

# Informe\_PEC1

Eva Álvarez Torres

2025-03-31

## TABLA DE CONTENIDOS

1	RESUMEN .....	1
2	OBJETIVOS.....	2
3	MÉTODOS.....	2
4	RESULTADOS .....	3
5	DISCUSIÓN.....	9
6	CONCLUSIONES .....	9
7	REFERENCIAS .....	9
8	ANEXO.....	10
8.1	Otros gráficos de interés: .....	10
8.2	Preguntas 1, 2 y 4 de la PEC .....	11

## 1 RESUMEN

Se analizaron datos de 39 pacientes sometidos a cirugía bariátrica (bypass gástrico o gastrectomía tubular) con el objetivo de evaluar la evolución de variables fisiológicas y metabolómicas a lo largo del tiempo. Se construyó un objeto SummarizedExperiment para organizar los datos y se realizó un análisis exploratorio mediante gráficos de tendencias, análisis de componentes principales, mapa de calor (heatmap) y pruebas estadísticas.

Se observó una disminución sostenida del peso, del índice de masa corporal (IMC) y de las medidas antropométricas, junto con una mejoría en la hemoglobina glicada (HBA1C). El análisis de varianza (ANOVA) reveló una interacción significativa entre el tiempo y el tipo de cirugía en la evolución del colesterol LDL, lo que sugiere que su trayectoria depende de la técnica quirúrgica empleada. En conjunto, los resultados indican un beneficio metabólico global de la cirugía bariátrica, con un perfil lipídico más favorable tras el bypass gástrico.

## 2 OBJETIVOS

Se analizaron los datos de 39 pacientes obesos sometidos a cirugía bariátrica con los siguientes objetivos:

1. Analizar si la cirugía bariátrica tiene un impacto positivo en la salud metabólica de los pacientes.
2. Evaluar si existen diferencias en la respuesta fisiológica y metabolómica en función de la técnica quirúrgica utilizada (bypass vs. tubular).
3. Explorar la posible correlación entre la pérdida de peso y la evolución de metabolitos clave como LDL, HDL y hemoglobina glicada.

Se decidió focalizar el análisis en estos tres metabolitos debido a su relevancia clínica en el contexto del síndrome metabólico, el riesgo cardiovascular y el control glucémico. Estos indicadores son de uso habitual en la evaluación de pacientes sometidos a cirugía bariátrica, y su evolución puede aportar información directa sobre la eficacia del tratamiento.

## 3 MÉTODOS

Se utilizó un dataset importado desde el [repositorio recomendado en el enunciado](#) en formato .csv analizado en RStudio mediante un documento RMarkdown. Inicialmente, el desarrollo se realizó de forma local, y posteriormente se generó un script en R con control de versiones mediante Git. Se creó un [repositorio GitHub](#) donde incluir el proyecto completo y cuyo enlace está disponible en el *apartado 7: Referencias*.

El dataset incluía 39 pacientes y 690 variables, de las cuales 6 correspondían a metadatos (edad, sexo, grupo (según fenotipo saludable o no), tipo de cirugía) y las restantes a medidas fisiológicas y metabolómicas repetidas en cuatro tiempos: T0 (antes de la cirugía), T2, T4 y T5.

Se creó un objeto SummarizedExperiment, que permite almacenar de forma estructurada los datos cuantitativos (assay) y metadatos (colData). Este objeto se utiliza habitualmente en bioinformática por su versatilidad y estructura clara.

Se llevó a cabo un análisis exploratorio estructurado en:

1. Descripción de los metadatos (tablas de frecuencia, histogramas, boxplots).
2. Exploración homogeneidad de grupos de tratamiento
  - a. Análisis de Componentes Principales (PCA) para explorar la homogeneidad de los grupos en T0.

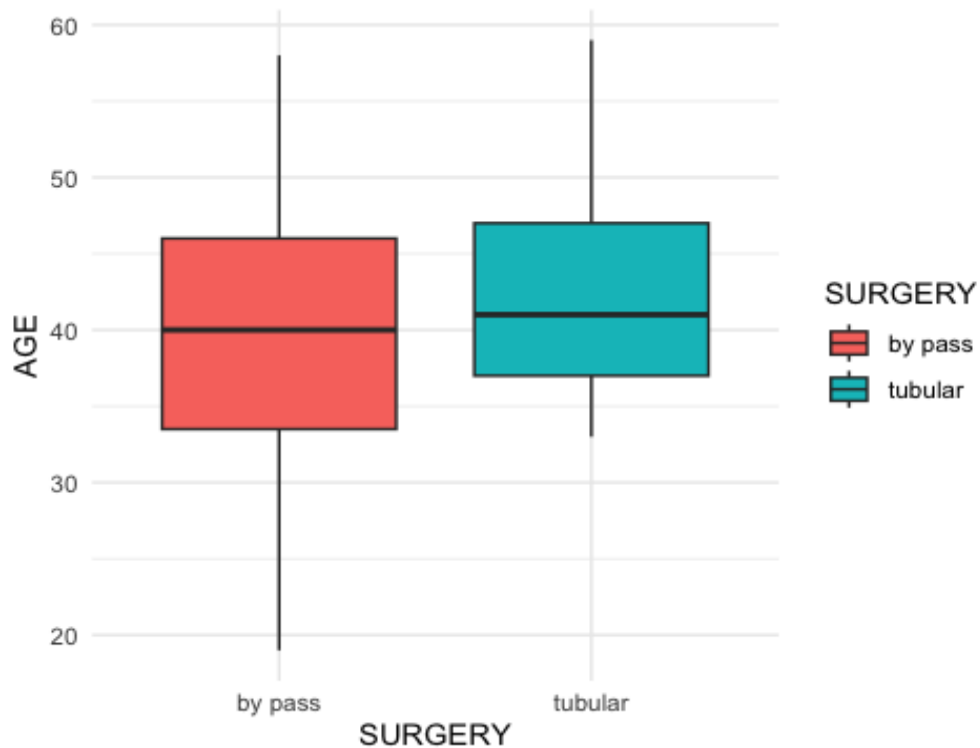
- b. Selección de los 30 metabolitos con mayor variabilidad y representación de mapa de calor (heatmap) para evaluar posibles clusters.
  - c. Identificación de los 10 metabolitos con mayor diferencia entre cirugías en T0.
3. Gráficos de evolución temporal para variables clave como PESO, IMC, CINTURA, CAD, HBA1C, LDL y HDL, estratificados por grupo y tipo de cirugía.
  4. Análisis de correlación entre pérdida de peso y metabolitos en T5.
  5. Comparación de cambios en LDL entre técnicas quirúrgicas mediante t-test y análisis de la varianza.

Se utilizaron las siguientes herramientas: tidyverse, reshape2, ggplot2, stats, SummarizedExperiment, Bioconductor.

En el [repositorio](#) se podrá consultar el script en R, el informe en formato RMarkdown y pdf, el objeto de clase, los datos, así como un resumen de los metadatos y los metadatos en formato markdown.

## 4 RESULTADOS

La muestra incluyó 39 pacientes, 27 mujeres y 12 hombres, distribuidos en 24 pacientes en el grupo 1 (saludables metabólicamente) y 15 en el grupo 2 (no saludables metabólicamente). En cuanto al tipo de cirugía, 26 fueron sometidos a bypass gástrico y 13 a gastrectomía tubular. La distribución por edad fue homogénea entre grupos quirúrgicos (*figura 1*).



*Figura 1. Distribución de edad por tipo de cirugía*

El análisis de componentes principales (PCA) a tiempo 0 (T0) no mostró una separación clara entre los grupos ni entre tipos de cirugía, lo que sugiere una adecuada homogeneidad basal entre las cohortes. Esta observación se refuerza con el mapa de calor generado a partir de los 30 metabolitos con mayor variabilidad, el cual evidenció patrones de expresión heterogéneos, pero sin agrupamientos sistemáticos atribuibles a la técnica quirúrgica o al grupo. (Figuras 2 y 3).

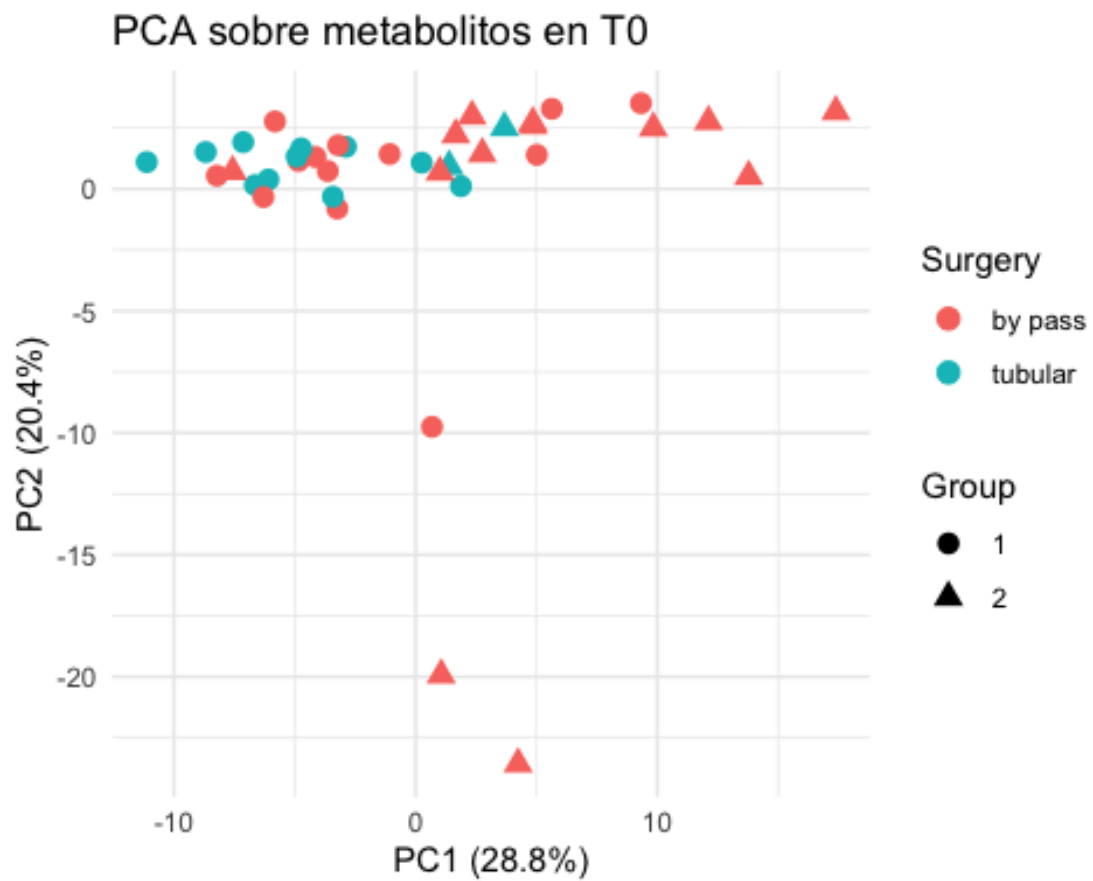
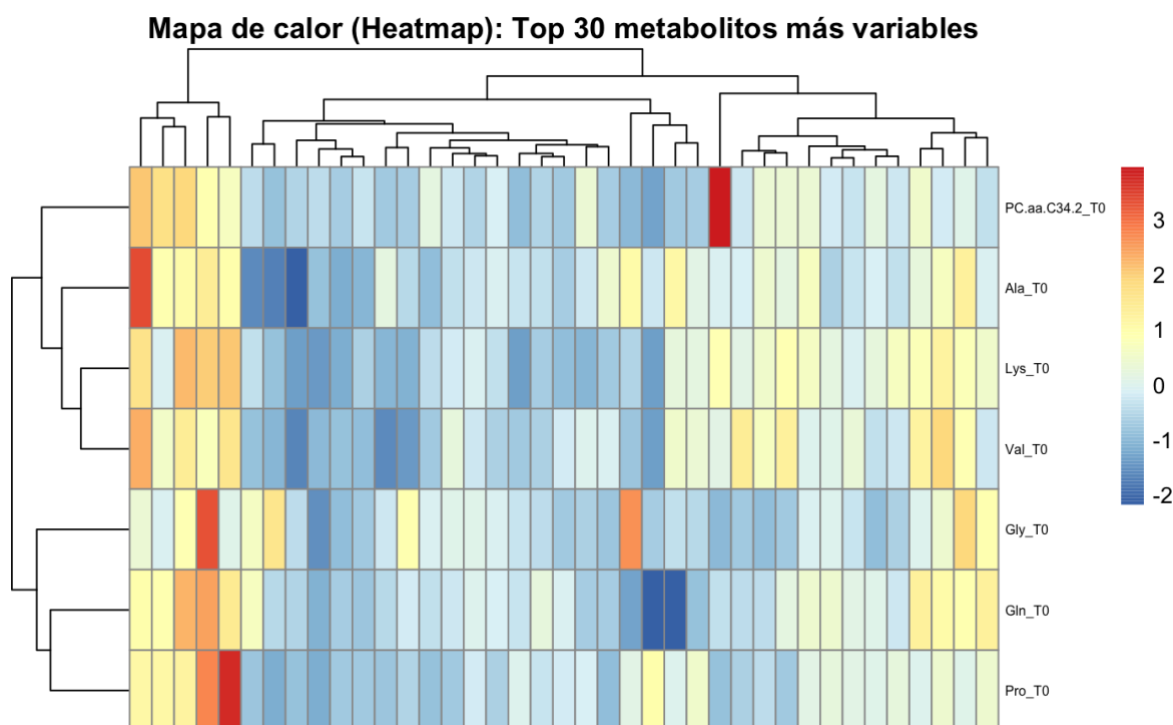


Figura 2. PCA sobre metabolitos a tiempo 0 ( $T_0$ )



*Figura 3. Heatmap de los 30 metabolitos más variables*

Aunque el análisis global mediante PCA y mapa de calor (heatmap) no mostró una separación clara entre los grupos quirúrgicos en el momento basal (T0), las pruebas *t* realizadas individualmente para cada metabolito identificaron diferencias estadísticamente significativas en algunos casos concretos, como C14.1, Ala, HOMA y varios lysoPC. Esto sugiere una cierta heterogeneidad metabólica de base en variables específicas.

No obstante, al no observarse agrupamientos sistemáticos ni patrones de clustering claros, puede considerarse que los grupos son globalmente homogéneos al inicio del estudio.

Las variables clínicas clave —peso, IMC, perímetro de cintura (CINT), y circunferencia abdominal (CAD)— mostraron una disminución progresiva y sostenida a lo largo del tiempo en ambos grupos y tipos de cirugía. Las curvas de tendencia fueron similares entre técnicas, sugiriendo una eficacia comparable en términos de pérdida ponderal y reducción de medidas antropométricas. (*Figuras 4 y 5*)

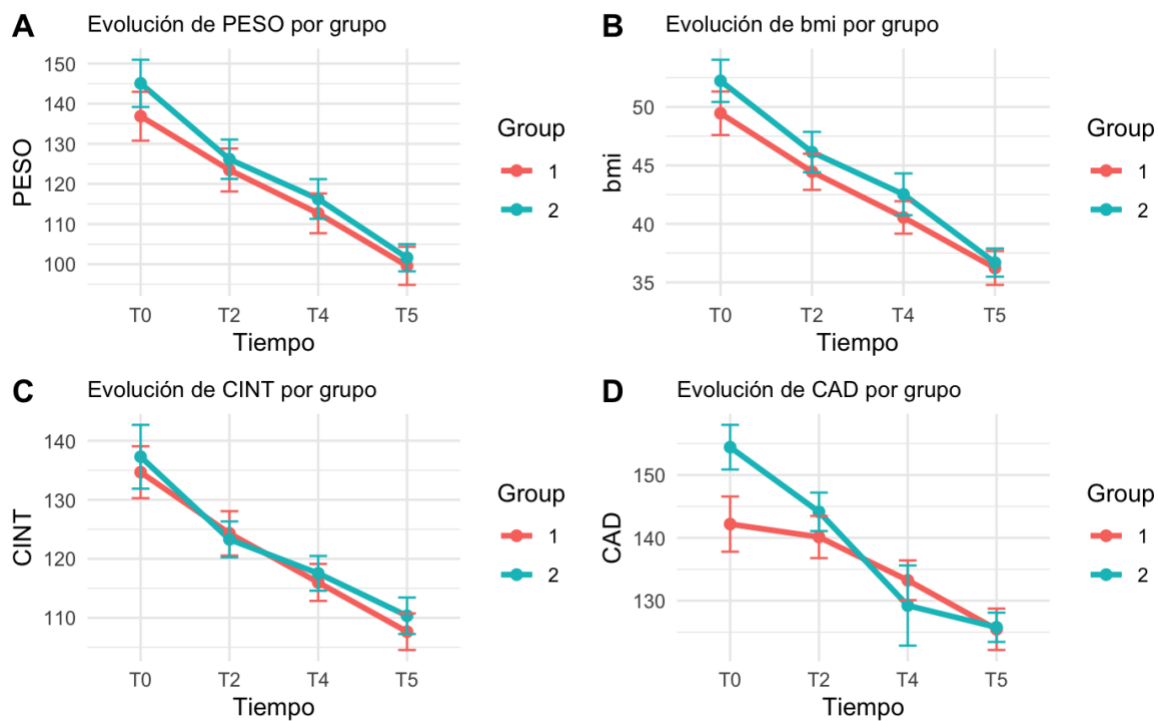


Figura 4. Evolución temporal de medidas antropométricas en función del grupo

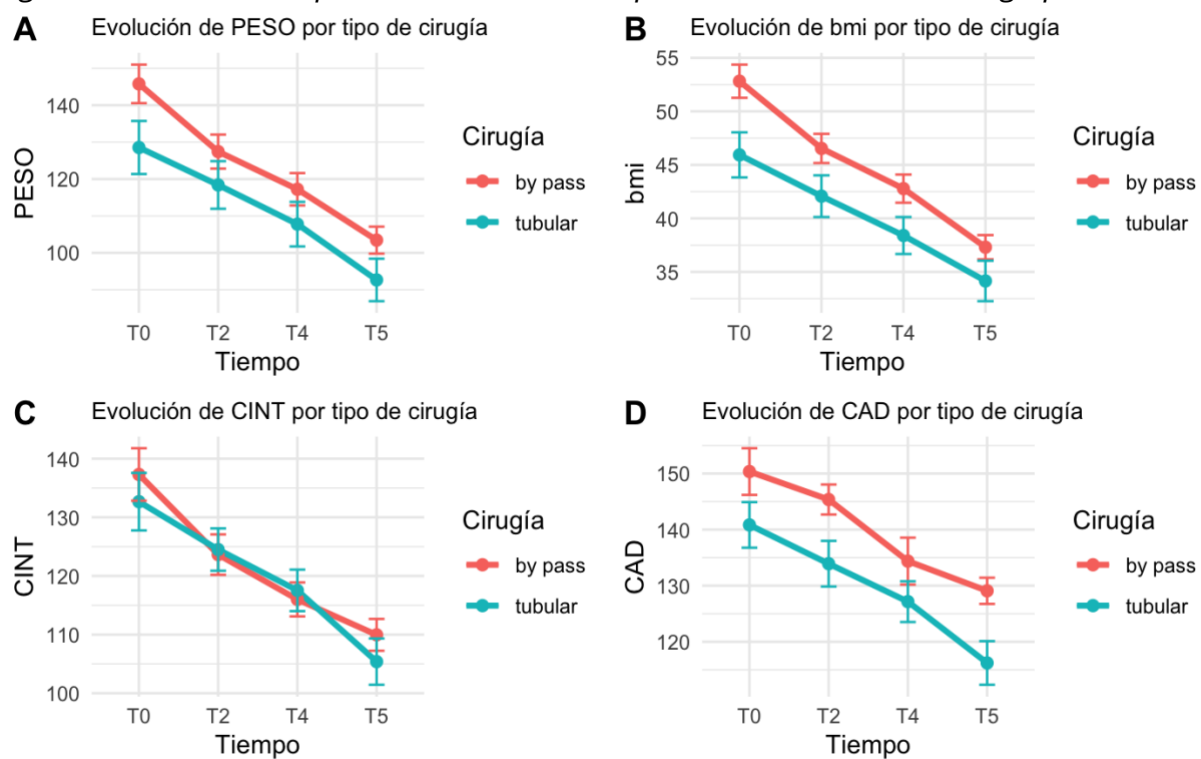


Figura 5. Evolución temporal de medidas antropométricas en función de la técnica quirúrgica

HBA1C mejoró de forma consistente tras la intervención.

En cuanto al colesterol HDL, se observó una tendencia general al alza, especialmente tras el bypass gástrico. Por el contrario, el LDL mostró una evolución divergente según el tipo de cirugía: (Figura 6)

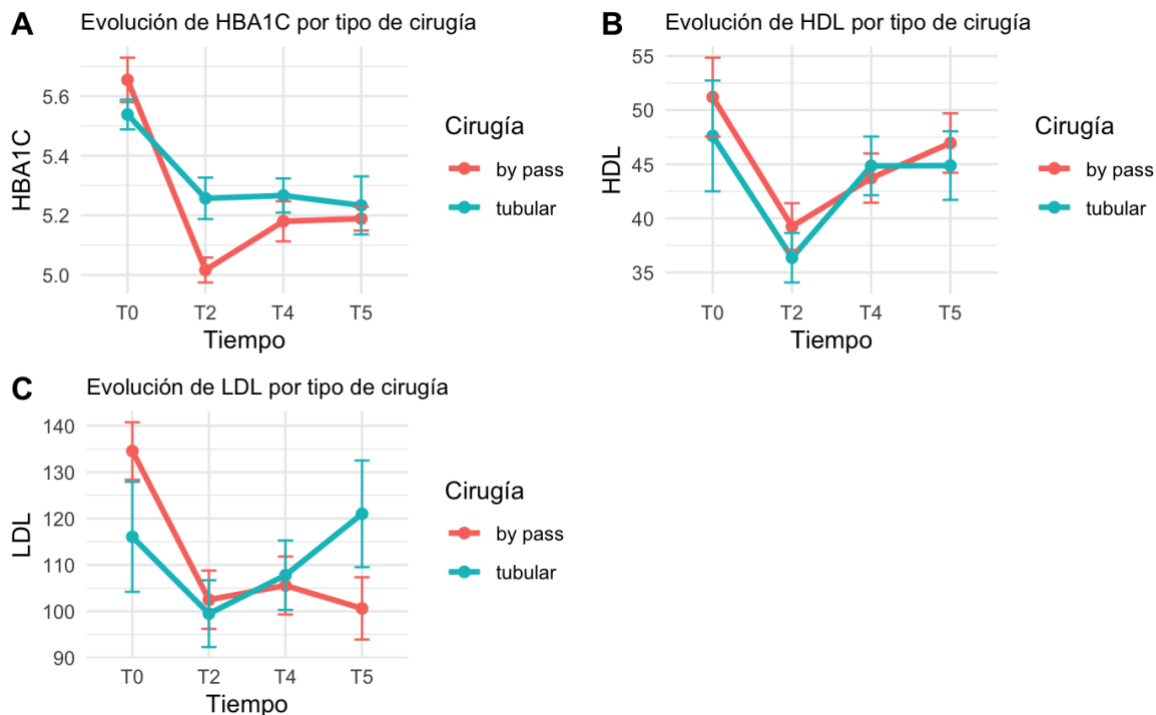


Figura 6. Evolución temporal de metabolitos clave en función de la técnica quirúrgica

- Bypass gástrico: disminución media de 34 mg/dL entre T0 y T5.
- Gastrectomía tubular: aumento medio de 2 mg/dL en el mismo periodo.

Aunque el t-test de comparación de cambios en LDL entre técnicas no alcanzó significación estadística ( $p = 0.083$ ), sí mostró una tendencia relevante desde el punto de vista clínico. Esta observación fue respaldada por un análisis de varianza (ANOVA) que reveló una interacción significativa entre el tiempo y el tipo de cirugía ( $p = 0.047$ ), lo que sugiere trayectorias de evolución distintas del LDL según la técnica empleada.

Las correlaciones entre pérdida de peso y LDL, HDL y HBA1C no fueron significativas ( $p > 0.05$ ), aunque se observó una tendencia negativa para HDL.



## 5 DISCUSIÓN

Este análisis exploratorio revela que la cirugía bariátrica tiene un impacto positivo en variables asociadas al síndrome metabólico, incluyendo peso, IMC, circunferencia abdominal y HBA1C.

Los resultados también sugieren que el tipo de técnica quirúrgica podría influir en la evolución de algunos parámetros bioquímicos, especialmente el colesterol LDL. En los pacientes sometidos a bypass gástrico se observó una disminución considerable de LDL, mientras que en aquellos con gastrectomía tubular hubo una ligera elevación. Aunque la diferencia final entre grupos no alcanzó significación estadística mediante t-test, el análisis de varianza sí reveló una interacción significativa entre el tipo de cirugía y la evolución temporal del LDL, lo que sugiere que la trayectoria de cambio difiere según la técnica empleada. Este descenso de LDL podría suponer una disminución añadida al riesgo cardiovascular más allá del efecto de la propia pérdida de peso.

La ausencia de diferencias iniciales sugiere que los efectos observados son atribuibles a la intervención y no a condiciones basales distintas.

Entre las limitaciones del estudio se encuentra el tamaño muestral reducido, especialmente en el grupo tubular, lo cual puede limitar la potencia estadística. Además, algunos metabolitos no estaban disponibles en todos los tiempos, dificultando análisis completos.

## 6 CONCLUSIONES

La cirugía bariátrica se asoció a una mejora en parámetros clave de salud metabólica, incluyendo peso, IMC, circunferencia abdominal y hemoglobina glicada.

No se encontraron diferencias claras en los perfiles metabólicos basales entre los grupos quirúrgicos, lo que respalda la validez comparativa del estudio.

Aunque no se alcanzó significación estadística en el análisis final de LDL mediante t-test, el análisis de varianza mostró que su evolución a lo largo del tiempo difiere según la técnica quirúrgica, apuntando a un posible mayor beneficio del bypass gástrico sobre el perfil lipídico.

No se identificaron correlaciones significativas entre la pérdida de peso y los niveles finales de LDL, HDL o HBA1C, lo que sugiere que otros factores podrían mediar los efectos metabólicos de la cirugía.

## 7 REFERENCIAS

1. Origen de los datos: <https://github.com/nutrimetabolomics/metaboData>

2. Enlace a github con el resto de documentos:  
<https://github.com/evaat52/Alvarez-Torres-Eva-PEC1>

## 8 ANEXO

### 8.1 Otros gráficos de interés:

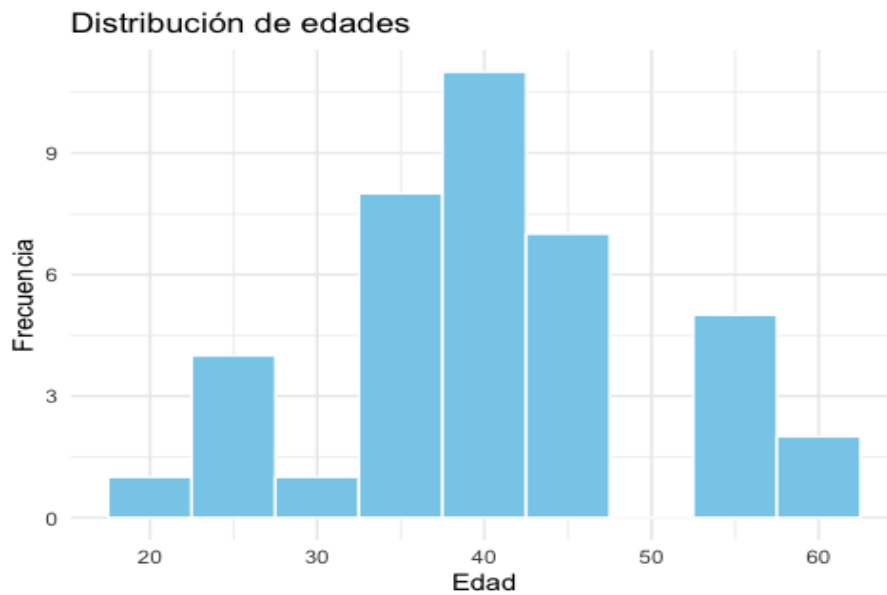


Figura 7. Histograma de edad

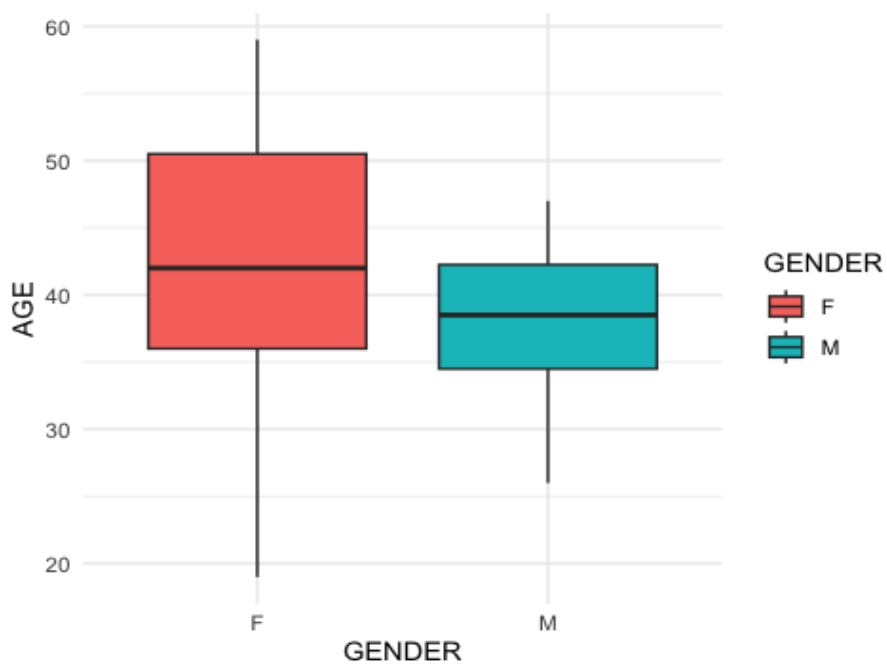


Figura 8. Distribución de edad por sexo

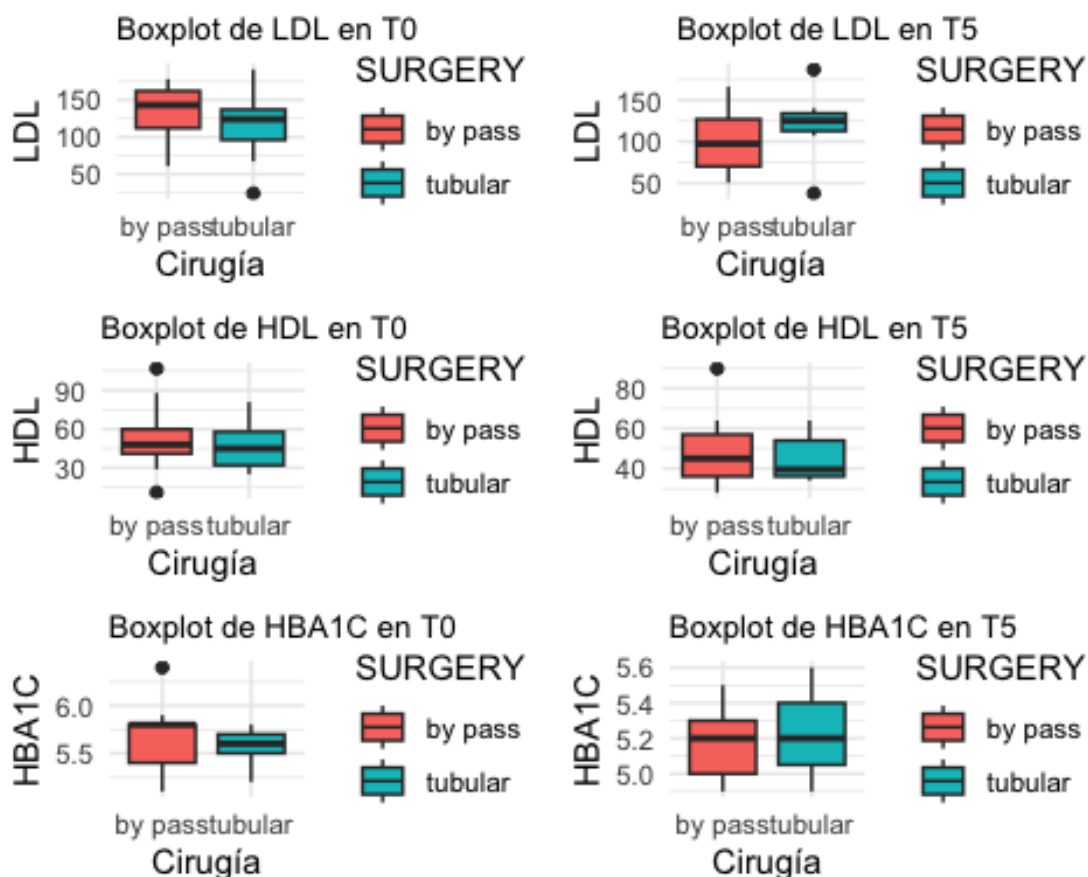


Figura 9. Metabolitos en T0 y T5 comparativos por cirugía

## 8.2 Preguntas 1, 2 y 4 de la PEC

### 8.2.1 Pregunta 1: descargar el dataset

```
library(httr) # Para descargar Los archivos
##
## Attaching package: 'httr'
##
## The following object is masked from 'package:Biobase':
##
##     content
library(readr) # Para Leer archivos CSV
library(tidyverse) # Para manipulación de datos
library(S4Vectors) # Para función DataFrame
options(repos = c(CRAN = "https://cloud.r-project.org"))
install.packages("BiocManager")
```

```
##
## The downloaded binary packages are in
## /var/folders/66/2ffnbw5x5yjcw3_td8ms7_rw0000gn/T//RtmpFAZknc/download
ed_packages

# Instalar el paquete SummarizedExperiment
BiocManager::install("SummarizedExperiment")

## 'getOption("repos")' replaces Bioconductor standard repositories, see
## 'help("repositories", package = "BiocManager")' for details.
## Replacement repositories:
## CRAN: https://cloud.r-project.org

## Bioconductor version 3.19 (BiocManager 1.30.25), R 4.4.0 (2024-04-24)

## Warning: package(s) not installed when version(s) same as or greater t
han current; use
## `force = TRUE` to re-install: 'SummarizedExperiment'

## Old packages: 'lattice', 'mgcv'

library(SummarizedExperiment)

# URL del archivo CSV en GitHub en versión raw
file_url <- "https://raw.githubusercontent.com/nutrimetabolomics/metaboDa
ta/refs/heads/main/Datasets/2018-MetabotypingPaper/DataValues_S013.csv"

# Descargar el archivo CSV
download.file(file_url, destfile = "metabolomics_dataset.csv", mode = "wb
")

# Leer el archivo CSV en R
metabolomics_data <- read.csv("metabolomics_dataset.csv")

# Para verificar la estructura del dataset usamos: str(metabolomics_data)

# Para ver las primeras filas del dataset usamos: head(metabolomics_data)

dim(metabolomics_data) # Verifica el número de filas y columnas
## [1] 39 696
```

### 8.2.2 Pregunta 2: Crear SummarizedExperiment object y concepto

```
# Extraer la información de Los pacientes
colData <- metabolomics_data[, 1:6] # Variables cualitativas
colData <- Dataframe(colData) # Convertir a Dataframe de Bioconductor
dim(colData)

## [1] 39 6

head(colData)
```

```
## DataFrame with 6 rows and 6 columns
##      X.1 SUBJECTS SURGERY AGE GENDER Group
## <integer> <integer> <character> <integer> <character> <integer>
## 1      1      1    by pass    27      F      1
## 2      2      2    by pass    19      F      2
## 3      3      3    by pass    42      F      1
## 4      4      4    by pass    37      F      2
## 5      5      5    tubular    42      F      1
## 6      6      6    by pass    24      F      2

# Extraer la matriz de datos (variables cuantitativas)
assay_data <- as.matrix(metabolomics_data[, 7:ncol(metabolomics_data)])
dim(assay_data)

## [1] 39 690

head(assay_data[, 1:5]) # Para ver Los primeros valores

##      MEDDM_T0 MEDCOL_T0 MEDINF_T0 MEDHTA_T0 GLU_T0
## [1,]      0      0      0      1      85
## [2,]      0      0      0      0      78
## [3,]      0      0      0      0      75
## [4,]      0      0      0      0      71
## [5,]      0      0      0      0      82
## [6,]      0      0      0      0      71

# Transponer la matriz de datos para que las muestras estén en filas
assay_data <- t(assay_data)

# Verificar dimensiones después de la transposición
dim(assay_data)

## [1] 690 39

# SummarizedExperiment
se <- SummarizedExperiment(
  assays = list(counts = assay_data),
  colData = colData
)
se

## class: SummarizedExperiment
## dim: 690 39
## metadata(0):
## assays(1): counts
## rownames(690): MEDDM_T0 MEDCOL_T0 ... SM.C24.0_T5 SM.C24.1_T5
## rowData names(0):
## colnames: NULL
## colData names(6): X.1 SUBJECTS ... GENDER Group
```

SummarizedExperiment funciona como un contenedor de matrices, donde las filas representan los metabolitos de interés y las columnas las observaciones. Un objeto

de esta clase puede contener uno o más assays, cada uno con múltiples matrices de datos, permitiendo almacenar diferentes tipos de mediciones.

Su principal diferencia con ExpressionSet es que este último es más antiguo y menos flexible, pensado específicamente para transcriptómica. ExpressionSet solo admite una única matriz de expresión, mientras que SummarizedExperiment permite trabajar con múltiples matrices y maneja de forma más estructurada los metadatos de filas (rowData) y columnas (colData).

### 8.2.3 Pregunta 4: guardar el SummarizedExperiment Object

```
dir.create("data", showWarnings = FALSE)
save(se, file = "SummarizedExperiment.Rda")
```