y su vulnerabilidad económica.
Quintil de ingreso	Clasificación ordinal en cinco grupos de igual tamaño de acuerdo con el ingreso corriente total per cápita del hogar. Quintil 1 = más bajo, Quintil 5 = más alto.	Permite analizar desigualdades económicas a nivel individual y segmentar la población según el nivel de ingresos, facilitando comparaciones de bienestar.
Entidad federativa (edo)	Variable de factor que representa a los 32 estados de la república	Variable que permite identificar de qué estado de la República Mexicana se encuentra el individuo contemplado en el análisis.

Fuente: Elaboración propia con base en ENIG, 2022

En términos formales, el modelo de *Double Machine Learning* (DML) aplicado para estimar el impacto de las ENT sobre la productividad ajustada (prod_22) puede expresarse mediante la ecuación (45 - 46):

$$\log Y_i = g(X_i) + \theta D_i + \zeta_i \quad (45)$$

$$D_i = m(X_i) + v_i \quad (46)$$

Donde:

$\log Y_i$: productividad individual transformada logarítmicamente

D_i : variable de tratamiento binaria (presencia de ENT).

X_i vector de covariables observados (Consumo de alimentos, tabaquismo, gasto en salud, gasto catastrófico en salud, peso, sexo, edad, condiciones crónicas de salud, quintil de ingreso, entidad federativa).

$g(X_i)$: función desconocida que modela el resultado en función de las covariables.

$m(X_i)$: función desconocida que modela la asignación del tratamiento condicional a las covariables.

θ : parámetro de interés, correspondiente al efecto causal promedio de las ENT sobre la productividad laboral.

ζ_i, v_i : errores residuales.

Para estimar θ , se realiza una corrección de doble descomposición de los efectos (Ver ecuación 47 y 48):

$$\zeta_i = Y_i - \hat{g}(X_i) \quad (47)$$

$$\hat{v}_i = D_i - \hat{m}(X_i) \quad (48)$$

Luego se estima el parámetro causal θ mediante regresión de los residuos (Ver ecuación 49):

$$\hat{\theta} = \frac{\sum_{i=1}^n \hat{\zeta}_i \hat{v}_i}{\sum_{i=1}^n \hat{v}_i^2} \quad (49)$$

Las funciones $\hat{g}(X_i)$ y $\hat{m}(X_i)$ se estiman mediante modelos de *machine learning* (como *Random Forest*, *Gradient Boosting* o *Lasso*).

Se utiliza validación cruzada en dos etapas (*sample-splitting* y *cross-fitting*) para evitar *overfitting* y garantizar propiedades asintóticas válidas.

El estimador $\hat{\theta}$ obtenido es ortogonal, es decir, robusto frente a pequeños errores en la estimación de $\hat{g}$ y $\hat{m}$.

En el enfoque DML, se utilizaron algoritmos de *Random Forest* para modelar funciones de regresión. Se evaluó el desempeño del modelo utilizando métricas como el RMSE. Se compararon múltiples algoritmos (*Random Forest*, SVM, *XGBoost*, regresión lineal, KNN) seleccionando el de menor error. La optimización de hiperparámetros del *Random Forest* se realizó mediante búsqueda aleatoria limitada a 10 iteraciones.

La Tabla 25 presenta los resultados de las estimaciones obtenidas mediante el enfoque de DML, utilizando diferentes algoritmos de aprendizaje automático tanto para el modelado de la variable de tratamiento (presencia de ENT) como para la función de resultado (productividad laboral). La variable respuesta es cuantitativa continua y corresponde a la productividad individual, medida como el logaritmo del ingreso por hora trabajada. Esta transformación logarítmica permite estabilizar la varianza y facilita la interpretación de los efectos como cambios relativos. En este contexto, el parámetro estimado θ representa el cambio promedio en la productividad, expresado en términos porcentuales aproximados, atribuible a la presencia de enfermedades no transmisibles (ENT), después de controlar por variables confusoras mediante aprendizaje automático. Específicamente, un valor de θ puede interpretarse como la variación porcentual en el ingreso por hora trabajada asociada a las ENT; los algoritmos de ML permiten controlar por covariables y al final aíslan solo el efecto del tratamiento sobre la respuesta.

Tabla 25

Resultados del modelo Double Machine Learning según algoritmo utilizado

Modelo de ML utilizado	RMSE
<i>Random Forest</i>	0.4714
<i>XGBoost</i>	0.4732
<i>SVM</i>	0.4923
<i>KNN</i>	0.5324

Fuente: Elaboración propia con base en la ENIGH 2022

En todos los modelos aplicados se observa una penalización negativa en la productividad laboral asociada a las ENT. Esto sugiere una relación adversa entre estas condiciones de salud y el desempeño económico individual. Entre los modelos evaluados, KNN presentó el mayor error de predicción, con un RMSE de 0.5324, mientras que el modelo *Random Forest* mostró el mejor desempeño, con un RMSE de 0.4714.

Se procedió a eliminar del análisis a los individuos con ENT que además presentan condiciones crónicas, debido a que este subgrupo introducía un efecto no deseado en las estimaciones de productividad. Al excluir estos casos, se busca obtener estimaciones más limpias y representativas del impacto real de las ENT en la PEA. Esta depuración permite reducir el sesgo introducido por la distorsión del presentismo, lo cual mejora la validez de los modelos predictivos

empleados y permite identificar con mayor precisión el efecto neto de las ENT sobre la productividad laboral. Los resultados subsecuentes reflejan un patrón más coherente y conservador, alineado con los supuestos teóricos y con las evidencias previas sobre el costo económico de estas enfermedades.

Al aplicar el modelo de *Double Machine Learning* utilizando *Random Forest* como algoritmo base para la estimación de los efectos de tratamiento y las funciones de control, se obtuvieron resultados consistentes con las hipótesis planteadas. En este contexto, el modelo de aprendizaje automático doble (DML) permite aislar el efecto del tratamiento respecto a la productividad y variables confusoras sobre la variable respuesta asociada a la presencia de enfermedades no transmisibles (ENT). Dado que la variable de resultado se encuentra en escala logarítmica, los coeficientes estimados se interpretan como variaciones porcentuales aproximadas en el ingreso por hora trabajada. En esta primera estimación, se observó una disminución promedio de 9.49% en la productividad ajustada atribuible a las ENT. No obstante, al repetir el análisis excluyendo específicamente al subgrupo de personas con ENT crónicas, debido a su posible efecto distorsionador relacionado con fenómenos como el presentismo, el efecto promedio se incrementó a 9.54%. Esta estimación se obtuvo utilizando los mejores hiperparámetros identificados: 695 árboles, 10 variables seleccionadas por división y un tamaño mínimo de nodo de 10.

Aunque la magnitud de esta diferencia puede parecer moderada, el patrón observado es relevante desde el punto de vista analítico. Los resultados muestran una mayor penalización al excluir a los pacientes con condiciones crónicas, lo cual sugiere que dicho subgrupo podría estar introduciendo un sesgo de atenuación en las estimaciones. Este hallazgo refuerza la necesidad de segmentar adecuadamente la muestra cuando se analizan impactos en productividad derivados de condiciones de salud, especialmente cuando los efectos no se expresan de manera homogénea entre subgrupos clínicamente diferenciados.

La consistencia en la dirección y significancia de los resultados entre ambos enfoques refuerza la robustez del análisis y sugiere que las ENT afectan negativamente la productividad laboral. Además, la diferencia observada entre los modelos que incluyen y excluyen al grupo "ENT con crónica" resalta la complejidad del impacto de las condiciones crónicas en la productividad. Este subgrupo parece tener un comportamiento particular que debe ser analizado

con mayor profundidad, ya que podría estar relacionado con estrategias adaptativas o características únicas que atenúan parcialmente los efectos negativos de las ENT.

Con el propósito de profundizar en los efectos diferenciales de las ENT en la productividad laboral, se procedió a realizar un análisis secundario enfocado en comparar a los individuos con ENT que también presentan condiciones crónicas con aquellos que padecen ENT, pero sin complicaciones crónicas asociadas, que fue asignado como un procedimiento de calibración del factor de penalización. El estudio parte del reconocimiento de que las ENT representan un determinante relevante de la productividad laboral, al asociarse con pérdidas económicas considerables en la PEA. Si bien el modelo de DML permitió estimar un efecto causal promedio de -9.54% en la productividad laboral de personas con ENT respecto a aquellas sin dichas enfermedades, se consideró pertinente profundizar en los efectos diferenciales asociados a la presencia de condiciones crónicas adicionales.

El planteamiento metodológico parte de la hipótesis de que las condiciones crónicas generan un fenómeno particular de presentismo, entendido como la asistencia al trabajo con una reducción significativa en el rendimiento. Bajo esta premisa, se procedió a ajustar la estimación de productividad a través de un factor de penalización que captura el deterioro adicional en individuos con ENT y cronicidad, en comparación con aquellos que solo presentan ENT sin complicaciones asociadas.

La lógica del ajuste se formaliza de la siguiente manera. Sea P_{ref} la productividad horaria promedio de un individuo sin ENT, y δ la penalización causal promedio estimada mediante DML para el grupo con ENT (-9.549%) resolviendo mediante la ecuación (50):

$$P_{ENT} = (1 + \delta) \times P_{ref} \quad (50)$$

Para el caso de las personas con ENT y enfermedades crónicas asociadas, se introdujo un coeficiente de presentismo $\theta \in (0,1]$, de forma que la productividad ajustada se expresa de acuerdo con la ecuación (51):

$$P_{ENT+cr} = \theta \times P_{ENT} \quad (51)$$

El parámetro θ representa la fracción de productividad retenida frente al grupo con ENT. Su complemento $(1-\theta)$ refleja la pérdida adicional atribuible al presentismo crónico.

Una vez establecido este marco, se implementó un procedimiento iterativo de calibración para estimar empíricamente el valor de θ . La estrategia consistió en modificar las horas laborales declaradas en la ENIGH, penalizando proporcionalmente a los individuos con cronicidad, y recalculando a partir de estas horas ajustadas una nueva medida de productividad de acuerdo con la ecuación (52):

$$P_i^\theta = \frac{\text{ingreso}_i}{4 \times \text{horas}_i^\theta} \quad (52)$$

$$\text{Donde: } \text{horas}_i^\theta = \begin{cases} \theta \times \text{hor}_i, & \text{si el individuo presenta cronicas} \\ \text{hor}_i, & \text{el resto de los casos} \end{cases}$$

Posteriormente, los datos ajustados se incorporaron a un nuevo modelo de DML, del cual se extrajo un coeficiente estimado $\hat{\beta}(\theta)$. El valor de θ se actualizó de manera iterativa hasta que la diferencia entre $\hat{\beta}(\theta)$ y el coeficiente objetivo β^* se encontró dentro de una tolerancia predefinida. De esta forma, el problema se formuló tal como se observa en la ecuación (53):

$$\theta^* = \arg \min_{\theta \in [0.6, 1]} (\hat{\beta}(\theta) - \beta^*)^2 \quad (53)$$

Para llevar a cabo la estimación se aplicaron dos enfoques complementarios:

- **Método de bisección**, en el que se partió de un intervalo inicial $[0.6, 1.0]$, evaluando iterativamente puntos medios y ajustando los límites del intervalo según el signo de la diferencia entre $\hat{\beta}(\theta) - \beta^*$.
- **Optimización unidimensional**, mediante la función *optimize()* de R, la cual minimiza de manera directa la función objetivo definida por la diferencia cuadrática.

Ambos procedimientos convergieron a valores similares para el factor de penalización, con resultados en el rango de $\theta \approx 0.8125$. Esto implica que las personas con ENT y cronicidad mantienen alrededor del 81.25% de su productividad respecto al grupo con ENT sin comorbilidades, lo que equivale a una pérdida adicional del 18.75% atribuible al presentismo crónico.

Los resultados del modelo DML ajustado mostraron que las ENT están asociadas con una penalización promedio de -11.44% en la productividad ajustada, con un error estándar de 0.3404. Este efecto es estadísticamente significativo ($p = 9.2 \times 10^{-9}$, IC $[-14.593, -8.398]$), lo que confirma

el impacto negativo de las ENT en la productividad laboral. El modelo también fue evaluado en términos de precisión predictiva, obteniendo un RMSE de 0.3404, lo que refleja una desviación promedio razonable entre las predicciones y los valores observados e indica una precisión adecuada al diferenciar entre individuos con y sin NCD. La Tabla 26 resume el impacto estimado bajo los modelos DML con y sin ajuste, incluyendo métricas de precisión como el RMSE.

Tabla 26

Impacto estimado de ENT sobre la productividad ajustada (modelos DML)

Modelo DML	Impacto Promedio	Intervalos de confianza	RMSE
Sin ajuste (base)	-9.499%	[-12.054, -5.770]	0.4714
Con ajuste por presentismo	-11.44%	[-14.593, -8.398]	0.3838

Fuente: Elaboración propia con base en la ENIGH 202

Estimación del impacto económico de las ENT sobre la productividad laboral

El análisis de la ENIGH 2022 permitió desarrollar una estimación rigurosa sobre el impacto económico de las ENT en la productividad laboral en México. La productividad laboral fue definida como el ingreso laboral por horas trabajadas a la semana, un indicador que refleja la eficiencia con la que cada persona convierte su tiempo laboral en ingresos monetarios. Mediante la aplicación de un modelo de DML, se calculó el efecto causal asociado a ENT sobre la productividad, lo que permitió controlar múltiples variables de confusión y evitar la sobreestimación o subestimación derivada de simples comparaciones de medias entre grupos.

Los resultados obtenidos mostraron que la presencia de ENT se asocia con una reducción proporcional de 11.44% en la productividad. En términos absolutos, la productividad promedio en personas sin ENT fue de 111.77 pesos por hora, mientras que para aquellos con ENT se redujo a 98.97 pesos por hora, lo que implica una pérdida de 12.8 pesos por hora trabajada. Este valor constituye la base para el cálculo del costo económico agregado.

A diferencia de aproximaciones simplificadas, en este caso se consideró tanto el promedio de horas trabajadas a la semana como el número de trabajadores con ENT ajustado por el factor de expansión de la ENIGH, lo cual asegura que la estimación refleje la dimensión poblacional y no únicamente la muestra. El costo total anual estimado por pérdida de productividad puede estimarse de acuerdo con la ecuación (54).

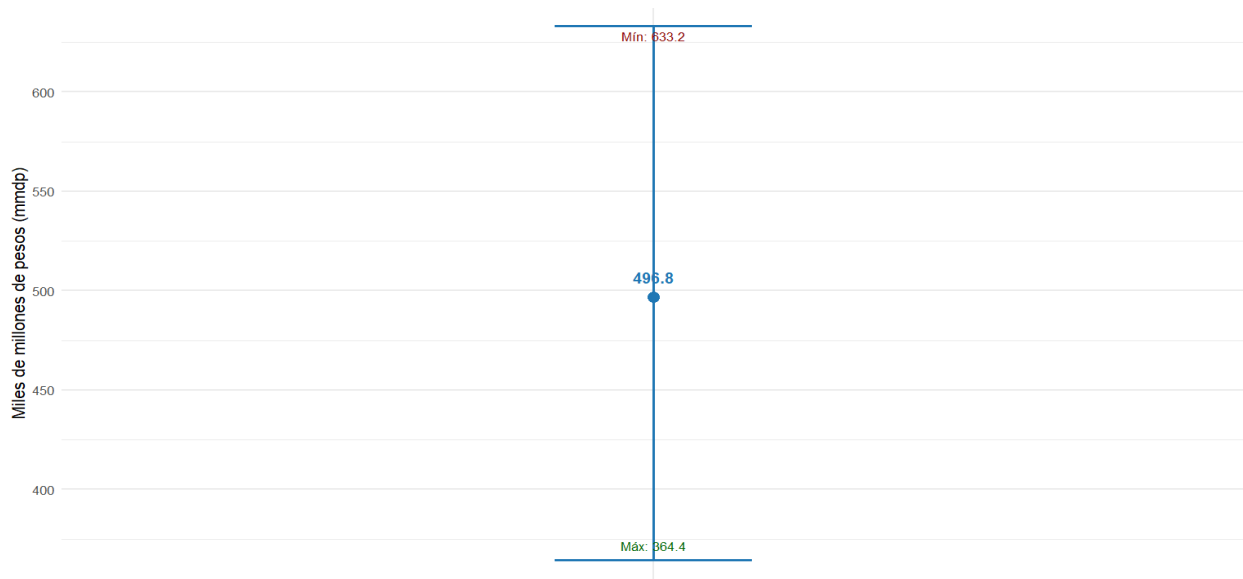
$$Costo\ Total\ Anual = \sum_{i=1}^{n_{ENT}} [\bar{Y}_{sin\ ENT} - \bar{Y}_{con\ ENT}] \times h_i \times f_i \times 52 \quad (54)$$

Donde $\bar{Y}_{sin\ ENT} - \bar{Y}_{con\ ENT}$ corresponde a la diferencia en productividad horaria (pesos/hora), h_i representa las horas trabajadas por semana del individuo i , f_i es el factor de expansión asociado a cada observación y n_{ENT} que corresponde al número de individuos con ENT en la muestra. De manera simplificada, este planteamiento equivale a multiplicar la diferencia promedio de productividad horaria por el promedio de horas semanales, por el número ponderado de personas con ENT y por el número de semanas en el año.

Al aplicar esta fórmula con los valores observados en la ENIGH 2022, se obtiene un costo sobre la pérdida de productividad total semanal de aproximadamente 9.5 mil millones de pesos, que al proyectarse a lo largo de las 52 semanas del año equivale a una pérdida de 496,817 millones de pesos anuales atribuibles al estilo de vida, presentismo y alimentación de personas con ENT en la población económicamente activa de México. Podemos observar en la Figura 26 los valores de pérdida anuales teniendo en cuenta los intervalos del parámetro estimado para pérdida de productividad.

Figura 26

Costo anual estimado de pérdida de productividad con intervalos de confianza 95%



Fuente: Elaboración propia con base en ENIGH 2022

La magnitud de este impacto adquiere una dimensión más clara al contrastarse con agregados macroeconómicos e institucionales relevantes. En primer lugar, si se compara con el Producto Interno Bruto (PIB) de México en 2022, estimado en 29.45 billones de pesos, la pérdida de productividad atribuible a las ENT representa aproximadamente el 1.68% del PIB nacional, una proporción comparable a la contribución anual de sectores económicos completos.

En segundo término, al contrastar este monto con el gasto público total en salud, que en 2022 ascendió a cerca de 1.5 billones de pesos (equivalente al 5.1% del PIB, INEGI, 2022), la pérdida estimada por ENT representa casi un tercio del gasto público anual en salud. Esta comparación pone de relieve que los costos económicos indirectos derivados de la disminución en productividad laboral son del mismo orden de magnitud que los recursos fiscales destinados a financiar el sistema de salud.

Adicionalmente, el monto estimado de 496 mil millones de pesos resulta comparable al presupuesto anual conjunto de instituciones públicas estratégicas, como el Instituto Mexicano del Seguro Social (IMSS) o la Secretaría de Salud a nivel federal, lo que ilustra que las pérdidas asociadas a las ENT no constituyen un fenómeno marginal, sino un componente estructural con efectos sustantivos sobre la sostenibilidad económica y fiscal del país.

Este resultado ofrece una aproximación metodológicamente más rigurosa del impacto económico de las ENT, al fundamentarse en estimaciones del *Average Treatment Effect* (ATE) obtenidas mediante modelos de aprendizaje automático con enfoque causal, que controlan simultáneamente múltiples variables de confusión. A diferencia de los enfoques tradicionales basados en simples diferencias de medias, esta estrategia incorpora heterogeneidad individual y permite estimar de forma más robusta el efecto neto de las ENT sobre la productividad laboral.

En conjunto, estos hallazgos refuerzan la relevancia de integrar herramientas de aprendizaje automático con metodologías de inferencia causal para cuantificar fenómenos complejos como el impacto económico de las ENT. La evidencia generada no solo permite dimensionar con mayor precisión las pérdidas asociadas a estas enfermedades, sino que también proporciona una base empírica sólida para el diseño de estrategias costo-efectivas de prevención y promoción de la salud laboral, cuyos beneficios potenciales podrían traducirse en ganancias económicas comparables a incrementos sustanciales del gasto público en salud.

CAPÍTULO V. Conclusiones y Recomendaciones

Las enfermedades no transmisibles (ENT) constituyen uno de los principales desafíos para la salud pública y el desarrollo económico de los países en vías de desarrollo, tanto por su prevalencia creciente como por su impacto directo en la productividad laboral. La presente investigación abordó este fenómeno de manera integral mediante la aplicación de técnicas estadísticas avanzadas, modelos de aprendizaje automático e inferencia causal, con el objetivo de identificar los determinantes estructurales y conductuales de estas enfermedades y su efecto sobre la productividad en México y en un grupo comparativo de países en desarrollo.

En una primera fase, el análisis se centró en la construcción y evaluación de índices derivados de componentes económicos, alimentarios y de hábitos de vida. A través del análisis de componentes principales (PCA) se logró sintetizar información compleja en dimensiones fundamentales que reflejan la disponibilidad de recursos económicos y la calidad de los hábitos nutricionales y de salud. Este procedimiento permitió no solo facilitar la comparación entre países, sino también establecer un marco de referencia para clasificar su riesgo relativo frente a las ENT. Los resultados evidenciaron que la combinación de estructuras económicas débiles, patrones alimentarios desfavorables y deficiencias en la inversión en salud tiende a incrementar de manera sustancial la carga de enfermedades crónicas en la población. De esta manera, se observó que la interacción entre condiciones materiales de vida y hábitos de salud se configura como un factor explicativo clave de la prevalencia de ENT, dando respuesta a la hipótesis 2 al mostrar la asociación significativa entre las dimensiones económicas, los hábitos de salud y la incidencia de estas enfermedades.

La identificación de tres dimensiones principales: índice económico, índice de alimentación saludable e índice de alimentación no saludable, permitió agrupar a los países en clústeres diferenciados. Este ejercicio mostró que las naciones con menores recursos económicos y mayor consumo de alimentos ultraprocesados presentan tasas de mortalidad superiores, mientras que los países con mejores niveles económicos y con mayor presencia de patrones alimentarios saludables exhiben menores niveles de mortalidad. Estos hallazgos ofrecen evidencia empírica de que la adopción de hábitos positivos, particularmente en la alimentación, ejerce un efecto protector frente a la mortalidad por ENT. Con ello, se aporta evidencia que corrobora la hipótesis 3, al

mostrar que los hábitos de salud positivos inciden favorablemente en la salud y en la capacidad de mantener la productividad laboral.

El análisis de conglomerados complementó esta perspectiva al mostrar que la heterogeneidad entre países en desarrollo es considerable incluso entre aquellos con ingresos nacionales similares. Esta heterogeneidad se reflejó en las diferencias en políticas públicas, acceso a servicios de salud y estrategias preventivas, indicando que la respuesta a las ENT no puede ser uniforme. La evidencia sugiere que los determinantes laborales y los hábitos de vida deben analizarse en conjunto con los factores económicos, ya que la interacción de estas dimensiones genera configuraciones de riesgo diferenciadas. En este sentido, los resultados fortalecen la idea central de que la economía, los hábitos de salud y las condiciones laborales se encuentran estrechamente interrelacionados en la explicación de las ENT.

El análisis también mostró que algunos países en desarrollo se encuentran en una posición más vulnerable no solamente por su nivel económico absoluto, sino por la combinación de malas prácticas alimenticias, bajos niveles de actividad física y alta prevalencia de factores de riesgo como el tabaquismo. Este enfoque integral permitió establecer una base sólida para la interpretación de los resultados en las siguientes fases del estudio y evidenció la utilidad de integrar herramientas estadísticas y de aprendizaje automático en la caracterización del riesgo poblacional.

La investigación incorporó enfoques modernos de análisis cuantitativo, incluyendo modelos de regresión penalizada como *Lasso*, *Ridge* y redes elásticas. Estas técnicas no solo permitieron una selección más refinada de variables relevantes, sino que mejoraron el rendimiento de los modelos frente a técnicas clásicas. Particularmente, las redes elásticas mostraron su eficiencia al reducir en un 38% el número de variables necesarias para explicar el fenómeno, al tiempo que incrementaron el ajuste del modelo en un 12% y disminuyeron el problema de multicolinealidad. Estos resultados validan el uso de estas técnicas en entornos de alta dimensionalidad y correlación entre predictores, como suele ser el caso de los estudios poblacionales de salud.

A partir de este estudio, se identificó que múltiples factores de riesgo están estrechamente asociados con el aumento de la mortalidad por ENT, entre los que destacan el índice de masa corporal promedio elevado, el consumo excesivo de carne procesada, bebidas azucaradas, así como el tabaquismo. Estos elementos configuran un patrón de riesgo claramente nocivo reflejando

estilos de vida poco saludables que, con el tiempo, repercuten negativamente en la esperanza de vida y la calidad de vida de las personas.

Frente a estos factores de riesgo, se evidenció también el efecto protector de algunos elementos dietéticos. La inclusión en la dieta de alimentos ricos en ácidos grasos *omega-3*, como el pescado, así como el consumo de grasas saludables, granos integrales y fibra, se asoció con una reducción significativa de la mortalidad por ENT. Estos hallazgos coinciden con la literatura internacional y confirman la importancia de promover políticas alimentarias que incentiven el acceso y consumo de alimentos saludables. La baja actividad física emergió como un factor determinante, particularmente en su relación con la mortalidad cardiovascular y su interacción con indicadores como el IMC. Aunque con un impacto relativamente moderado, el consumo de alcohol y el tabaquismo también se identificaron como elementos contribuyentes al deterioro de la salud, lo que refuerza la necesidad de intervenciones en salud pública focalizadas.

Asimismo, ciertos factores específicos mostraron una asociación estrecha con enfermedades particulares. Por ejemplo, la contaminación por metano, la exposición a asma ocupacional y otros indicadores relacionados con la calidad del aire demostraron una relación directa con las enfermedades respiratorias crónicas. Del mismo modo, el consumo desproporcionado de bebidas azucaradas y la presencia de niveles elevados de glucosa en sangre fueron más relevantes en el caso de la diabetes *mellitus*, revelando una dimensión alimentaria y metabólica del problema. Estos hallazgos refuerzan la afirmación de que los hábitos positivos en alimentación y actividad física generan beneficios tangibles sobre la salud, apoyando la hipótesis 3 al demostrar que la adopción de prácticas saludables ejerce un efecto protector frente a la enfermedad.

En cuanto a los determinantes económicos, el análisis de los Años de Vida Ajustados por Discapacidad (AVAD o DALYs) brindó una aproximación al impacto que las ENT tienen sobre la productividad laboral. Variables como las largas jornadas laborales se correlacionaron negativamente con enfermedades cardiovasculares y diabetes *mellitus*, mientras que la inversión gubernamental en salud no mostró un impacto significativo en la mejora de los indicadores en estos países. Este resultado contrasta con las evidencias disponibles en países desarrollados donde dichas inversiones suelen generar efectos positivos medibles. Por lo tanto, se destaca la necesidad de repensar no solo los montos invertidos, sino también la calidad, destino y eficiencia de dichas

inversiones. En este contexto, políticas enfocadas en la mejora de hábitos alimenticios, la promoción de la actividad física y la regulación de la carga laboral emergen como estrategias prioritarias para mejorar los resultados en salud y productividad.

Complementariamente, el uso de modelos de SEM proporcionó una perspectiva integradora sobre las relaciones causales entre hábitos de salud, factores biomédicos, consumo, estrés y las ENT en la PEA de México. Los modelos SEM presentaron adecuados índices de ajuste global, respaldando su validez empírica. La construcción de una variable latente de segundo orden denominada “hábitos de salud negativos” permitió comprender cómo diferentes prácticas y condiciones de vida interactúan entre sí para incrementar el riesgo de padecer una ENT.

Los resultados del SEM mostraron que dicha variable latente tiene un efecto causal significativo y positivo sobre la presencia de ENT. Específicamente, se encontró que los componentes con mayor peso dentro de este constructo fueron el perfil individual (edad y nivel educativo), los biomarcadores (como presión arterial sistólica, circunferencia de cintura e IMC), así como el consumo de alcohol y tabaco. Se observó además que mayores ingresos y mayor actividad física vigorosa se asocian con una reducción en la manifestación de hábitos de salud negativos, lo cual señala el rol protector que tienen el nivel socioeconómico y el estilo de vida activo. Otro hallazgo importante fue que, si bien existe una relación entre las ENT y los hábitos de salud, esta no debe considerarse bidireccional de manera automática, ya que el modelo propuesto no fue diseñado para capturar relaciones recursivas, sino efectos unidireccionales causales desde los hábitos hacia la enfermedad.

Dentro de las dimensiones evaluadas, la presión arterial sistólica y la circunferencia de cintura fueron los indicadores más consistentes dentro del grupo de biomarcadores. En el ámbito bioquímico, destacaron los niveles altos de triglicéridos y colesterol LDL como predictores claves. En el componente de estrés, el dormir bien emergió como el principal factor protector, en tanto que la sensación de cansancio y las horas de sueño insuficientes contribuyeron negativamente. En el consumo de sustancias, el alcohol fue más determinante que el tabaco, lo que apunta hacia la necesidad de ajustar el enfoque de las campañas preventivas. Los resultados obtenidos permiten afirmar que las ENT representan un factor que reduce de manera directa la productividad laboral, respondiendo a la hipótesis 1 al mostrar el efecto negativo de las ENT.

En la tercera fase del estudio, se evaluó el impacto directo de las ENT en la productividad laboral utilizando técnicas de inferencia causal mediante DML a partir de datos de la ENIGH 2022. El análisis reveló una disminución promedio de 11.44% en la productividad de los trabajadores con ENT, en comparación con sus pares sin ENT, lo que representa un porcentaje relevante que permite estimar una pérdida económica nacional anual de aproximadamente 496 mil millones de pesos mexicanos.

Un hallazgo clave fue que los individuos con ENT crónicas presentaron una reducción promedio menor que aquellos con ENT no crónicas, posiblemente debido a factores como adherencia al tratamiento, adaptación laboral o apoyo institucional. En contraste, el grupo con ENT sin condiciones crónicas mostró la mayor pérdida de productividad, lo que indica una mayor vulnerabilidad y la urgencia de diseñar intervenciones específicas.

Este análisis permitió distinguir el efecto diferenciado que tienen las condiciones crónicas dentro del conjunto de ENT, revelando que las personas con cronicidad asociada presentan una disminución adicional del 18.75% en su productividad en comparación con quienes padecen ENT sin complicaciones crónicas. Al incorporar este ajuste en un modelo de DML, se obtuvo una reducción promedio de -11.44% pesos por hora trabajada, con una precisión estadística robusta ($p = 9.4 \times 10^{-8}$) y una capacidad predictiva adecuada (RMSE = 0.3838) corroborando la respuesta a la hipótesis 1 con el valor del efecto negativo de estas ENT sobre la productividad laboral de la PEA.

La comparación de modelos predictivos mostró que *Random Forest* ofreció el mejor desempeño en la predicción de productividad ajustada, con un RMSE de 0.4714, superando a *XGBoost*, SVM y otros métodos. Este resultado refuerza la utilidad de enfoques de aprendizaje automático para capturar relaciones no lineales y heterogéneas entre variables de salud y productividad.

Los hallazgos obtenidos permiten dar respuesta integral a la hipótesis 2, consolidando la idea de que la interacción entre condiciones económicas, hábitos de salud y factores laborales constituye la explicación más sólida de la carga de ENT que trasciende los límites del sector salud, afectando directamente la productividad laboral. Esta merma en la productividad repercute en el crecimiento económico, dado que la fuerza laboral constituye un insumo esencial para la generación de bienes y servicios; por lo tanto, una reducción sostenida en la capacidad productiva

se traduce en menores niveles de competitividad, innovación y dinamismo económico a nivel nacional. El hecho de que la pérdida atribuible a ENT sea mayor que todo el valor económico generado por el sector salud en un año revela una paradoja relevante: mientras que se destinan recursos a la atención y tratamiento de estas enfermedades, la economía pierde simultáneamente una cantidad aún mayor en términos de productividad no realizada.

Desde el punto de vista metodológico, el uso de DML asegura que estas estimaciones no se limiten a simples asociaciones, sino que capturen el efecto causal promedio controlando la heterogeneidad de la muestra y las múltiples covariables que influyen en la relación entre salud y productividad. Además, la incorporación del factor de expansión de la ENIGH otorga validez externa a los resultados, ampliando su aplicabilidad a nivel poblacional.

Las implicaciones de estos hallazgos son profundas. En el ámbito de la política pública, subrayan la urgencia de fortalecer las estrategias de prevención, diagnóstico temprano y control de ENT, pues la inversión en salud puede generar retornos económicos significativos al reducir la pérdida de productividad. Para las empresas, los resultados evidencian la necesidad de implementar programas de salud ocupacional y bienestar laboral, que no solo mejoran la calidad de vida de los trabajadores, sino que también reducen costos asociados al presentismo y a la baja productividad. Finalmente, para la investigación científica, este estudio constituye un ejemplo del valor que ofrecen los métodos de aprendizaje automático en la estimación de efectos causales complejos, aportando una base empírica sólida para diseñar intervenciones costo-efectivas.

Esta investigación demuestra que las ENT representan una amenaza multifactorial que afecta simultáneamente la salud pública, el desarrollo económico y la eficiencia laboral. Se requiere un enfoque integral y basado en evidencia que combine estrategias de prevención, diagnóstico temprano, promoción de estilos de vida saludables y uso de tecnologías digitales. Entre ellas, destacan los dispositivos portátiles (*wearables*) como herramientas clave para el monitoreo en tiempo real capaces de detectar desviaciones antes de que se consoliden en condiciones crónicas (Arturo & Roman, 2025). Las futuras investigaciones podrían profundizar en la integración de estos dispositivos con modelos de aprendizaje automático para personalizar estrategias de intervención y maximizar tanto la salud de los trabajadores como la sostenibilidad económica de los países en desarrollo. Este enfoque híbrido, técnico y humano, se vislumbra como el camino más sólido para enfrentar los desafíos que las ENT imponen en el siglo XXI.

CAPÍTULO VI. Contribución Académica, Empresarial y Nacional

6.1 Contribución Académica

El presente estudio representa una contribución académica significativa al campo de la salud pública y la economía laboral al profundizar en la relación entre las ENT y la productividad laboral mediante el uso de herramientas avanzadas como el aprendizaje automático. La investigación ofrece un marco metodológico robusto que combina modelos econométricos tradicionales con técnicas modernas como *Random Forest* y *Elastic Net*, lo que permite identificar de manera precisa las variables más relevantes y establecer patrones complejos de asociación entre las dimensiones económicas, sociales y de salud.

Además, la integración de métodos multivariantes y modelado predictivo refuerza la capacidad de los investigadores para analizar grandes bases de datos, abordando desafíos como la falta de normalidad en las variables y la multicolinealidad. Este enfoque no solo enriquece la literatura existente, sino que también abre nuevas líneas de investigación sobre cómo el uso de dispositivos *wearables* y tecnologías avanzadas puede mejorar la monitorización de la salud y la prevención de ENT.

Finalmente, el estudio se convierte en un referente para futuras investigaciones al centrarse en la PEA, un segmento crítico para las economías de países en desarrollo. Sus hallazgos son relevantes no solo para el análisis académico, sino también para informar políticas públicas y estrategias empresariales, estableciendo un puente entre la teoría y la práctica.

6.2 Contribución Empresarial

El impacto de este estudio en el ámbito empresarial radica en la propuesta de estrategias basadas en evidencia para mitigar los efectos de las ENT en la productividad laboral. Las empresas enfrentan retos significativos derivados del ausentismo y el presentismo asociados con estas enfermedades, lo que resulta en pérdidas económicas y disminución en la eficiencia organizacional. Este trabajo ofrece a las empresas un marco analítico que identifica los principales factores de riesgo y hábitos de salud que afectan el desempeño de sus empleados, permitiéndoles diseñar programas de salud más efectivos.

A partir de la estimación del impacto económico agregado de las ENT sobre la productividad laboral, cercano a 496 mil millones de pesos anuales, es posible reinterpretar la gestión de la salud ocupacional desde una lógica de retorno de inversión, más que como un gasto discrecional. Desde esta perspectiva, la inversión en programas preventivos, sistemas de monitoreo continuo y herramientas digitales de seguimiento de la salud puede entenderse como una forma de mantenimiento preventivo del capital humano, análoga a las inversiones que las empresas realizan para preservar su infraestructura física o tecnológica.

La implementación de tecnologías como dispositivos *wearables*, plataformas de analítica avanzada y modelos predictivos basados en aprendizaje automático ofrece a las empresas la posibilidad de monitorear en tiempo casi real indicadores clave de salud, actividad física y comportamiento, permitiendo la detección temprana de patrones de riesgo. Estos sistemas tienen un costo marginal relativamente bajo cuando se comparan con las pérdidas operativas asociadas al deterioro progresivo de la productividad laboral causado por las ENT. La adopción de este tipo de estrategias preventivas puede traducirse en beneficios tangibles para las organizaciones, tales como la reducción de costos médicos directos e indirectos, una mayor estabilidad en la fuerza laboral, mejoras en el clima organizacional y una mayor capacidad de retención del talento. Asimismo, el uso de datos objetivos y personalizados permite optimizar la asignación de recursos, enfocando las intervenciones en los grupos de mayor riesgo y maximizando así la eficiencia de las inversiones en salud.

En conjunto, los hallazgos de este estudio permiten articular un argumento sólido para el sector empresarial: la inversión en salud ocupacional basada en datos y tecnología no debe concebirse como un gasto adicional, sino como una estrategia de eficiencia operativa y sostenibilidad organizacional, capaz de proteger el valor del capital humano y de mitigar pérdidas económicas sustanciales asociadas a las ENT. De esta manera, el enfoque propuesto traduce la evidencia médica y epidemiológica en un lenguaje directamente aplicable a la toma de decisiones estratégicas dentro de las empresas.

6.3 Contribución Nacional

A nivel nacional, esta investigación aporta evidencia crucial para abordar uno de los mayores retos de salud pública en México: el impacto de las ENT sobre la Población Económicamente Activa (PEA). Los hallazgos muestran que las ENT generan una disminución en

la productividad laboral, lo que a su vez implica repercusiones económicas que trascienden el ámbito del sector salud.

Derivado de esta evidencia, se plantea la necesidad de reforzar políticas públicas orientadas a la prevención y promoción de la salud, con énfasis en la reducción de factores de riesgo como el sedentarismo, el tabaquismo y las malas prácticas alimenticias, todos altamente prevalentes en el país.

Como alternativa, desde la Secretaría de Economía y la Secretaría de Hacienda y Crédito Público podrían explorarse esquemas de incentivos fiscales que reconozcan el gasto en salud preventiva como una inversión estratégica. Actualmente, la Ley del Impuesto sobre la Renta permite la deducción de ciertos gastos médicos y, en el caso de las empresas, el pago de primas de seguros de gastos médicos mayores a favor de sus trabajadores. Sin embargo, este marco podría ampliarse para incluir explícitamente acciones de prevención y promoción de la salud, así como tecnologías de apoyo (por ejemplo, programas de bienestar laboral, servicios preventivos, dispositivos de monitoreo personal). Un esquema de este tipo favorecería que las inversiones en salud sean vistas no solo como un costo, sino como un mecanismo para reducir el impacto económico de las ENT y, con ello, mejorar la competitividad del país.

En este sentido, el uso de tecnologías de monitoreo continuo de salud, como los dispositivos *wearables*, no constituye un hallazgo directo del presente estudio, pero sí representa una propuesta futura para fortalecer las estrategias de prevención a nivel nacional. Estas tecnologías, al permitir la detección temprana de riesgos y fomentar estilos de vida más saludables, podrían contribuir a mejorar los indicadores de salud de la población y, de manera indirecta, a mitigar las pérdidas de productividad asociadas a las ENT.

Finalmente, los resultados de esta investigación ofrecen insumos relevantes para la formulación de políticas públicas que promuevan entornos laborales más saludables y sostenibles. La colaboración entre el sector público y privado podría traducirse en un cambio estructural hacia una economía más resiliente, donde la salud de los trabajadores se reconozca como un activo estratégico para el desarrollo económico y social del país.

CAPÍTULO VII. Limitantes de la Investigación y Futuras Direcciones de la Investigación

Aunque el presente estudio se sustenta en un enfoque empírico riguroso y en el uso de bases de datos ampliamente reconocidas, como la ENSANUT y la ENIGH, es importante reconocer una serie de limitaciones inherentes a su diseño y a las fuentes de información utilizadas. Estas limitaciones no invalidan los resultados obtenidos, pero sí delimitan su alcance interpretativo y abren oportunidades claras para futuras investigaciones.

En primer lugar, la investigación es de naturaleza observacional y se basa principalmente en encuestas de corte transversal. Si bien estas fuentes son representativas a nivel nacional y constituyen insumos fundamentales para el análisis de la población económicamente activa, presentan restricciones relacionadas con la disponibilidad de variables, la periodicidad de levantamiento y la distribución estadística de algunos indicadores. En particular, el diseño transversal de la ENIGH impide observar la evolución temporal de la productividad laboral y del estado de salud en un mismo individuo, lo que limita la identificación de trayectorias dinámicas y de efectos acumulativos de las enfermedades no transmisibles a lo largo del tiempo.

Si bien el estudio incorpora metodologías avanzadas de inferencia causal, como el enfoque de DML, que permiten controlar de manera robusta un amplio conjunto de variables de confusión y aproximar el efecto neto de las ENT sobre la productividad laboral, dichas estimaciones deben interpretarse como efectos causales promedio, condicionados a los supuestos del modelo y a la estructura transversal de los datos. En este sentido, aunque el uso de DML fortalece sustancialmente la validez interna frente a métodos econométricos tradicionales, no sustituye la evidencia que podría derivarse de diseños longitudinales, experimentales o cuasi-experimentales.

Adicionalmente, la heterogeneidad entre las bases de datos internacionales y nacionales utilizadas, así como las diferencias metodológicas en los procesos de recopilación y estandarización de la información, representan un reto adicional, particularmente al realizar comparaciones entre países en desarrollo con contextos institucionales, sanitarios y económicos distintos. A ello se suman las limitaciones derivadas del acceso incompleto a información detallada sobre atención médica, historia clínica individual y trayectorias laborales, variables que podrían enriquecer el análisis de mecanismos causales específicos.

Asimismo, el modelo empírico desarrollado en esta investigación se centra en factores biológicos, conductuales y socioeconómicos observables a nivel individual, tales como biomarcadores, hábitos de consumo, actividad física y condiciones económicas. Sin embargo, la literatura especializada reconoce de manera consistente que el diseño del sistema de trabajo, incluyendo la carga mental y física, la organización de los turnos, la duración de la jornada, las condiciones ergonómicas del puesto y el grado de control sobre el trabajo, constituye un determinante estructural clave en la generación de fatiga, estrés crónico y deterioro progresivo de la salud, los cuales actúan como precursores relevantes de las ENT.

La ausencia de variables organizacionales y psicosociales del trabajo en las bases de datos utilizadas limita la capacidad del modelo para capturar de manera explícita la responsabilidad del sistema productivo en la creación de condiciones que favorecen la aparición y progresión de enfermedades crónicas. En consecuencia, los resultados no deben interpretarse como una atribución exclusiva de la pérdida de productividad a decisiones o comportamientos individuales, sino como una estimación parcial condicionada a la información disponible y al nivel de agregación de los datos.

No obstante, la aplicación de modelos avanzados de aprendizaje automático y técnicas de reducción de dimensionalidad ha demostrado ser una herramienta eficaz para manejar problemas de falta de normalidad, alta colinealidad y complejidad funcional, permitiendo aprovechar de manera óptima la información disponible y extraer evidencia empírica robusta incluso en contextos de datos complejos.

A partir de estas limitaciones, futuras líneas de investigación podrían beneficiarse de la incorporación de datos longitudinales o paneles administrativos que permitan analizar la evolución temporal de la productividad y del estado de salud de los individuos, así como evaluar con mayor precisión los efectos de intervenciones preventivas en salud laboral. De igual forma, resulta fundamental integrar explícitamente variables relacionadas con la organización del trabajo y el entorno laboral, como la carga cognitiva, el estrés laboral percibido, el trabajo por turnos, la intensidad del trabajo y las condiciones ergonómicas, dentro de modelos de aprendizaje automático e inferencia causal, evitando así una visión puramente biomédica del fenómeno.

Finalmente, la combinación de datos administrativos, registros clínicos, información empresarial y datos fiscales, junto con información continua proveniente de dispositivos

wearables, abriría nuevas posibilidades para analizar trayectorias dinámicas de salud y productividad, evaluar intervenciones preventivas en tiempo casi real y estimar con mayor precisión los efectos causales de modificaciones en el diseño del trabajo. De esta manera, futuras investigaciones podrían contribuir al desarrollo de modelos integrales de salud laboral que reconozcan la responsabilidad compartida entre individuos, organizaciones y sistemas productivos.

Referencias

- Aguilar, C.A. (1999). Promoción de la salud para la prevención de las enfermedades crónico-degenerativas vinculadas con la alimentación y el estilo de vida. *Salud Comunitaria y promoción de la salud*. ICEPSS Editores.
- Aguilar, C.A., Mehta, R., Rojas, R., Gómez, F.J., Olaiz, G. & Rull, J.A. (2005). Management of the metabolic syndrome as a strategy for preventing the macrovascular complications of type 2 diabetes: controversial issues. *Curr Diabetes Em. I*(2):145- 158. <https://doi.org/10.2174/1573399054022767>
- Allen, L. N., Wigley, S., & Holmer, H. (2021). Implementation of non-communicable disease policies from 2015 to 2020: a geopolitical analysis of 194 countries. *The Lancet Global Health*, 9(11), e1528-e1538. [https://doi.org/10.1016/S2214-109X\(21\)00359-4](https://doi.org/10.1016/S2214-109X(21)00359-4)
- Al-Qurishi, M., Al-Rakhami, M., Al-Qershi, F., Hassan, M. M., Alamri, A., Khan, H. U., & Xiang, Y. (2015). A framework for cloud-based healthcare services to monitor noncommunicable diseases patient. *International journal of distributed Sensor Networks*, 11(3), 985629. <https://doi.org/10.1155/2015/985629>
- Álvarez Castaño, L. S. (2009). Los determinantes sociales de la salud: más allá de los factores de riesgo. *Revista Gerencia y Políticas de Salud*, 8(17), 69-79. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-70272009000200005
- Alwashmi, M., Mugford, G., Abu-Ashour, W. & Nuccio, M. (2019) A Digital Diabetes Prevention Program (Transform) for Adults with Prediabetes: Secondary Analysis. *JMIR Diabetes*, 4(3). <https://doi.org/10.2196/13904>
- AMIS, Asociación Mexicana de Instituciones de Seguros. (2018). *Aseguramiento en salud: Propuesta del sector asegurador privado*. <https://sitio.amis.com.mx/aseguramiento-en-salud-2/>
- Anderson, L. M., Quinn, T. A., Glanz, K., Ramirez, G., Kahwati, L. C., Johnson, D. B., ... & Katz, D. L. (2019). The effectiveness of worksite nutrition and physical activity interventions for controlling employee overweight and obesity: *A systematic review*. *BMC Public Health*, 19, 8033. <https://doi.org/10.1186/s12889-019-8033-1>
- Arredondo, A., Orozco, E., Alcalde-Rabanal, J., Navarro, J., & Azar, A. (2018). Challenges on the epidemiological and economic burden of diabetes and hypertension in Mexico. *Revista de Saúde Pública*, 52(23), 1–12. <https://doi.org/10.11606/S1518-8787.2018052000293>
- Arturo, DM.S., Rodriguez-Aguilar, R. (2024). Dimensions Related to NCD in Developing Countries During Working Age Using Ridge, Lasso, and Elastic Net Regressions. In *International Conference on Computer Science and Health Engineering* (pp. 225-240). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-67440-2_18
- Arturo, DM.S., Roman, RA. (2025). Leveraging mHealth and Artificial Intelligence for Enhanced Health Indicators, A TwiMV Framework Proposal. In: Kose, U., Alzubi, J. (eds) *IoT Technologies and Wearables for HealthCare*.

ICWH 2024. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 639. Springer, Cham. https://doi.org/10.1007/978-3-031-95571-6_6

Athey, S., & Imbens, G. W. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27), 7353–7360. <https://doi.org/10.1073/pnas.1510489113>

Azalea, I. K. (2022). *Causal-Comparative Research* (ex post facto research).

Baig, M. M., GholamHosseini, H., Moqem, A. A., Mirza, F., & Lindén, M. (2017). A systematic review of wearable patient monitoring systems—current challenges and opportunities for clinical adoption. *Journal of medical systems*, 41(7), 1-9. <https://doi.org/10.1007/s10916-017-0760-1>

Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology and Health*. 13, 623-649. <https://doi.org/10.1080/08870449808407422>

Benjamin, E. J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., ... & American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2019). Heart disease and stroke statistics—2019 update: a report from the American Heart Association. *Circulation*, 139(10), e56-e528. <https://doi.org/10.1161/CIR.0000000000000659>.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16(1), 78–117. <https://doi.org/10.1177/0049124187016001004>

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.

Brassey, J., Herbig, B., Jeffery, B., & Ungerman, D. (2023). *Reframing employee health: Moving beyond burnout to holistic health*. McKinsey Health Institute. Disponible en: <https://www.mckinsey.com/mhi/our-insights/reframing-employee-health-moving-beyond-burnout-to-holistic-health>

Breiman, L. (2001). *Random Forests*. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

Budreviciute, A., Damiani, S., Sabir, D. K., Onder, K., Schuller-Goetzburg, P., Plakys, G., ... & Kodzius, R. (2020). Management and prevention strategies for non-communicable diseases (NCDs) and their risk factors. *Frontiers in public health*, 8, 574111. <https://doi.org/10.3389/fpubh.2020.574111>

CDC, Centro para el control y prevención de enfermedades. (2022) *Diabetes en español. Información básica*. Recuperado el 6 de Octubre de 2022 en <https://www.cdc.gov/diabetes/spanish/basics/getting-tested.html#:~:text=Los%20valores%20de%20az%C3%BAcar%20en,mayores%20indican%20que%20tiene%20diabetes.&text=Esta%20prueba%20mide%20sus%20niveles,un%20l%C3%ADquido%20que%20contiene%20glucosa>.

CENETEC (2021). Promoción, Prevención, Diagnóstico y Tratamiento de la Hipertensión Arterial en Primer Nivel de Atención. *Guía de Práctica Clínica: Evidencias y Recomendaciones*. <http://www.cenetec-difusion.com/CMGPC/GPC-IMSS-076-21/ER.pdf>.

- Chaker, L., Falla, A., van der Lee, S.J., Muka, T., Imo, D., Jaspers, L., Colpani, V., Mendis, S., Chowdhury, R., Bramer, W. M., Pazoki, R., & Franco, O. H. (2015). The global impact of non-communicable diseases on macro-economic productivity: a systematic review. *European Journal of Epidemiology*, 30(5), 357-395. <https://doi.org/10.1007/s10654-015-0026-5>.
- Chaparro Narváez, P. E., & Guerrero, J. (2001). Condiciones de trabajo y salud en conductores de una empresa de transporte público urbano. *Revista de salud pública*, 3(2), 171-187. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0124-00642001000200005.
- Chapman, K. R., Mannino, D. M., Soriano, J. B., Vermeire, P. A., Buist, A. S., Thun, M. J., ... & Beasley, R. (2006). Epidemiology and costs of chronic obstructive pulmonary disease. *European Respiratory Journal*, 27(1), 188-207. <https://doi.org/10.1183/09031936.06.00024505>
- Chapman, Larry S. (2005). The Art of Health promotion: Meta-Evaluation of Worksite Health Promotion Economic Return Studies: 2005 Update. *American Journal of Health promotion*. 26(4). <https://doi.org/10.4278/ajhp.26.4.tahp>.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>
- Chen, Y. H., & Sawan, M. (2021). Trends and challenges of wearable multimodal technologies for stroke risk prediction. *Sensors*, 21(2), 460. <https://doi.org/10.3390/s21020460>.
- Christian, A. K., Daniel, E., & Sanuade, O. A. (2025). The silent epidemic: unravelling NCD risk clusters and socioeconomic determinants in Zambia. *BMC Public Health*, 25(1), 2535. <https://doi.org/10.1186/s12889-025-23769-y>
- Córdova-Villalobos, J.A., Barriguete-Meléndez, J.A., Lara-Esqueda, A., Barquera, S., Rosas-Peralta, M., Hernández-Ávila, M., León-May, M.A. & Aguilar-Salinas, C.A. (2008). Las enfermedades crónicas no transmisibles en México: sinopsis epidemiológica y prevención integral. *Salud Pública de México*, 50 (5): 419-427. Recuperado en 30 de marzo de 2022, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342008000500015&lng=es&tlng=es.
- Comisión Económica para América Latina y el Caribe (CEPAL). (2016). *Productividad y brechas estructurales en México*. Naciones Unidas. Recuperado en 5 de enero de 2026, de <https://repositorio.cepal.org/entities/publication/92f3efab-8fb3-4fc4-80c3-e85b59bb3097>.
- De la Torre Díez, I., Garcia-Zapirain, B., Méndez-Zorrilla, A., & López-Coronado, M. (2016). Monitoring and follow-up of chronic heart failure: a literature review of eHealth applications and systems. *Journal of medical systems*, 40(7), 1-9. <https://doi.org/10.1007/s10916-016-0537-y>.
- Diez Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York academy of sciences*, 1186(1), 125-145. <https://doi.org/10.1111/j.1749-6632.2009.05333.x>

- Ding, D., Lawson, K. D., Kolbe-Alexander, T. L., Finkelstein, E. A., Katzmarzyk, P. T., Van Mechelen, W., & Pratt, M. (2016). The economic burden of physical inactivity: a global analysis of major non-communicable diseases. *The Lancet*, 388(10051), 1311-1324. [https://doi.org/10.1016/S0140-6736\(16\)30383-X](https://doi.org/10.1016/S0140-6736(16)30383-X)
- Domínguez-Miranda, S. A., & Rodríguez-Aguilar, R. (2023a). Health 4.0, prevention, and health promotion in companies: systematic literature review. In *Computer Science and Engineering in Health Services: 6th EAI International Conference Proceedings, COMPSE 2022, Mexico City, July 28, 2022* (p. 217). Springer International Publishing. https://doi.org/10.1007/978-3-031-34750-4_13
- Domínguez-Miranda, S. A., & Rodríguez-Aguilar, R. (2024a). Machine learning models in health prevention and promotion and labor productivity: A co-word analysis. *Iberoamerican Journal of Science Measurement and Communication*, 4(1), 1-16. <https://doi.org/10.47909/ijsmc.85>
- Domínguez-Miranda, S. A., & Rodríguez-Aguilar, R. (2024b). Related factors with NCD in developing countries: economic, diet and risk factors dimensions. *EAI Endorsed Transactions on Pervasive Health and Technology*, 10. <https://doi.org/10.4108/eetpht.10.3499>
- Domínguez-Miranda, S. A., Rodríguez-Aguilar, R., & Velazquez Salazar, M. (2025). Modeling the Relation Between Non-Communicable Diseases and the Health Habits of the Mexican Working Population: A Hybrid Modeling Approach. *Mathematics*, 13(6), 959. <https://doi.org/10.3390/math13060959>
- Domínguez-Miranda, S.A., & Rodríguez-Aguilar, R. (2023b). The Economic Dimensions of the Non-communicable Diseases: A Panel Data Study. In *International Conference on Intelligent Computing & Optimization*, 5(134-144). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-50158-6_14
- Domínguez-Miranda, S.A.; Rodríguez-Aguilar, R.; Velazquez-Salazar, M. (2026). Exploring the Influence of NCD on the Health Habits of the Mexican Working Population. In *Computer Science and Engineering in Health Services, 8th EAI International Conference on Computer Science and Engineering in Health Services, COMPSE 2024*. EAI/Springer Innovations in Communication and Computing. Springer, Cham. ISBN: 978-3-031-94629-5. https://doi.org/10.1007/978-3-031-94630-1_15
- Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford university press.
- Eng, J. Y., Moy, F. M., & Bulgiba, A. (2016). Impact of a Workplace Health Promotion Program on Employees' Blood Pressure in a Public University. *PLoS ONE*, 11(2), e0148307. <https://doi.org/10.1371/journal.pone.0148307> <https://pubmed.ncbi.nlm.nih.gov/2648307/>
- ENIGH (2022). *Encuesta Nacional de Ingresos y Gasto de los Hogares*. Recopilado el 1 de mayo de 2024 en: <https://www.inegi.org.mx/programas/ENIGH/nc/2022/>
- ENSANUT (2022). *Encuesta Nacional de Salud y Nutrición continua 2022*. Recopilado el día 29/09/2023 de <https://ENSANUT.insp.mx/encuestas/ENSANUTcontinua2022/descargas.php>

- Epp J. (1996). Lograr la salud para todos: un marco para la promoción de la Salud. En: Promoción de la Salud: una antología. *Organización Panamericana de la Salud*. (557) 25-36. <https://iris.paho.org/bitstream/handle/10665.2/3318/Promocion%20de%20la%20salud%20una%20antologia.pdf?sequence=1&isAllowed=y>.
- Erazo, L. R., Nisenbaum, I. S., & Gutiérrez, J. N. (2002). *Psicología y salud*. UNAM. https://books.google.com.mx/books?hl=es&lr=&id=xzAZH2_Ohu4C&oi=fnd&pg=PA7&dq=Psicolog%C3%A4+y+salud+en+M%C3%A9xico+reynoso+libro&ots=MjIHP9cvLQ&sig=-boxjq10JB9u3emzbKR2ndYw97s&redir_esc=y#v=onepage&q=Psicolog%C3%A4da%20y%20salud%20en%20M%C3%A9xico%20reynoso%20libro&f=false.
- Evans, J. R. (2021). *Business analytics: methods, models, and decisions*.
- Ezzati, M., & Riboli, E. (2013). Behavioral and dietary risk factors for noncommunicable diseases. *New England Journal of Medicine*, 369(10), 954-964. <https://doi.org/10.1056/NEJMr1203528>
- FAO. (2023). Food balances. Recuperado el 2023/03/29. <https://www.fao.org/faostat/en/#data/FBSH>.
- Feenstra, R. C., Inklaar, R. & Timmer, M.P. (2022), The Next Generation of the Penn World Table. *American Economic Review*, 105(10), 3150-3182. Last accessed 06/01/2023. <https://www.rug.nl/ggdc/productivity/pwt/>.
- Fertman, C. I., & Allensworth, D. D. (2016). *Health promotion programs: from theory to practice*. John Wiley & Sons. https://d1wqtxts1xzl7e.cloudfront.net/79564003/20181220034953health_promotion_program-with-cover-page-v2.pdf?Expires=1662614297&Signature=YSvNSN38uhfrOBTjGxBTX8SXSyKS6cKsa4f4sTuHRVGlvnlTCxbk-8SkV5xp4Fat6Maa13frrtV9XM4skmXa2K4G0AvuGaHD98R~G-QgWhdlwsothllT1qulfr23NafMGwo9YI9ea1SoiWJnK10lBWMpp7xKBCPoqW5GUHECnWaVp9RmBQ~0dFngb8i2FvthQAretyZz1LAXwWnuaqem0lpSTxMlnlkFpG16k4QKyLULIPZnsUosrahpH9TZXi7OY7quxvAoCR9qNVpJWUo4jwscCzYm1zAbCHq2sTJ35e2ZBvJ2VIZKR73osbjQb7kUJdERZUDHQ~PSbs5OQdvu8w&key-Pair-Id=APKAJLOHF5GGSLRBV4ZA.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1), 1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929880/>.
- Garvey, W. & Arathuzik, G. (2016). *Lifestyle Therapy for Diabetes Mellitus*. Mechanick, J., Kushner, R. (eds) *Lifestyle Medicine*. Springer, Cham. https://doi.org/10.1007/978-3-319-24687-1_22.
- Gil-Girbau, M., Pons-Vigués, M., Rubio-Valera, M., Murrugarra, G., Masluk, B., Rodríguez-Martín, B., ... & Berenguera, A. (2021). Modelos teóricos de promoción de la salud en la práctica habitual en atención primaria de salud. *Gaceta sanitaria*, 35(1), 48-59. <https://doi.org/10.1016/j.gaceta.2019.06.011>.
- Goetzel, R. Z., Pei, X., Tabrizi, M., Henke, R. M., Kowlessar, N., Nelson, C. F., & Metz, R. D. (2011). Ten modifiable health risk factors are linked to more than one-fifth of employer-employee health care spending. *Health Affairs*, 30(2), 247–256. <https://doi.org/10.1377/hlthaff.2010.0810>

- Haldar, S.K. & Mallik, G. (2010): Does human capital cause economic growth? A case study of India, *International Journal of Economic Sciences and Applied Research. Kavala Institute of Technology*, 3(1) 7-25. <http://hdl.handle.net/10419/66597>.
- Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and generalizations*. CRC press.
- Herbert, T. B., & Cohen, S. (1993). Depression and immunity: a meta-analytic review. *Psychological bulletin*, 113(3), 472. <https://doi.org/10.1037/0033-2909.113.3.472>.
- Hesterberg, T., Choi, N. H., Meier, L., & Fraley, C. (2008). Least angle and ℓ_1 penalized regression: A review. *Statist. Surv.* 2, 61-93. <https://doi.org/10.1214/08-SS035>.
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons. Recopilado el 1 de mayo de 2024 en: https://www.researchgate.net/profile/Andrew-Cucchiara/publication/261659875_Applied_Logistic_Regression/links/542c7eff0cf277d58e8c811e/Applied-Logistic-Regression.pdf
- Hosseinpoor, A. R., Bergen, N., Kunst, A., Harper, S., Guthold, R., Rekve, D., ... & Chatterji, S. (2012). Socioeconomic inequalities in risk factors for non-communicable diseases in low-income and middle-income countries: results from the World Health Survey. *BMC public Health*, 12(1), 1-13. <https://doi.org/10.1186/1471-2458-12-912>
- Hoyle, R. H. (Ed.). (2012). *Handbook of structural equation modeling*. Guilford Press.
- IDF. (2017). *Diabetes atlas*, 8th edition. Brussels: International Diabetes Federation, Executive Office. https://diabetesatlas.org/upload/resources/previous/files/8/IDF_DA_8e-EN-final.pdf.
- IHME - Institute for Health Metrics and Evaluation (2020). *Global Burden of Disease Study 2019*, Results. <https://vizhub.healthdata.org/gbd-results/>.
- IMSS, Instituto Mexicano del Seguro Social (2022). *Entornos laborales Seguros y Saludables*. Obtenida el 23 de abril de 2024 en: https://www.imss.gob.mx/sites/all/statics/elssa/infografias/Desarrollo_Conceptual_ELSSA.pdf
- INEGI (2020). *Módulo de Práctica Deportiva y Ejercicio Físico (MOPRADEF)*. Resultados noviembre. <https://bit.ly/3pqXbsk>.
- INEGI, Instituto Nacional de Estadística y Geografía (2015). *Cálculo de los índices de productividad laboral y del costo unitario de la mano de obra 2015 / Instituto Nacional de Estadística y Geografía*. 2^a. Ed.-- México: INEGI. ISBN 978-607-739-754-0. Obtenida el 22 de abril de 2024 en: https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825078829.pdf
- Instituto Nacional de Estadística y Geografía (INEGI). (2020). *Cuenta satélite del sector salud de México (CSSSM)*. Anexo 1: PIB ampliado del sector salud [Banco de Indicadores, clave 6207067077]. Recopilado el 29 de marzo de 2022 en: https://www.inegi.org.mx/contenidos/programas/saludsat/2018/datosabiertos/cssalud_anual_csv.zip

- Instituto Nacional de Estadística y Geografía (INEGI). (2023). *Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022: Diseño de muestra y factores de expansión*. INEGI. <https://www.inegi.org.mx/programas/enigh/nc/2022/>
- Jaspers, L., Colpani, V., Chaker, L., van der Lee, S.J., Muka, T., Imo, D., Mendis, S., Chowdhury, R., Bramer, W. M., Falla, A., Pazoki, R., & Franco, O. H. (2015). The global impact of non-communicable diseases on households and impoverishment: a systematic review. *European Journal of Epidemiology*, 30(3), 163-188. <https://doi.org/10.1007/s10654-014-9983-3>.
- Jeet, G., Thakur, J. S., Prinja, S., & Singh, M. (2017). Community based interventions for non-communicable diseases prevention and control in low- and middle-income countries: A systematic review. *PLoS ONE*, 12(8), e0180640. <https://doi.org/10.1371/journal.pone.0180640>
- Jensen, P. H., Jensen, C., & Andersen, L. L. (2024). *Cost-effectiveness of return-to-work interventions for employees with musculoskeletal and psychological disorders: An economic evaluation*. *Journal of Occupational Rehabilitation*. <https://di.org/10.1016/j.jval.2024.06.008>
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Karasek Jr, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative science quarterly*, 285-308. <https://doi.org/10.2307/2392498>
- Kickbusch, I. (2003). The contribution of the World Health Organization to a new public health and health promotion. *American journal of public health*, 93(3), 383-388. <https://doi.org/10.2105/AJPH.93.3.383>
- Kirsten, W. (2010). Making the link between health and productivity at the workplace—A global perspective. *Industrial health*, 48(3), 251-255. <https://doi.org/10.2486/indhealth.48.251>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling (4th ed.)*. The Guilford Press
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications. <https://doi.org/10.25336/csp29418>
- Koopmanschap, M. A., Rutten, F. F., van Ineveld, B. M., & Van Roijen, L. (1995). The friction cost method for measuring indirect costs of disease. *Journal of health economics*, 14(2), 171-189. [https://doi.org/10.1016/0167-6296\(94\)00044-5](https://doi.org/10.1016/0167-6296(94)00044-5).
- Lalonde, M. (1974). *A new perspective on the health of Canadians*. Minister of Supply and Services Canada. <http://www.phac-aspc.gc.ca/ph-sp/pdf/perspect-eng.pdf>.
- Li, S., Wang, J., Zhang, B., Li, X., & Liu, Y. (2019). Diabetes mellitus and cause-specific mortality: a population-based study. *Diabetes & metabolism journal*, 43(3), 319-341. <https://doi.org/10.4093/dmj.2018.0060>.

- Lopez-Sanchez, G. F., López-Bueno, R., Villaseñor-Mora, C., & Pardhan, S. (2022). Comparison of diabetes *mellitus* risk factors in Mexico in 2003 and 2014. *Frontiers in Nutrition*, 9, 894904. <https://doi.org/10.3389/fnut.2022.894904>
- Louppe, G. (2014). Understanding *Random Forests*. *Cornell University Library*, 10. Recopilado el 6 de mayo de 2024 en: https://www.researchgate.net/profile/Gilles-Louppe/publication/264312332_Understanding_Random_Forests_From_Theory_to_Practice/links/54ae38ea0cf2213c5fe427b7/Understanding-Random-Forests-From-Theory-to-Practice.pdf
- Mahmood, S. S., Levy, D., Vasan, R. S., & Wang, T. J. (2014). The Framingham Heart Study and the epidemiology of cardiovascular disease: A historical perspective. *The Lancet*, 383(9921), 999–1008. [https://doi.org/10.1016/S0140-6736\(13\)61752-3](https://doi.org/10.1016/S0140-6736(13)61752-3)
- Mancia, G., Fagard, R., Narkiewicz, K., Redán, J., Zanchetti, A., Böhm, M., ... & Zannad, F. (2013). 2013 Practice guidelines for the management of arterial hypertension of the European Society of Hypertension (ESH) and the European Society of Cardiology (ESC): ESH/ESC Task Force for the Management of Arterial Hypertension. *Journal of hypertension*, 31(10), 1925-1938. <https://doi.org/10.3109/08037051.2013.817814>.
- Mansyur, M. (2021). Occupational Health, Productivity and Evidence-Based Workplace Health Intervention. *Acta Medica Philippina*, 55(6). <https://doi.org/10.47895/amp.v55i6.4273>.
- Marques, A. (2012). *Atividades Física e Saúde. A Randomized da randomized de estilos de vida ativos e saudáveis*. Chiado Editora.
- Marques, A., Almeida, B., Bordado, J., Cortés, P., & Gómez, F. (2018). Determinantes de la Salud: La importancia de promover estilos de vida activos. *Journal of Sport Pedagogy and Research*, 4(3), 65-73. https://www.researchgate.net/profile/Francisco-Gomez-51/publication/336073979_Determinantes_de_la_Salud_La_importancia_de_promover_estilos_de_vida_activos/links/5e8b43e24585150839c4f7ca/Determinantes-de-la-Salud-La-importancia-de-promover-estilos-de-vida-activos.pdf.
- Marqués, F (2020). Los programas de salud y bienestar en el trabajo como palanca para la sostenibilidad empresarial. *Boletín Económico De ICE*, (3119) <https://doi.org/10.32796/bice.2020.3119.6955>
- Mayne, T. J., Howard, K., & Brandt-Rauf, P. W. (2004a). Measuring and evaluating the effects of disease on workplace productivity. *Journal of Occupational and Environmental Medicine*, 46(6), S1-S2. <https://doi.org/10.1097/01.jom.0000126688.57275.6d>
- Mayne, T., Beresford, T. P., & Overton, J. (2004b). Absenteeism and presenteeism: Health-related productivity losses associated with depression. *Journal of Occupational and Environmental Medicine*, 46(6), 606–611. <https://doi.org/10.1097/01.jom.0000128174.44883.04>
- Méndez-Hernández, P., Dosamantes-Carrasco, D., Siani, C., Flores, Y. N., Arredondo, A., Lumbreras-Delgado, I., Granados-García, V. M., Denova-Gutiérrez, E., Gallegos-Carrillo, K., & Salmerón, J. (2012). A workplace physical activity program at a public university in Mexico can reduce medical costs associated with type 2

diabetes and hypertension. *Salud Pública de México*, 54(1), 20–27.
[researchgate.net+2scielosp.org+2scielo.org.mx+2](https://researchgate.net/2scielosp.org+2scielo.org.mx+2)

NICE. (2022). Clinical guideline. Hypertension in adults: diagnosis and management: National Institute for Health and Care Excellence (NICE). <https://www.nice.org.uk/guidance/ng136/resources/hypertension-in-adults-diagnosis-and-management-pdf-66141722710213>

Norma Oficial Mexicana. (2022). *Norma Oficial Mexicana NOM-035-STPS-2018, Factores de riesgo psicosocial en el trabajo-Identificación, análisis y prevención*. Recuperado el 12 de octubre de 2022 de <https://www.gob.mx/stps/articulos/norma-oficial-mexicana-nom-035-stps-2018-factores-de-riesgo-psicosocial-en-el-trabajo-identificacion-analisis-y-prevencion#:~:text=La%20NOM%20035%20tiene%20como,en%20los%20centros%20de%20trabajo>.

OCDE. (2019). *Health at Glance 2018: OECD indicators*. <https://www.oecd.org/mexico/health-at-a-glance-mexico-ES.pdf>

OCDE. (2023). *Working age population (indicator)*. doi:10.1787/d339918b-en. <https://data.oecd.org/pop/working-age-population.htm>

OMS (2022). *Actividad física*. Obtenido el 22 de abril de 2024 en: <https://www.who.int/es/news-room/fact-sheets/detail/physical-activity>

OMS. (1986). *Carta de Ottawa para la Promoción de la Salud*. Recuperado el 17 de abril del 2024 en: https://d1wqtxts1xzle7.cloudfront.net/50565360/Carta_de_Ottawa-libre.pdf?1480218969=&response-content-disposition=inline%3B+filename%3DCarta_de_Ottawa_Para_la_Promocion_de_la.pdf&Expires=1713372484&Signature=baCD9H7M3ZYbd11sy9z-DQ9nJwQCdUyCg0KuGdZPC8xzNY-o7T3csEAMBdppPE8Z~JQM0TD-ONXWvMm3QbdbwDm~vu-oZC5V2KRL5py-1JEFNS2oxfGM1kz1Kntlo-cXykE6Db43fKz1dpkxOFVEk3yvvDOYSG1DVsfEMAWeTCjbd0oi1wObtU5I5zyf~s7xbfpDGf7fL33chFgEoO3Knyt41amLnAFX58zZuUTgZQBjPlgeqzN64yPw5yALDjGvAlPjc9kuKFBSPkEP9DxQo5wanwI2Fv2AGNS2Ep1M2bCltilxKAbyMW9Gt35dEVquaC66yVmXj0gDHbLfnDjDg_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

OPS, Organización Panamericana de la Salud (2023). *Economics of NCDs*. Recuperado el 30 de septiembre de 2025. <https://www.paho.org/en/topics/economics-ncds>

OPS, Organización Panamericana de la salud. (2021). *La carga de las enfermedades respiratorias crónicas en la Región de las Américas, 2000-2019*. Organización Panamericana de la Salud. [https://www.paho.org/es/enlace/carga-enfermedades-respiratorias-cronicas#:~:text=En%20toda%20la%20regi%C3%B3n%20en,a%C3%B1os%20\(47%25\)%20en%20mujeres](https://www.paho.org/es/enlace/carga-enfermedades-respiratorias-cronicas#:~:text=En%20toda%20la%20regi%C3%B3n%20en,a%C3%B1os%20(47%25)%20en%20mujeres)

OPS, Organización Panamericana de la Salud. (2022). *Curso de vida saludable*. Recuperado el día 6 de octubre de 2022 de <https://www.paho.org/es/temas/curso-vida-saludable>

- Parati, G., Torlasco, C., Omboni, S., & Pellegrini, D. (2017). Smartphone applications for hypertension management: a potential game-changer that needs more control. *Current hypertension reports*, 19(6), 1-9. <https://doi.org/10.1007/s11906-017-0743-0>
- Patel, R., & Saini, A. (2024). Evaluating the return on investment (ROI) of mental health programs in corporate environments: A multicenter study. *Journal of Mental Health Issues and Behavior*, 46(1), 11–18. <https://doi.org/10.55529/jmhib.46.1.11>
- Pfeffer, F., Kaufer-Horwitz, M. & Rodríguez, M. (2015). Nutrición en el curso de la vida. *Nutrición del adulto*. Editorial Médica Panamericana. 369-397 <https://www.berri.es/pdf/NUTRIOLOG%C3%8DA%20M%C3%89DICA/9786079356415>
- Pineros, M., Znaor, A., Mery, L., & Bray, F. (2017). A global cancer surveillance framework within noncommunicable disease surveillance: making the case for population-based cancer registries. *Epidemiologic reviews*, 39(1), 161-169. <https://doi.org/10.1093/epirev/mxx003>
- Proper, K. I., & van Oostrom, S. H. (2023). The effectiveness of workplace health promotion interventions on productivity outcomes: A systematic review of systematic reviews. *Systematic Reviews*, 12(1), 45. <https://doi.org/10.1186/s13643-023-02258-6>
- Puyol, Á. (2012). Ética, equidad y determinantes sociales de la salud. *Gaceta sanitaria*, 26, 178-181. <https://doi.org/10.1016/j.gaceta.2011.08.007>
- Ramírez, R., León, R., Gutiérrez, T., Mier, N., Banda, O., & Delabra, M. (2023). Effectiveness of an online Chronic Disease Self-Management Program (CDSMP) on physical activity and depressive symptoms in people with non-communicable diseases in Mexico and Peru. *Healthcare*, 11(8), 1168. <https://doi.org/10.3390/healthcare11081168>
- Reynoso, L. y Seligson, I (2002). *Psicología y salud en México*. UNAM/CONACYT.
- Ritchie, H., Roser, M., & Spooner, F. (2023). *Burden of disease*. *Our World in Data*. <https://ourworldindata.org/burden-of-disease>
- Rongen, A., Robroek, S. J., van Lenthe, F. J., & Burdorf, A. (2013). Workplace health promotion: a meta-analysis of effectiveness. *American journal of preventive medicine*, 44(4), 406-415. <https://doi.org/10.1016/j.amepre.2012.12.007>
- Rosado, I. S. M., Párraga, M. J. C., & García, M. C. E. (2018). Factores biológicos, psicológicos y sociales que afectan la salud mental. *Revista Caribeña de Ciencias Sociales*. <https://www.eumed.net/rev/caribe/2018/03/factores-salud-mental.html>
- Rose, S., Paul, C., Boyes, A., Kelly, B., & Roach, D. (2017). Stigma-related experiences in non-communicable respiratory diseases: a systematic review. *Chronic respiratory disease*, 14(3), 199-216. <https://doi.org/10.1177/1479972316680847>

- Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health education monographs*, 2(4), 328-335.
<https://doi.org/10.1177/109019817400200403>
- Rozjabek, H., Fastenau, J., LaPrade, A., & Sternbach, N. (2020). Adult obesity and health-related quality of life, patient activation, work productivity, and weight loss behaviors in the United States. *Diabetes, metabolic syndrome, and obesity: targets and therapy*, 13, 2049. <https://doi.org/10.2147%2FDMSO.S245486>
- Saha, S. (2013). Impact of health on productivity growth in India. *Inter J Eco, Finance & Manag*, 2(4).
https://www.ejournalofbusiness.org/archive/vol2no4/vol2no4_6.pdf
- Sánchez-Torres, D. A. (2017). Accesibilidad a los servicios de salud: debate teórico sobre determinantes e implicaciones en la política pública de salud. *Revista Médica del Instituto Mexicano del Seguro Social*, 55(1), 82-89.
<https://www.redalyc.org/journal/4577/457749297021/457749297021.pdf>
- Schenker, J. D., & Rumrill Jr, P. D. (2004). Causal-comparative research designs. *Journal of vocational rehabilitation*, 21(3), 117-121.
- Schreiber-Gregory, D. N. (2018). Ridge Regression and multicollinearity: An in-depth review. *Model Assisted Statistics and Applications*, 13(4), 359-365. <https://doi.org/10.3233/MAS-180446>.
- Schultz, T. P. (1997). Assessing the productive benefits of nutrition and health: An integrated human capital approach. *Journal of Econometrics*, 77(1), 141-158. [https://doi.org/10.1016/S0304-4076\(96\)01810-6](https://doi.org/10.1016/S0304-4076(96)01810-6)
- Schumacker, R. E., & Lomax, R. G. (2016). A beginner's guide to structural equation modeling (4ta ed.). Routledge.
- Scichilone, N., Benfante, A., Bocchino, M., Braidò, F., Paggiaro, P., Papi, A., ... & Sanduzzi, A. (2015). Which factors affect the choice of the inhaler in chronic obstructive respiratory diseases? *Pulmonary Pharmacology & Therapeutics*, 31, 63-67. <https://doi.org/10.1016/j.pupt.2015.02.006>
- Shamah-Levy, T., Vielma-Orozco, E., Heredia-Hernández, O., Romero-Martínez, M., Mojica-Cuevas, J., Cuevas-Nasu, L., Santaella-Castell, J.A. & Rivera-Dommarco, J. (2020). *Encuesta Nacional de Salud y Nutrición 2018-19: Resultados Nacionales*. Instituto Nacional de Salud Pública.
- Shikdar, A. A., & Sawaqed, N. M. (2003). Worker productivity, and occupational health and safety issues in selected industries. *Computers & industrial engineering*, 45(4), 563-572. [https://doi.org/10.1016/S0360-8352\(03\)00074-3](https://doi.org/10.1016/S0360-8352(03)00074-3)
- Shlens, J. (2014). A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*.
<https://doi.org/10.48550/arXiv.1404.1100>
- Siddique, H. M. A., Mohey-ud-din, G., & Kiani, A. (2020). Human health and worker productivity: evidence from middle-income countries. *International Journal of Innovation, Creativity and Change*, 14(11), 523-544. Available at SSRN: <https://ssrn.com/abstract=3748998>

- Siegrist, J. (2010). Effort-reward imbalance at work and cardiovascular diseases. *International journal of occupational medicine and environmental health*, 23(3), 279. <https://doi.org/10.2478/v10001-010-0013-8>
- Sistema de información de la Secretaría de Salud – DGIS. <http://sinaiscap.salud.gob.mx:8080/DGIS/#>
- Sistema epidemiológico de México (2023). Anuario de Morbilidad 1984-2022. Recopilado el 22 de abril de 2024 en: <https://epidemiologia.salud.gob.mx/anuario/html/glosario.html#:~:text=Es%20la%20presencia%20constante%20o,de%20un%20%C3%A1rea%20geogr%C3%A1fica%20determinada.&text=Es%20una%20alteraci%C3%B3n%20del%20estado,por%20un%20agente%20etiolo%C3%B3gico%20espec%C3%ADfico.>
- Song, Z., & Baicker, K. (2019). Effect of a workplace wellness program on employee health and economic outcomes: A randomized clinical trial. *JAMA*, 321(15), 1491–1501. <https://doi.org/10.1001/jama.2019.0258>
- SS, Secretaría de Salud. (2013a). *Estrategia nacional para la prevención y el control de sobrepeso, la obesidad y la diabetes*. https://www.gob.mx/cms/uploads/attachment/file/19109/estrategia_sobrepeso_diabetes_obesidad.pdf
- SS, Secretaría de Salud. (2013b). *Prevención y control de la obesidad y riesgo cardiovascular 2013-2018*. https://www.gob.mx/cms/uploads/attachment/file/37677/PAE_PrevencionControlObesidadRiesgoCardiovascular2013_2018.pdf
- Sunny, J. S., Patro, C. P. K., Karnani, K., Pingle, S. C., Lin, F., Anekoji, M., ... & Ashili, S. (2022). Anomaly detection framework for wearables data: a perspective review on data concepts, data analysis algorithms and prospects. *Sensors*, 22(3), 756. <https://doi.org/10.3390/s22030756>
- Svedberg, P., Bardage, C., Sandin, S., & L. Pedersen, N. (2006). A prospective study of health, life-style and psychosocial predictors of self-rated health. *European journal of epidemiology*, 21, 767-776. <https://doi.org/10.1007/s10654-006-9064-3>
- Tang, K. (2015). Estimating productivity costs in health economic evaluations: a review of instruments and psychometric evidence. *Pharmacoeconomics*, 33(1), 31-48. <https://doi.org/10.1007/s40273-014-0209-z>
- Tompa, E., Dolinschi, R., & De Oliveira, C. (2006). Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *Journal of Occupational Rehabilitation*, 16(3), 367-392. <https://doi.org/10.1007/s10926-006-9035-2>
- Wang, Y., & Wang, J. (2020). Modelling and prediction of global non-communicable diseases. *BMC public health*, 20(1-13). <https://doi.org/10.1186/s12889-020-08890-4>
- WBOD, World Bank Open Data. (2023). *Global Development Data*. Last accessed 06/01/2023. <https://data.worldbank.org/>
- WHO, World Health Organization (2018). "Healthy Diet." Disponible en: <https://www.who.int/news-room/factsheets/detail/healthy-diet>

- WHO, World Health Organization. (2013). *Plan of Action for the Prevention and Control of Noncommunicable Diseases in the Americas 2013-2019*. <https://www.paho.org/hq/dmdocuments/2015/plan-accion-prevencion-control-ent-americas.pdf>
- WHO, World Health Organization. (2020). *The top 10 causes of death*. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- WHO, World Health Organization. (2021a). *Non Communicable diseases*. <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases#:~:text=Cardiovascular%20diseases%20account%20for%20most,of%20all%20premature%20NCD%20deaths.>
- WHO, World Health Organization. (2021b). *World Health Statistics 2021*. <https://apps.who.int/iris/rest/bitstreams/1356343/retrieve>
- WHO, World Health Organization. (2021c). *Cardiovascular diseases (CVDs)*. [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- WHO, World Health Organization. (2021d). *Declaración política de la tercera reunión de alto nivel de la Asamblea General sobre la prevención y el control de las enfermedades no transmisibles: informe del director general* (No. EB148/7). https://apps.who.int/iris/bitstream/handle/10665/359406/B148_7-sp.pdf?sequence=1
- WHO, World Health Organization. (2022a). *Chronic obstructive pulmonary disease (COPD)*. [https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-\(copd\)](https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd))
- WHO, World Health Organization. (2022b). *Diabetes* https://www.who.int/health-topics/diabetes#tab=tab_1
- WHO, World Health Organization. (2022c). *Cancer*. https://www.who.int/health-topics/cancer#tab=tab_1
- WHO, World Health Organization. (2022d). *Cancer. Datos y cifras*. <https://www.who.int/es/news-room/fact-sheets/detail/cancer>
- WHO, World Health Organization. (2023). *THE GLOBAL HEALTH OBSERVATORY*. Recuperado el 1 de noviembre de 2023. [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/mean-IMC-\(kg-m-\)-\(age-standardized-estimate\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/mean-IMC-(kg-m-)-(age-standardized-estimate))
- Williams, J., Allen, L., Wickramasinghe, K., Mikkelsen, B., Roberts, N., & Townsend, N. (2018). A systematic review of associations between non-communicable diseases and socioeconomic status within low-and lower-middle-income countries. *Journal of global health*, 8(2). <https://doi.org/10.7189%2Fjogh.08.020409>
- Wipfli, H., Zacharias, K. D., Hundal, N. R., Reynales Shigematsu, L. M., Bahl, D., Arora, M., Bassi, S., & Kumar, S. (2018). Workplace wellness programming in low- and middle-income countries: A qualitative study of corporate key informants in Mexico and India. *Globalization and Health*, 14(55). <https://doi.org/10.1186/s12992-018-0362-9>

- Wong, S. S., Yu, M., Loprinzi, P. D., & Hongu, N. (2014). Smart applications to track and record PA: implications for obesity treatment. *Smart Homecare Technology and TeleHealth*, 2, 77. <http://dx.doi.org/10.2147/SHTT.S41484>
- World data (2023). *Developing countries*. Recopilado el 06/23/2023. <https://www.worlddata.info/developing-countries.php>
- World Health Organization. (2022). *Noncommunicable Diseases Progress Monitor 2022*. Ginebra: Organización Mundial de la Salud.
- Zhang, H., Jiang, Y., Nguyen, H. D., Poo, D. C. C., & Wang, W. (2017). The effect of a smartphone-based coronary heart disease prevention (SBCHDP) programme on awareness and knowledge of CHD, stress, and cardiac-related lifestyle behaviours among the working population in Singapore: a pilot randomized controlled trial. *Health and quality of life outcomes*, 15(1), 1-13. <https://doi.org/10.1186/s12955-017-0623-y>
- Zhang, W., Bansback, N., & Anis, A. H. (2011). *Measuring and valuing productivity loss due to poor health: A critical review*. *Social Science & Medicine*, 72(2), 185–192. <https://doi.org/10.1016/j.socscimed.2010.10.026>
- Zhao, Y., Atun, R., Oldenburg, B., McPake, B., Tang, S., Mercer, S. W., ... & Lee, J. T. (2020). Physical multimorbidity, health service use, and catastrophic health expenditure by socioeconomic groups in China: an analysis of population-based panel data. *The Lancet Global Health*, 8(6), e840-e849. [https://doi.org/10.1016/S2214-109X\(20\)30127-3](https://doi.org/10.1016/S2214-109X(20)30127-3)
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2), 301-320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

ANEXOS

ANEXO I. Productos de Investigación derivado de la Tesis Doctoral

1. Domínguez-Miranda, S. A., & Rodríguez-Aguilar, R. (2023). Health 4.0, prevention, and health promotion in companies: a systematic review. In *Computer Science and Engineering in Health Services: 6th EAI International Conference Proceedings, COMPSE 2022, Mexico City, July 28, 2022* (p. 217). Springer Nature. https://doi.org/10.1007/978-3-031-34750-4_13
2. Domínguez-Miranda, S. A., & Rodríguez-Aguilar, R. (2023). The Economic Dimensions of Non-communicable Diseases: A Panel Data Study. In *International Conference on Intelligent Computing & Optimization* (pp. 134-144). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-50158-6_14
3. Domínguez Miranda, S. A., & Rodríguez Aguilar, R. (2024). Machine learning models in health prevention and promotion and labor productivity: A co-word analysis. *Iberoamerican Journal of Science Measurement and Communication*, 4(1), 1-16. <https://doi.org/10.47909/ijsmc.85>
4. Arturo, D. M. S., & Rodriguez-Aguilar, R. (2024). Dimensions Related to NCD in Developing Countries During Working Age Using Ridge, Lasso, and Elastic Net Regressions. In *International Conference on Computer Science and Engineering in Health Engineering* (pp. 225-240). EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-031-67440-2_18
5. Dominguez-Miranda, S. A., & Rodriguez-Aguilar, R. (2024). Related factors with NCD in developing countries: economic, diet, and risk factors dimensions. *EAI Endorsed Transactions on Pervasive Health and Technology*, 10. <https://doi.org/10.4108/eetpht.10.3499>
6. Domínguez-Miranda, S.A.; Rodríguez-Aguilar, R.; Velazquez Salazar, M. (2025) Modeling the Relation Between Non-Communicable Diseases and the Health Habits of the Mexican Working Population: A Hybrid Modeling Approach. *Mathematics* 13(959). <https://doi.org/10.3390/math13060959>
7. Arturo, DM.S., Roman, RA. (2025). Leveraging mHealth and Artificial Intelligence for Enhanced Health Indicators, A TwiMV Framework Proposal. In: Kose, U., Alzubi, J. (eds)

IoT Technologies and Wearables for HealthCare. ICWH 2024. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 639. Springer, Cham. https://doi.org/10.1007/978-3-031-95571-6_6

8. Domínguez-Miranda, S.A.; Rodríguez-Aguilar, R.; Velazquez-Salazar, M. (2026). Exploring the Influence of NCD on the Health Habits of the Mexican Working Population. In *Computer Science and Engineering in Health Services, 8th EAI International Conference on Computer Science and Engineering in Health Services, COMPSE 2024*. EAI/Springer Innovations in Communication and Computing. Springer, Cham. ISBN: 978-3-031-94629-5. Manuscript accepted for publication. https://doi.org/10.1007/978-3-031-94630-1_15.

Ponencias en Congresos

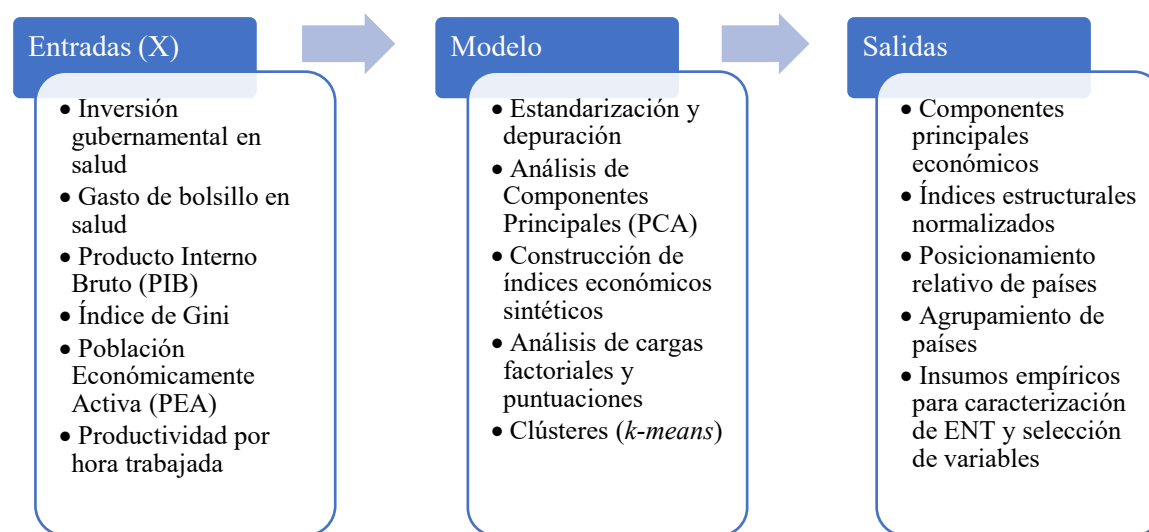
1. **Presentation of research paper:** “Health 4.0, prevention and promotion within companies” during the EAI COMPSE 2022 international conference on October 6, 2022 (Virtual mode).
2. **Presentation of research paper:** “The economic dimensions of noncommunicable diseases: a panel data study” during the International Conference on Intelligent Computing and Optimization, Thailand, in April 2023.
3. **Presentation of research paper:** “Dimension of NCD in developing countries during working age using Ridge, Lasso, and elastic network regressions” during the EAI COMPSE 2023 international conference on November 16, 2023, at UNAM, México.
4. **Presentation of research paper:** “Exploring the Influence of NCD on the Health Habits of the Mexican Working Population” during the EAI COMPSE 2024 international conference on September 19, 2024, in Laredo, Texas.
5. **Presentation of research paper:** “Leveraging mHealth and Artificial Intelligence for Enhanced Health Indicators, A TwiMV Framework Proposal” during the 5th EAI international conference on wearables in healthcare on December 2nd, 2024, in Isparta, Turkey.

ANEXO II. Diagramas de inferencia

Con el fin de facilitar la lectura e interpretación de los resultados, las Tablas 27-32 presentan diagramas de influencia que sintetizan la relación entre las variables analizadas y las técnicas metodológicas empleadas en cada etapa del estudio. Siguiendo el enfoque propuesto por Evans (2021), el diagrama permite identificar con claridad qué conjuntos de variables actúan como determinantes, cuáles constituyen variables intermedias o latentes y cuáles corresponden a los resultados de interés. En particular, se visualiza cómo los factores estructurales de salud, alimentación y economía alimentan los modelos exploratorios y de selección de variables (PCA, *Ridge*, *Lasso*, *redes elásticas* y clústeres), cómo los hábitos de salud y variables clínicas se integran en el modelado causal mediante ecuaciones estructurales, y cómo la condición de ENT opera como variable de tratamiento en la estimación del impacto causal sobre la productividad laboral a través del enfoque de *Double Machine Learning*. Esta representación gráfica complementa la operacionalización detallada de variables y aporta claridad sobre el rol específico que cada conjunto de variables desempeña dentro del diseño metodológico del estudio.

Tabla 27

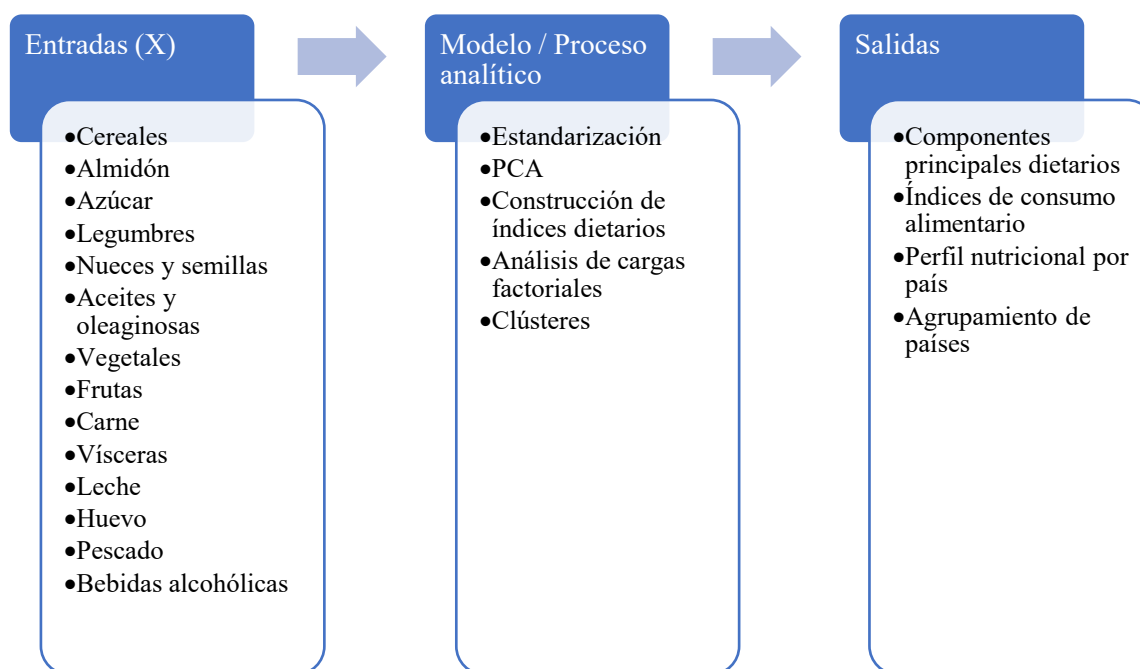
Diagrama de Inferencia I. Factores estructurales asociados a ENT – Dimensión económica (PCA)



Fuentes: OCDE, 2023; WBOD, 2023; Feenstra et al, 2022

Tabla 28

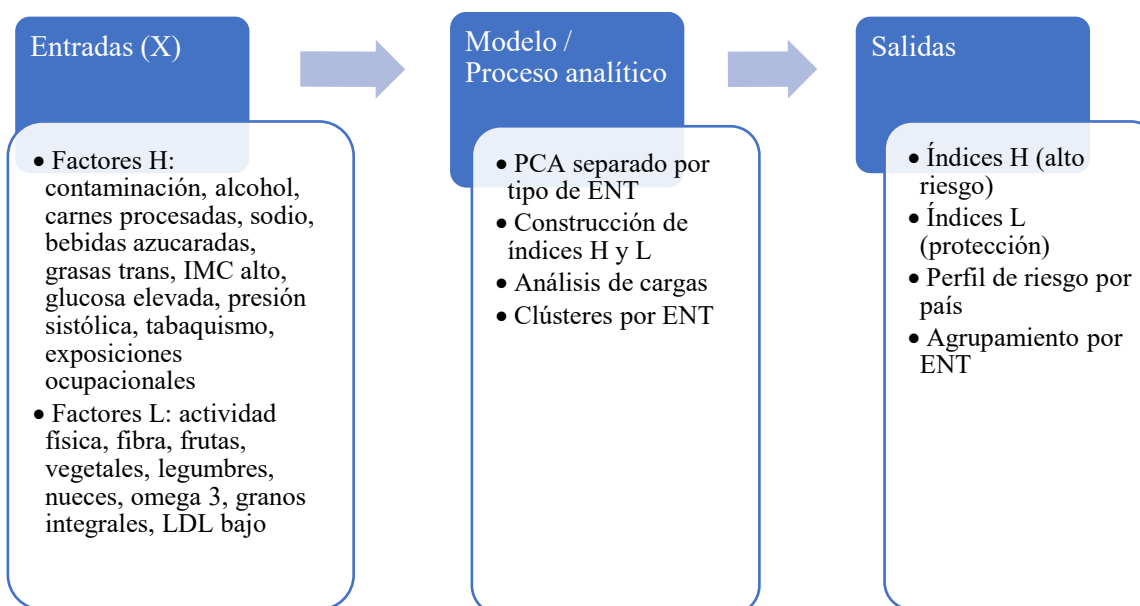
Diagrama de Inferencia 2. Patrones de consumo alimentario asociados a ENT (PCA)



Fuente: FAO, 2023

Tabla 29

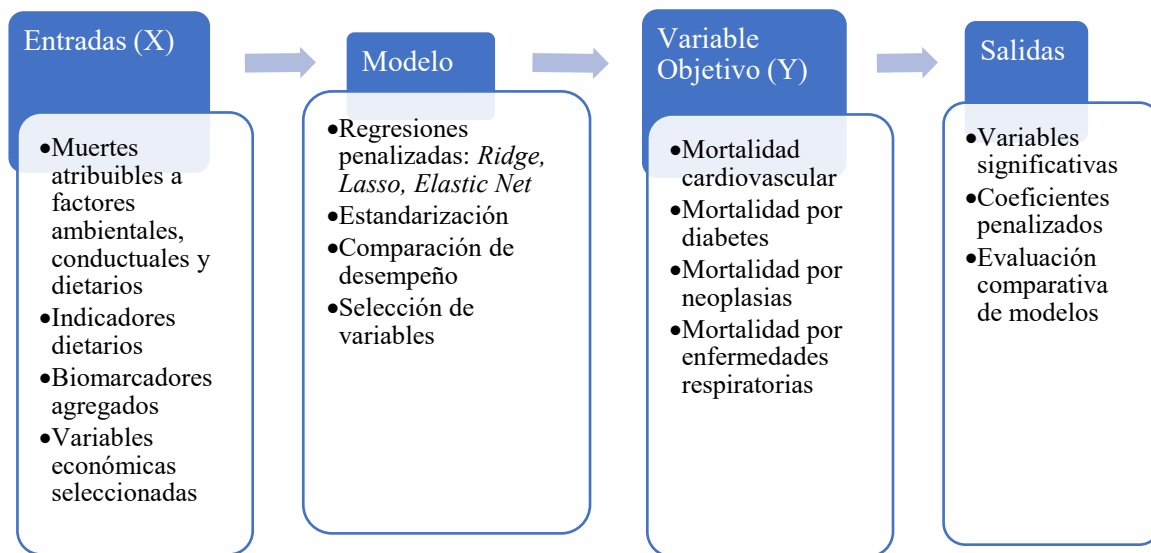
Diagrama de Inferencia 3. Factores de riesgo H y L asociados a ENT (PCA por enfermedad)



Fuente: WHO, 2023; IHME, 2020

Tabla 30

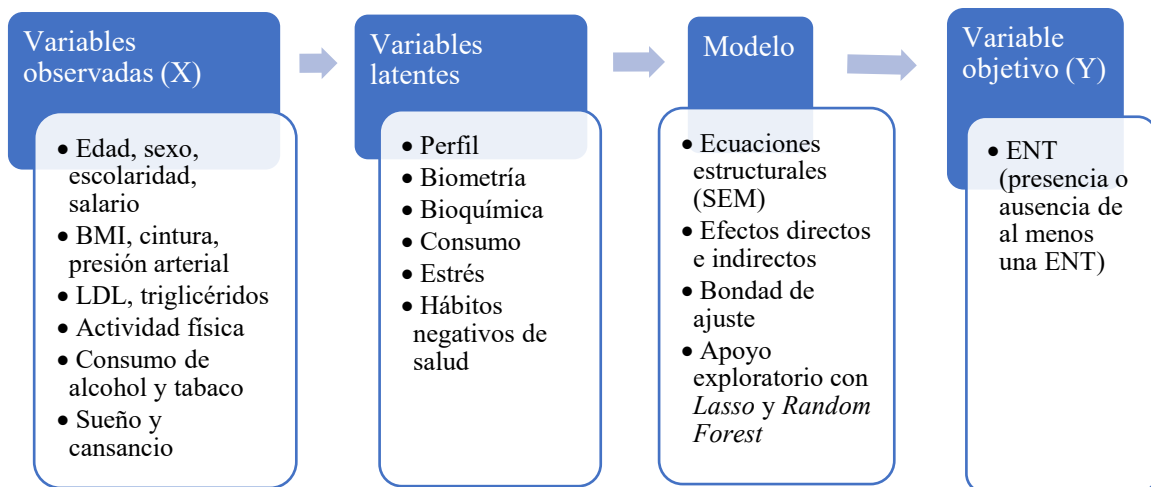
Diagrama de Inferencia 4. Efectos de factores de riesgo sobre mortalidad por ENT



Fuente: WHO, 2023; IHME, 2020

Tabla 31

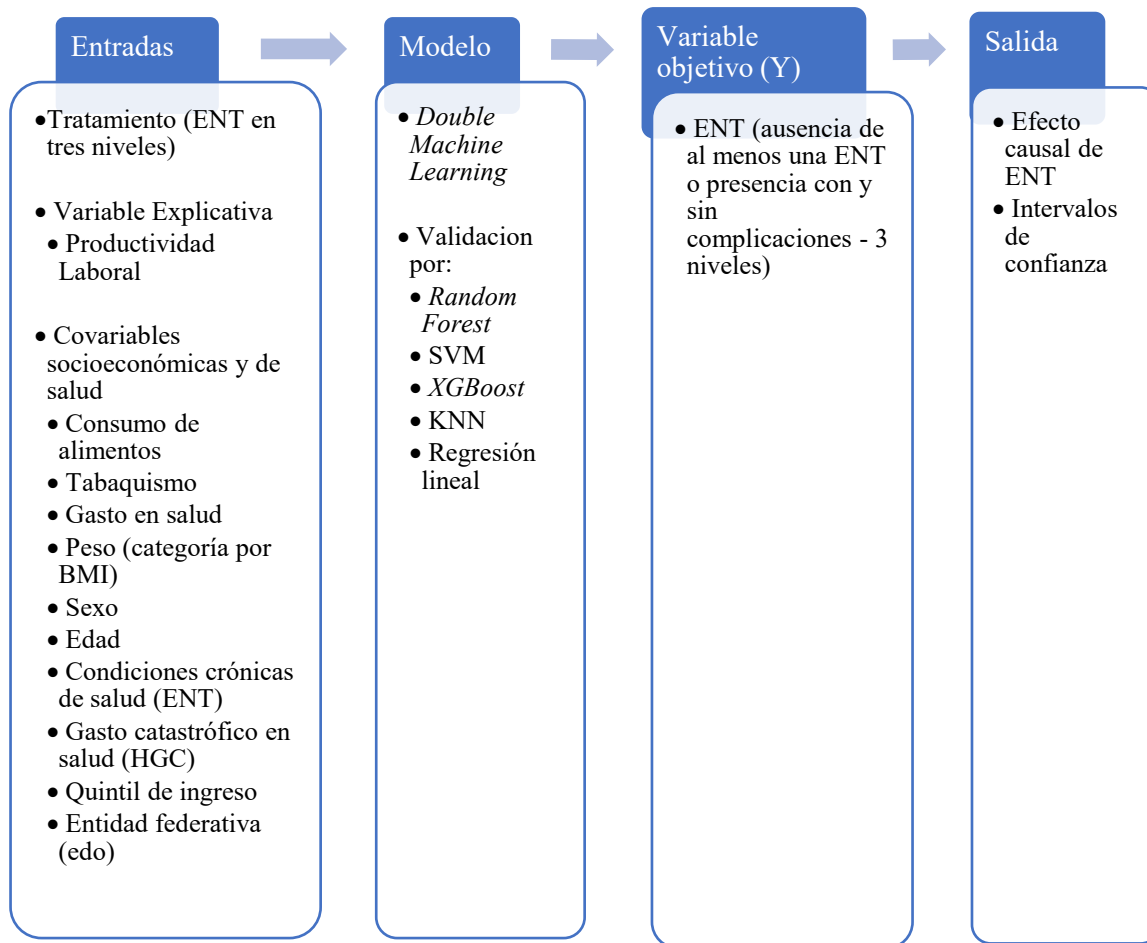
Diagrama de Inferencia 5. Hábitos de salud y ENT en la PEA de México (SEM)



Fuente: ENSANUT, 2022

Tabla 32

Diagrama de Inferencia 6. Impacto causal de las ENT en la productividad laboral (DML)



Fuente: ENIGH, 2022