INSTITUTO MEXICANO DEL SEGURO SOCIAL

CENTRO MÉDICO DE OCCIDENTE “LIC. IGNACIO GARCÍA TÉLLEZ”

DIVISIÓN DE GENÉTICA



PROTOCOLO PARA OBTENCIÓN DEL GRADO DE MAESTRÍA EN CIENCIAS EN BIOINGENIERÍA Y CÓMPUTO INTELIGENTE

ONCOLOGÍA DE PRECISIÓN PARA EL TRATAMIENTO DE CÁNCER DE MAMA

PRESENTA:

Ing. Eduardo Ruiz Robles

INVESTIGADOR RESPONSABLE:

Dra. Martha Patricia Gallegos Arreola

Guadalajara, Jalisco, Enero de 2025

**IDENTIFICACIÓN DE LOS INVESTIGADORES**

**INVESTIGADOR RESPONSABLE**

Dra. Martha Patricia Gallegos Arreola

Investigador Titular B, Adscrita a la División de Genética del

Centro de Investigación Biomédica de Occidente (CIBO). Instituto Mexicano del

Seguro Social (IMSS).

Teléfono:

Correo electrónico:

**INVESTIGADORES ASOCIADOS**

Dr. Germán Andrés Preciat González

Principal Investigador

Módulo M, Centro Universitario de Ciencias Exactas e Ingenierías

Blvd.Marcelino García Barragán 1421, Col. Olímpica, C.P. 44430, Guadalajara, Jal.

Teléfono:

Correo electrónico: german.preciat@alumnos.udg.mx

Dr. Ernesto Borrayo Carbajal

Profesor e investigador asociado C

Dept. Bioingeniería traslacional

Módulo M, Centro Universitario de Ciencias Exactas e Ingenierías

Blvd.Marcelino García Barragán 1421, Col. Olímpica, C.P. 44430, Guadalajara, Jal.

Teléfono:

Correo electrónico: ernesto.borrayo@alumnos.udg.mx

Dra. Ana Maria Puebla Pérez

Profesor Investigador Titular C

Investigador Biomédico, CIBO, IMSS, 30 años en la Universidad de Guadalajara

Centro Universitario de Ciencias Exactas e Ingenierías

Blvd.Marcelino García Barragán 1421, Col. Olímpica, C.P. 44430, Guadalajara, Jal.

Teléfono:+52 (33) 1378 5900.

Correo electrónico: maria.puebla@academicos.udg.mx

Dra. Patricia Martins Conde

Investigador postdoctoral

Universidad de Luxemburgo, Centro de Biomedicina de Sistemas de Luxemburgo

Roudeneck,1, Boulevard du Jazz, L-4370 ESCH

Teléfono: (+352) 46 66 44 6266

Correo electrónico: patricia.martinsconde@uni.lu

**TESISTA**

Ing. Biomédico, Eduardo Ruiz Robles

Estudiante de la Maestría en Ciencias en Bioingeniería y Cómputo inteligente

Centro Universitario de Ciencias Exactas e Ingenierías

Blvd.Marcelino García Barragán 1421, Col. Olímpica, C.P. 44430, Guadalajara, Jal.

Teléfono: 3311154110

Correo electrónico: eduardo.ruiz7275@alumnos.udg.mx

Índice

[**Índice de abreviaturas 4**](#_v88993ebve25)

[**Resumen 5**](#_5qgzbm9hrt9j)

[**Marco teórico 6**](#_f3ulf6atcg9s)

[**Justificación 7**](#_ux5mnlabvg7t)

[**Planteamiento del problema 8**](#_fbnu6mkxqyy9)

[**Pregunta de investigación 8**](#_hq03msfp6e5h)

[**Objetivos 8**](#_e6c8ouzi6vz3)

[**Objetivos específicos 9**](#_bmrvuiy3d2i)

[**Hipótesis 9**](#_chp98pi81j2g)

[**Material y métodos 9**](#_6z0wuc594k3j)

[**Consideraciones éticas 10**](#_32ez9e78xgx1)

[1. Definición de Información Confidencial 11](#_gjdgxs)

[2. Uso de la Información Confidencial 11](#_30j0zll)

[3. Obligaciones de Confidencialidad 11](#_1fob9te)

[4. Medidas de Seguridad 11](#_3znysh7)

[5. Duración de la Confidencialidad 11](#_2et92p0)

[6. Excepciones a la Confidencialidad 12](#_tyjcwt)

[7. Sanciones por Incumplimiento 12](#_3dy6vkm)

[8. Destrucción de la Información 12](#_4d34og8)

[9. Jurisdicción 12](#_2s8eyo1)

[**Recursos, financiamiento y factibilidad 13**](#_bn5kn3c5or0n)

# 

# 

# 

# 

# 

# 

# Índice de abreviaturas

COBRA Constraint-based reconstruction and analysis

MEGs Modelos a escala genómica

SVM Support vector Machine

# Resumen

**Título:** Oncología de precisión para el tratamiento de cáncer de mama.

**Antecedentes:**

Aunque en los últimos años se han desarrollado modelos in silico integrando datos ómicos, su capacidad predictiva para aplicaciones clínicas o de salud pública sigue siendo limitada, especialmente en casos donde las condiciones clínicas requieren información complementaria más allá de lo genómico.2 Este proyecto busca superar este desafío mediante la incorporación de información clínica en modelos metabólicos de escala genómica utilizando COBRA Toolbox y la correlación de fenotipos metabólicos con datos clínicos a través de técnicas estadísticas y de machine learning, como SVM, redes neuronales y K-means. Esta integración tiene como objetivo identificar patrones predictivos, descubrir nuevos biomarcadores y optimizar estrategias de tratamiento personalizado en la oncología de precisión para el cáncer de mama.

**Objetivo general:**

El cáncer se caracteriza por una serie de cambios en las funciones de distintas capas biomoleculares siendo el metabolismo una de ellas, afectando el comportamiento de la célula y el microambiente tumoral. Los modelos metabólicos de escala genómica son reconstrucciones de las redes metabólicas de células que permiten estudiar el metabolismo celular. Estos se componen de las diferentes reacciones de metabolitos con equilibrio de masa y mediante asociaciones de genes a proteínas que participan en las reacciones. El metabolismo del cáncer puede ser estudiado mediante la combinación de MEGs y métodos de reconstrucción y análisis basado de restricciones (COBRA), que permite analizar cómo factores ambientales y genéticos afectan a un fenotipo en específico. COBRA representa las reacciones metabólicas de un sistema mediante una matriz estequiométrica, incluyendo las tasas de absorción y secreción de un sistema. Estas herramientas pueden adaptarse condiciones específicas de cáncer permitiendo estudiar su metabolismo, como un posible objetivo de fármaco contra el cáncer.1

El objetivo general de esta tesis es desarrollar modelos metabólicos a escala genómica (MEGs) in silico, integrando datos no ómicos como imágenes médicas, registros clínicos y características del microambiente tumoral mediante modelado el basado en restricciones con la herramienta COBRA Toolbox. Además, se buscará correlacionar fenotipos metabólicos con datos clínicos específicos, como niveles de glucosa, oxígeno y otros biomarcadores utilizando técnicas de aprendizaje automático como Support Vector Machines (SVM), k-means y redes neuronales. Al integrar esta información clínica en los modelos metabólicos, se busca mejorar la precisión en el diagnóstico y pronóstico de la enfermedad ofreciendo herramientas para definir tratamientos más personalizados y eficaces, así como ofrecer nuevas perspectivas en la investigación para el cáncer de mama.

**Material y métodos:**

Se recolectarán datos clínicos de pacientes con cáncer de mama, para su integración en modelos metabólicos de escala genómica, Una vez integrados los datos clínicos en los modelos, se correlacionarán fenotipos específicos con características clínicas particulares de cada paciente utilizando técnicas de machine learning.

**Recursos e infraestructura:**

Para llevar a cabo el proyecto, se requiere la recolección de datos clínicos de pacientes con cáncer de mama, los cuales serán analizados mediante técnicas de *machine learning* y ciencia de datos.

**Experiencia del grupo:**

El grupo de investigadores que participarán en el proyecto es multidisciplinario, habiendo investigadores enfocados en el modelado de sistemas metabólicos, análisis genómico, e inmunofarmacología.

**Tiempo para desarrollarse:**

El tiempo estimado para el desarrollo del proyecto es de 1 año

# Marco teórico

**Antecedentes**

Se han creado modelos mediante la integración de datos ómicos en los últimos años, sin embargo sólo pocos con buena capacidad predictiva para ser implementados en prácticas clínicas o dominios de salud pública, debido a que puede haber enfermedades o condiciones clínicas que no son completamente explicadas por datos genómicos y requieren información clínica adicional para mejorar la predicción.2

Este proyecto tiene como objetivo integrar la información clínica de los pacientes en modelos de escala genómica, haciendo uso del modelado basado en restricciones en COBRA Toolbox y correlacionar fenotipos metabólicos con datas clínicos mediante técnicas estadísticas, como correlación de pearson y machine learning, como SVM debido a su robustez en la clasificación de conjuntos de datos de alta dimensionalidad, redes neuronales, ya que son capaces de capturar relaciones no lineales en los datos y K-means, útil para agrupar pacientes con características metabólicas similares. Al correlacionar fenotipos metabólicos con datos clínicos específicos, este enfoque puede ayudar a identificar patrones valiosos para la predicción de respuestas a tratamientos, identificar biomarcadores, y mejorar las estrategias de tratamiento personalizado en oncología

**Cáncer de mama y su impacto en Salud Pública**

El cáncer de mama es el tipo de cáncer que más afecta a la mujer a nivel mundial, presentando en el 2020 más de 685,000 muertes en el mundo y 2.3 casos nuevos según la Organización Mundial de la Salud. Esta patología es heterogénea, manifestándose en distintos subtipos de cáncer y con características específicas en cada tumor, lo que hace que el tratamiento deba ser personalizado para cada paciente.4

El estudio del metabolismo en el cáncer es una de las áreas más exploradas de la biología, con más de cinco décadas de investigación. Esto se debe a que el metabolismo en las células cancerosas presenta alteraciones significativas en comparación con las células normales.10 Las células cancerosas, incluyendo las del cáncer de mama, suelen depender de la glicólisis aeróbica para generar la energía necesaria para su supervivencia y proliferación.8 Además, las modificaciones metabólicas juegan un papel esencial en la progresión de la enfermedad; por ejemplo, la desregulación de rutas como la del propanoato contribuye a un perfil celular más agresivo, incrementando el potencial metastásico en cánceres de mama y pulmón.9

El cáncer de mama se clasifica en cuatro subtipos moleculares basados en la presencia o ausencia de ciertos marcadores,como los receptores de progesterona (PR), el receptor 2 del factor de crecimiento epidérmico (HER2) y los receptores de estrógeno (ER) 8 Estos subtipos impactan de manera diversa el crecimiento y el comportamiento del tumor, generando la necesidad de tratamientos personalizados que se ajusten a las características específicas del tumor de cada paciente.4

**Aplicación de COBRA Toolbox para el modelado metabólico**

COBRA toolbox representa, de manera matemática y computacional, la relación entre genotipos y fenotipos de un sistema bioquímico, a través de restricciones basadas en leyes físico químicas, genéticas y del ambiente. Este enfoque ha sido aplicado para generar y analizar modelos metabólicos tanto normales como patológicos en el metabolismo humano.3 Un modelo metabólico a escala genómica se representa mediante una matriz estequiométrica de tamaño nxn donde las filas representan a metabolitos y las columnas reacciones, En estos modelos, el flujo de cada reacción se representa mediante un vector v que permite calcular cantidades de interés como la producción de biomasa, proporcionando información detallada sobre la contribución de cada metabolito y reacción en el sistema metabólico completo.2

El cáncer conlleva una serie de cambios complejos en el comportamiento metabólico de sus células y del microambiente tumoral, difíciles de medir debido a la variabilidad de tipos celulares y la dificultad de cuantificar ciertos metabolitos.

Los modelos metabólicos a escala genómica, permiten un análisis computacional de las redes metabólicas de estas células, facilitando el estudio detallado de los flujos metabólicos en el microambiente tumoral y ofreciendo una perspectiva integral para comprender mejor su dinámica y potenciales intervenciones terapéuticas.2

Uno de los métodos más utilizados en COBRA es el análisis de balance de flujo (FBA, por sus siglas en inglés). Este método asume que el sistema se encuentra en un estado estacionario y tiene como objetivo maximizar o minimizar una función biológica, como la producción de biomasa, que simula el crecimiento celular. A través del FBA, es posible evaluar el comportamiento de redes metabólicas bajo diferentes condiciones, permitiendo predecir el impacto de cambios en nutrientes o mutaciones en el metabolismo celular.2

**Aplicaciones de Machine Learning y Deep Learning**

Los datos clínicos pueden ser tanto estructurados como no estructurados, lo que implica que algunas variables pueden no tener una relación clara entre ellas y actuar de manera independiente. Aquí es donde el deep learning se vuelve especialmente útil, ya que puede identificar patrones en datos complejos y no estructurados mediante técnicas avanzadas de aprendizaje automático. Estas técnicas permiten que los modelos no solo aprendan de los datos, sino que también se adapten a condiciones cambiantes con el tiempo. El algoritmo de SVM (Support Vector Machine) es un algoritmo de aprendizaje supervisado que funciona creando líneas para separar datos en un espacio multidimensional, y que sirve para hacer algún tipo de clasificación, como la predicción de biomarcadores.8

Por otro lado, el clustering o agrupamiento es una técnica de aprendizaje no supervisado que no requiere datos etiquetados y resulta muy útil para identificar patrones en las características de los datos. Esta técnica busca descubrir características inherentes en los datos para crear grupos o clusters, generalmente basándose en similitudes entre ellos.8

Uno de los objetivos es identificar patrones en los datos que permitan correlacionar fenotipos metabólicos con características clínicas del paciente, como ciertos biomarcadores específicos o antecedentes clínicos.

Finalmente, se llevará a cabo una evaluación de la capacidad predictiva de los modelos desarrollados, con el fin de validar su efectividad en la correlación de fenotipos metabólicos con características clínicas específicas. Esta evaluación no solo buscará garantizar la precisión y fiabilidad de las predicciones, sino que también aspira a contribuir al avance del conocimiento científico en el ámbito de la biomedicina. Además, se espera que los resultados obtenidos sean divulgados en conferencias y publicaciones científicas, promoviendo así una mayor comprensión y aplicación de estos hallazgos en la práctica clínica y en futuras investigaciones.

# **Justificación**

El cáncer de mama es la principal causa de muerte en mujeres en México desde 2006. Entre 2010 y 2017, Guadalajara y Chapala destacaron con las tasas de mortalidad más altas del país, registrando 21.2 y 19.5 defunciones por cada 100,000 mujeres mayores de 20 años, respectivamente.11 En 2022, se notificaron 23,790 casos nuevos de cáncer de mama en personas de 20 años o más.12

La integración de modelos de IA en la medicina de precisión permite profundizar en la comprensión de la biología del cáncer, abordando la heterogeneidad tumoral y posibilitando estrategias de tratamiento personalizadas mediante el uso de datos multi-ómicos, como los genómicos, proteómicos, metabolómicos y transcriptómicos.4  En este proyecto, la incorporación de datos no ómicos, un enfoque con escasa bibliografía previa, representa una contribución innovadora en la investigación del cáncer y la medicina personalizada.

Un estudio por el MIT technology Review y GE Healthcare ha demostrado que 79% de clínicos creen que las herramientas de IA les han ayudado a optimizar procesos permitiendo brindar una atención más centrada al paciente, entre los beneficios clave para el uso de tecnologías de IA y ayudar en la toma de decisiones clínicas en oncología son las siguientes:

* Diagnóstico automatizado de cáncer
* Capacidades de predicción mejoradas
* Actualizaciones de datos en tiempo real
* Atención personalizada
* Mejores resultados, y
* Aumento de la eficiencia, mejora de resultados y reducción de los costos.5

# **Planteamiento del problema**

La diversidad clínica y genética de las enfermedades neoplásicas dificulta la capacidad de predecir con precisión el desarrollo de la enfermedad en un paciente en particular.

Existen diversas pruebas predictivas dirigidas a pacientes con cáncer de mama, sin embargo estas pruebas se basan en el análisis de expresión de varios genes presentes en el tejido tumoral, y varían tanto en su capacidad para predecir el comportamiento de la enfermedad como en los genes que evalúan.6

Si bien existen modelos de predicción clínica como Predict, CancerMath y el Nottingham Prognostic Index que apoyan la toma de decisiones sobre estrategias terapéuticas, muchos presentan fallas metodológicas, sesgos o poca transparencia en su desarrollo y validación.

En una revisión de los modelos utilizados para aplicación clínica, realizado el 2022 por Tom A. Hueting, et al, en Holanda, más del 95% presentó un riesgo de sesgo y el 31% no pudo ser aplicado a los pacientes.12 Mediante la integración de datos no ómicos de los pacientes, se busca obtener modelos más personalizados y precisos para reducir este sesgo y que sean aplicables para estrategias de tratamiento.

Por lo tanto, el proyecto de tesis busca correlacionar condiciones metabólicas específicas a esa variabilidad clínica del paciente, integrando la información clínica en MEGs y utilizando técnicas avanzadas de machine learning y análisis de datos clínicos de pacientes con cáncer de mama. Esto permitirá mejorar la precisión de las predicciones sobre el comportamiento de la enfermedad, desarrollando modelos más robustos y personalizados, que eventualmente podrían ser implementados en la práctica clínica para apoyar la toma de decisiones en el tratamiento del cáncer de mama.

# **Pregunta de investigación**

¿Qué tan efectivas son las correlaciones entre datos metabólicos y clínicos para predecir biomarcadores clave en el tratamiento personalizado del cáncer de mama?

**Objetivo general**Desarrollar modelos metabólicos a escala genómica (MEGs) personalizados que integren datos clínicos relevantes para analizar su relación con la progresión y/o pronóstico del cáncer en pacientes, empleando herramientas estadísticas y de machine learning.

**Objetivos específicos**

1. **Integración de datos clínicos y modelado metabólico**1.1 Recolectar y preprocesar datos clínicos de pacientes con cáncer de mama para su integración en el análisis.  
   1.2 Diseñar y construir modelos metabólicos a escala genómica (MEGs) personalizados que incorporen información clínica asociada a la progresión y/o pronóstico del cáncer.
2. **Análisis avanzado y adaptación al microambiente tumoral**2.1 Adaptar los modelos metabólicos a las características específicas del microambiente tumoral para mejorar su precisión.  
   2.2 Realizar análisis estadísticos y aplicar modelos de machine learning, como SVM, redes neuronales y K-means, para identificar patrones y correlaciones en los datos clínicos.

# **Hipótesis**

Existe una correlación significativa entre fenotipos metabólicos derivados de modelos metabólicos a escala genómica (MEGs) y características clínicas de los pacientes, lo que sugiere que hay una variabilidad en el metabolismo celular relacionada con la información clínica de cada paciente con cáncer de mama.

# **Material y métodos**

Para el desarrollo de este proyecto, se requiere un equipo de cómputo personal de alta capacidad que permita ejecutar modelos complejos de machine learning y Modelos de Escala Genómica (MEGs), así como una conexión a internet estable que facilite la transferencia y almacenamiento seguro de los datos. Además, será necesario contar con software especializado, como herramientas para análisis de datos, almacenamiento y modelado metabólico, destacando el uso de la COBRA Toolbox, que permitirá la manipulación y análisis de redes metabólicas de manera eficiente.

En la primera etapa, se llevará a cabo la obtención y preprocesamiento de datos clínicos. Los datos iniciales, que se encuentran en hojas físicas, serán digitalizados y convertidos a formato CSV para facilitar su análisis computacional. Este proceso incluirá la limpieza de datos, la eliminación de valores atípicos y el manejo de datos faltantes para asegurar la calidad y consistencia de la información.

Posteriormente, se realizará la integración de los datos clínicos en los MEGs. Esto implicará analizar metabolitos específicos asociados al cáncer, que serán utilizados como base para la creación de modelos metabólicos. Aquí, la COBRA Toolbox será fundamental para construir y ajustar los modelos metabólicos, permitiendo llevar a cabo simulaciones y análisis de flujos metabólicos bajo diferentes condiciones.

Una vez integrados los datos, se aplicarán técnicas de machine learning para identificar patrones de correlación. En esta etapa, se emplearán métodos de agrupación, como k-means, para encontrar patrones significativos en los fenotipos metabólicos.

Finalmente, se evaluará la capacidad predictiva de los modelos desarrollados. Esto incluirá pruebas rigurosas para validar su desempeño y precisión, utilizando métricas estándar como sensibilidad, especificidad. El uso de la COBRA Toolbox también contribuirá en esta etapa, al permitir simulaciones avanzadas y análisis de sensibilidad para evaluar el impacto de diferentes parámetros en los resultados metabólicos. Este enfoque integral asegura que los modelos no solo sean precisos, sino también útiles para aplicaciones futuras en oncología de precisión.

**Resultados esperados e Impacto:**

Los resultados esperados incluyen 1) identificar patrones y correlaciones claras entre fenotipos metabólicos y características clínicas de los pacientes; 2) Desarrollar MEGs de diferentes células tumorales mediante la integración de datos clínicos de los pacientes; 3) Correlacionar fenotipos metabólicos con características clínicas del paciente que ayuden a predecir la respuesta a un tratamiento en específico o pronóstico.

Los resultados podrían contribuir a una mejor comprensión de las rutas metabólicas que influyen en la progresión del cáncer o resistencia a cierto tratamiento, lo que puede permitir la identificación de nuevos biomarcadores metabólicos y a predecir respuestas a tratamiento, haciendo las terapias más efectivas y personalizadas.

Un desafío importante es la falta de uniformidad en los datos clínicos, que incluyen variables cualitativas y cuantitativas, lo cual complica su integración en los modelos metabólicos. Para abordar este problema, se emplearán técnicas de machine learning como SVM, redes neuronales y k-means para identificar patrones y variables clave que impacten en el metabolismo celular y puedan incorporarse en los MEGs.

# **Consideraciones éticas**

Se respetará la confidencialidad y seguridad de la información clínica de los pacientes con la que se estará llevando a cabo la investigación en todo momento, en base al documento anexado de contrato de confidencialidad firmado por el investigador a cargo, el investigador principal y el tesista.

# **Recursos, financiamiento y factibilidad**

Este proyecto se realizará en colaboración con investigadores de la Universidad de Guadalajara, una investigadora de la Universidad de Luxemburgo y en convenio con el Centro de Investigación Biomédica de Occidente (CIBO), institución que nos proporcionará la información clínica de los pacientes con la que se estará llevando a cabo la investigación respetando la confidencialidad y seguridad de los datos de los pacientes en todo momento.

Uno de los desafíos en el proyecto es la falta de uniformidad de los datos clínicos incluyendo variables cualitativas y cuantitativas, dificultando su integración en los modelos metabólicos.2 Por lo cual se plantea utilizar técnicas de análisis de datos y machine learning como SVM, redes neuronales y k means para encontrar información que impacten en el metabolismo celular y pueda incluirse en los MEGs.

Los recursos necesarios para llevar a cabo el proyecto son un equipo de cómputo personal y conexión a la red para el desarrollo de los modelos de machine learning y MEGs.

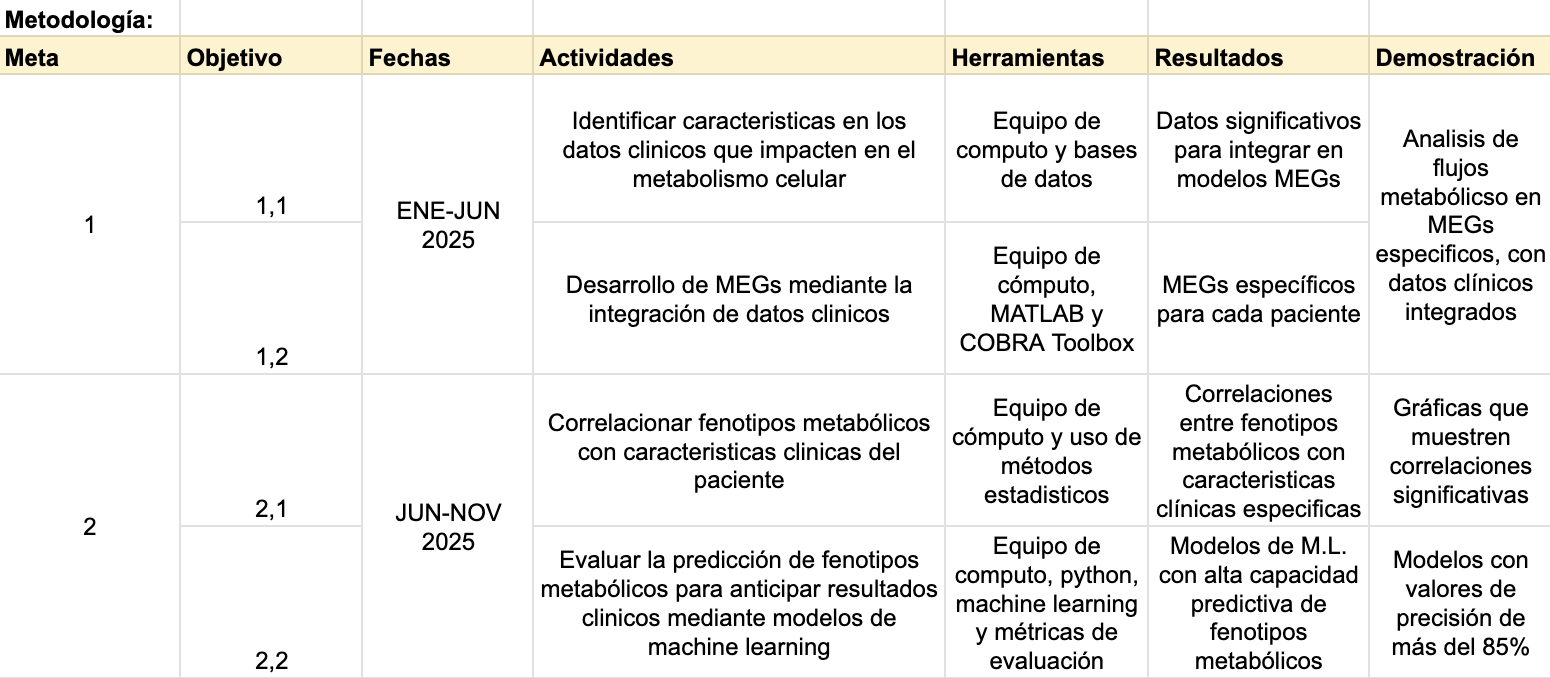
El código generado durante el proyecto será compartido en un repositorio de GitHub del Grupo de Ingeniería Metabólica in Silico (GIMI), lo que permitirá la reproducibilidad, transparencia y colaboración con otros investigadores.

**Referencias**

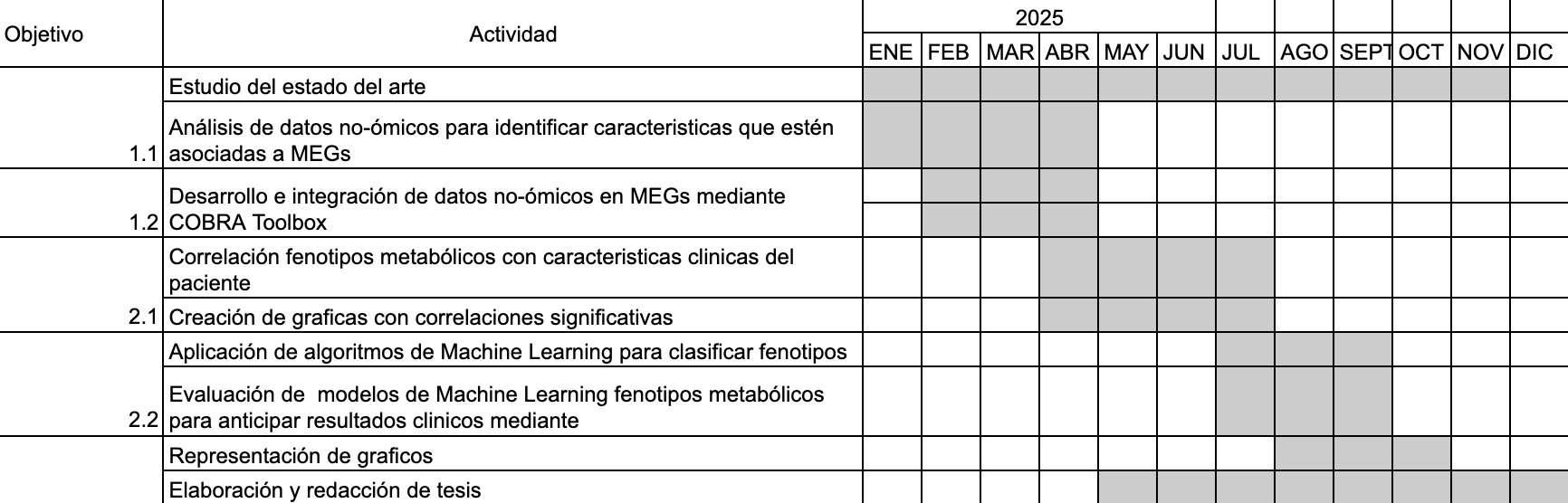
1. Ng, R. H., Lee, J. W., Baloni, P., & Diener, C. (2022). Constraint-based reconstruction and analyses of metabolic models: Open-source Python tools and applications to cancer. Frontiers in Oncology, 12, 914594. <https://doi.org/10.3389/fonc.2022.914594>
2. López de Maturana E, Alonso L, Alarcón P, Martín-Antoniano IA, Pineda S, Piorno L, Calle ML, Malats. (2019) N. Challenges in the Integration of Omics and Non-Omics Data. Genes. 2019; 10(3):238. <https://doi.org/10.3390/genes10030238>
3. Heirendt, L., Arreckx, S., Pfau, T. *et al.* Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc* 14, 639–702 (2019). <https://doi.org/10.1038/s41596-018-0098-2>
4. Papalexis, P., Georgakopoulou, V.E., Drossos, P.V., Thymara, E., Nonni, A., Lazaris, A.C. ... Thomopoulou, G. (2024). Precision medicine in breast cancer (Review). Molecular and Clinical Oncology, 21, 78. <https://doi.org/10.3892/mco.2024.2776>
5. Loaiza-Bonilla, A. 2022. La inteligencia artificial en oncología: contexto actual y una visión hacia la próxima década. *Medicina*. 43, 4 (ene. 2022), 527–534. DOI:<https://doi.org/10.56050/01205498.1642>
6. Reyna Villasmil, E. (2024). Utilidad en la prevención, tratamiento y pronóstico de la medicina de precisión en oncología. Revista Repertorio De Medicina Y Cirugía, 33(2), 117-123. <https://doi.org/10.31260/RepertMedCir.01217372.1279>
7. F. Franco, E., & J. Ramos, R. (2019). Aprendizaje de máquina y aprendizaje profundo en biotecnología: aplicaciones, impactos y desafíos. *Ciencia, Ambiente Y Clima*, *2*(2), 7–26. <https://doi.org/10.22206/cac.2019.v2i2.pp7-26>
8. Zou, J., Mai, C., Lin, Z., Zhou, J., & Lai, G. (2024). Targeting metabolism of breast cancer and its implications in T cell immunotherapy. Frontiers in Immunology, 15. <https://doi.org/10.3389/fimmu.2024.1381970>
9. Gomes, A.P., Ilter, D., Low, V. *et al.* Altered propionate metabolism contributes to tumour progression and aggressiveness. *Nat Metab* 4, 435–443 (2022). <https://doi.org/10.1038/s42255-022-00553-5>
10. Ralph J. DeBerardinis, Navdeep S. Chandel ,Fundamentals of cancer metabolism.Sci. Adv.2,e1600200(2016).DOI:10.1126/sciadv.1600200
11. Ramos-Herrera, Igor M., Reyna-Sevilla, Antonio, González Castañeda, Miguel E., Robles-Pastrana, Juan D., Herrera-Echauri, Dulce D., & González-Rivera, César A.. (2020). Cáncer de mama en Jalisco. Análisis espacial de la mortalidad en 2010-2017. Gaceta médica de México, 156(6), 542-548. Epub 27 de mayo de 2021.<https://doi.org/10.24875/gmm.20005546>
12. Instituto Nacional de Estadística y Geografía (INEGI). (2023, 17 de octubre). Comunicado de prensa número 595/23: Estadísticas a propósito del Día Internacional de la Lucha Contra el Cáncer de Mama (19 de octubre). Comunicación social. <https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2023/EAP_CMAMA23.pdf>
13. Hueting, T. A., van Maaren, M. C., Hendriks, M. P., Koffijberg, H., & Siesling, S. (2022). The majority of 922 prediction models supporting breast cancer decision-making are at high risk of bias. Journal of Clinical Epidemiology, 152, 238-247. https://doi.org/10.1016/j.jclinepi.2022.10.016.

**Anexos**

Metodología**:**

****

Cronograma de actividades:

****

**CONTRATO DE CONFIDENCIALIDAD**

Este contrato de confidencialidad (“Contrato”) se celebra entre:

**Dr. German Andres Preciat Gonzalez**, Investigador en la Universidad de Guadalajara, en adelante “*El Investigador Principal*”;

**Ing. Eduardo Ruiz Robles**, alumno del Dr. German Andres Preciat Gonzalez y colaborador en el proyecto de investigación, en adelante “*El Alumno*”;

**Dra. Martha Patricia Gallegos**, Investigadora en la División de Genética del CIBO, IMSS, en adelante "*La Proveedora de Datos*", quién será la responsable de proporcionar la información clínica, quedando su mención a título informativo.

Las partes acuerdan lo siguiente:

### **1. Definición de Información Confidencial**

La Información Confidencial incluye, pero no se limita a:

* Historias clínicas de pacientes con cáncer proporcionadas por *La Proveedora de Datos*.
* Datos personales identificables de los pacientes (nombre, fecha de nacimiento, direcciones, etc.).
* Resultados médicos, diagnósticos, pruebas, y cualquier otro dato relacionado con los pacientes.

### **2. Uso de la Información Confidencial**

La información confidencial solo se utilizará con el propósito de:

* Desarrollar modelos de análisis de datos y/o modelos metabólicos para la caracterización de la enfermedad de cáncer en los pacientes.

### **3. Obligaciones de Confidencialidad**

* *El Investigador Principa*l y *El Alumno* se comprometen a no divulgar, reproducir ni compartir la Información Confidencial, ni parcial ni totalmente, a terceros sin el consentimiento previo y por escrito de La Proveedora de Datos.
* *El Alumno* solo podrá consultar la información para los fines establecidos, y bajo ninguna circunstancia podrá compartirla o modificarla.

### **4. Medidas de Seguridad**

Las partes se comprometen a implementar las siguientes medidas de seguridad para proteger la Información Confidencial:

* Almacenamiento de los datos en servidores seguros con acceso restringido.
* Cifrado de la información para protegerla frente a accesos no autorizados.
* El acceso a los datos estará limitado exclusivamente al *Investigador Principal* y *El Alumno*.

### **5. Duración de la Confidencialidad**

Este contrato permanecerá en vigor durante el desarrollo del proyecto y continuará vigente de manera indefinida después de la finalización del mismo, manteniendo las obligaciones de confidencialidad sin límite de tiempo.

### **6. Excepciones a la Confidencialidad**

* Los resultados de los análisis y los modelos desarrollados podrán divulgarse públicamente, siempre y cuando no se incluyan datos clínicos identificables de los pacientes.

### **7. Sanciones por Incumplimiento**

En caso de incumplimiento del presente contrato por parte de *El Investigador Principal* o *El Alumno*, las sanciones podrán incluir:

* Acción disciplinaria por parte de la Universidad de Guadalajara, que podría resultar en la expulsión o terminación de la relación académica.
* Acciones legales en conformidad con las leyes aplicables de México, incluyendo sanciones económicas o penales por el uso indebido de datos personales y médicos conforme a la Ley Federal de Protección de Datos Personales en Posesión de los Particulares.

### 

### **8. Destrucción de la Información**

Al finalizar el proyecto, los datos clínicos deberán ser eliminados permanentemente de las computadoras y otros dispositivos de almacenamiento del *Investigador Principal* y *El Alumno*, salvo que exista una obligación legal de conservarlos.

### **9. Jurisdicción**

Este contrato se rige por las leyes de los Estados Unidos Mexicanos, y las partes se someten a la jurisdicción de los tribunales competentes en la ciudad de Guadalajara Jalisco para cualquier conflicto relacionado con este contrato.

**Firmado por:**

| Firma y Fecha  **Dr. German Andres Preciat Gonzalez** Investigador Principal | Firma y Fecha  **Ing. Eduardo Ruiz Robles** Alumno |
| --- | --- |