

# Búsqueda de biomarcadores bioquímicos de la enfermedad de Párkinson en muestras biológicas

**Mireia Colomer Rodriguez**

Máster Bioinformática y Bioestadística UOC-UB

Área 2. Subárea 1. Biología del desarrollo, cáncer, biología molecular y farmacología

**Ivette Olivares Castiñeira**

**Carles Ventura Royo**

05/06/2019



## FICHA DEL TRABAJO FINAL

<b>Título del trabajo:</b>	<i>Búsqueda de biomarcadores bioquímicos de la enfermedad de Párkinson en muestras biológicas</i>
<b>Nombre del autor:</b>	<i>Mireia Colomer Rodriguez</i>
<b>Nombre del consultor/a:</b>	<i>Ivette Olivares Castiñeira</i>
<b>Nombre del PRA:</b>	<i>Carles Ventura Royo</i>
<b>Fecha de entrega (mm/aaaa):</b>	06/2019
<b>Titulación::</b>	<i>Máster Bioinformática y Bioestadística</i>
<b>Área del Trabajo Final:</b>	<i>Biología del desarrollo, cáncer, biología molecular y farmacología</i>
<b>Idioma del trabajo:</b>	<i>Español</i>
<b>Palabras clave</b>	<i>Enfermedad de Párkinson, biomarcadores, análisis de datos, R.</i>

### **Resumen del Trabajo (máximo 250 palabras):**

La enfermedad de Párkinson es la segunda enfermedad neurodegenerativa más común después de la enfermedad de Alzheimer y se diagnostica a un 1% de los individuos mayores de 65 años. Actualmente, se detecta en un estadio tardío de la enfermedad una degeneración neuronal de aproximadamente el 60%. Por esto, es necesario un diagnóstico temprano de la enfermedad y es aquí donde los biomarcadores tienen un papel muy destacado. Un biomarcador es una característica que proporciona una información sobre el estado de una enfermedad y permite evaluar su evolución y los efectos de los tratamientos. Hay diferentes tipos de biomarcadores que se pueden clasificar a grandes rasgos en: clínicos, bioquímicos, genéticos y de escaneo.

Este trabajo se centra en la búsqueda de posibles biomarcadores bioquímicos que pueden estar presentes en muestras biológicas de fácil obtención como el plasma, el líquido cefalorraquídeo y la saliva. Se utilizan datos procedentes del estudio observacional BioFIND y los resultados se obtienen mediante estadística inferencial y R.

En este estudio se obtiene la importancia de la proteína alfa-sinucleína como potencial biomarcador tanto en plasma como en líquido cefalorraquídeo. Además, se observan claramente niveles más elevados de  $\beta$ 1-42 amiloide comparado con tau total y tau fosforilada. Por último, se encuentran diferencias significativas entre PINK1 y fractalquina en muestras de plasma entre los dos géneros.

Como conclusión, se destaca la importancia de la combinación de diferentes biomacadores potenciales en la enfermedad de Párkinson con el fin de optimizar el diagnóstico de la enfermedad.

**Abstract (in English, 250 words or less):**

Parkinson's disease is the second most common neurodegenerative disease after Alzheimer's disease and is diagnosed in 1% of individuals older than 65 years. Currently, at later stages of the disease is detected a neuronal degeneration of approximately 60%. Therefore, an early diagnosis of the disease is necessary, and this is where the biomarkers have a very prominent role. A biomarker is a characteristic that provides information about the state of a disease and allows to evaluate its evolution and the effects of the treatments. There are different types of biomarkers that can be classified broadly into: clinical, biochemical, genetic and scanning.

This work focuses on the search for possible biochemical biomarkers that may be present in readily available biological samples such as plasma, cerebrospinal fluid and saliva. Data from the BioFIND observational study are used and the results are obtained by inferential statistics and R.

In this study we obtain the importance of the alpha-synuclein protein as a potential biomarker in both plasma and cerebrospinal fluid samples. In addition, higher levels of amyloid  $\beta$ 1-42 compared to total tau and phosphorylated tau are clearly observed. Finally, significant differences between PINK1 and fractalkine are found between genders in plasma samples.

As a conclusion, the importance of the combination of different potential biomarkers in Parkinson's disease in order to optimize the diagnosis of the disease is highlighted.

# Índice

1. Introducción.....	1
1.1 Contexto y justificación del Trabajo.....	1
1.2 Objetivos del Trabajo.....	3
1.3 Enfoque y método seguido.....	3
1.4 Planificación del Trabajo.....	4
1.4.1 Tareas.....	4
1.4.2 Calendario.....	7
1.4.3 Hitos.....	8
1.4.4 Análisis de riesgos.....	9
1.5 Breve resumen de productos obtenidos.....	9
1.6 Breve descripción de los otros capítulos de la memoria.....	10
2. Biomarcadores bioquímicos.....	11
2.1. Alfa-sinucleína.....	11
2.1.2. Niveles alfa-sinucleína en plasma.....	13
2.1.3. Niveles alfa-sinucleína en líquido cefaloraquídeo.....	15
2.1.4. Niveles alfa-sinucleína en saliva.....	17
2.1.5. Efecto de la contaminación en sangre en los niveles de alfa-sinucleína en muestras biológicas.....	19
2.1.6. Conclusiones.....	21
2.2. Proteína $\beta$ 1-42 amiloide, tau total y tau fosforilada.....	22
2.2.1. Introducción.....	22
2.2.2. $\beta$ 1-42 amiloide.....	24
2.2.3. Tau total.....	26
2.2.4. Tau fosforilada.....	28
2.2.5. Conclusiones.....	30
2.3. Proteínas relevantes.....	30
2.3.1. Introducción.....	30
2.3.2. Epidermal growth factor (EGF).....	32
2.3.3. Apolipoprotein AI (ApoA1).....	33
2.3.4. Protein deglycase DJ-1 (DJ-1).....	33
2.3.5. Brain-derived neurotrophic factor (BDNF).....	34
2.3.6. Glial fibrillary acidic protein (GFAP).....	35
2.3.7. Fractalkine (Frac).....	35
2.3.8. Nicotinamide phosphoribosyltransferase (NAMPT).....	36
2.3.9. PTEN induced kinase 1 (PINK1).....	37
2.3.10. Leucine-rich repeat kinase 2 (LRRK).....	37
2.3.11. Conclusiones.....	38
3. Conclusiones.....	39
4. Glosario.....	40
5. Bibliografía.....	41
6. Anexos.....	43
6.1 Alfa-sinucleína en plasma.....	43
6.2 Alfa-sinucleína en líquido cefaloraquídeo.....	47
6.3 Alfa-sinucleína en saliva.....	50
6.4 Efecto contaminación hemoglobina.....	55

6.5 $\beta$ 1-42 amieloide.....	56
6.6 Tau total .....	60
6.7 Tau fosforilada.....	64
6.8 Epidermal growth factor (EGF).....	68
6.9 Apolipoprotein AI (ApoA1).....	69
6.10 Protein deglycase DJ-1 (DJ-1) .....	70
6.11 Brain-derived neurotrophic factor (BDNF) .....	71
6.12 Glial fibrillary acidic protein (GFAP).....	72
6.13 Fractalkine (Frac) .....	73
6.14 Nicotinamide phosphoribosyltransferase (NAMPT).....	74
6.15 PTEN induced kinase 1 (PINK1) .....	75
6.16 Leucine-rich repeat kinase 2 (LRRK).....	76

## **Lista de figuras**

Figura 1. Niveles alfa-sinucleína en muestras biológicas.

Figura 2. Niveles alfa-sinucleína en plasma entre PD vs Control.

Figura 3. Niveles alfa-sinucleína en plasma entre PD vs Control en hombres.

Figura 4. Niveles alfa-sinucleína en plasma entre PD vs Control en mujeres.

Figura 5. Niveles alfa-sinucleína en líquido cefalorraquídeo entre PD vs Control.

Figura 6. Niveles alfa-sinucleína en líquido cefalorraquídeo entre PD vs Control en hombres.

Figura 7. Niveles alfa-sinucleína en líquido cefalorraquídeo entre PD vs Control en mujeres.

Figura 8. Niveles alfa-sinucleína en saliva entre PD vs Control.

Figura 9. Niveles alfa-sinucleína en saliva entre PD vs Control en hombres.

Figura 10. Niveles alfa-sinucleína en saliva entre PD vs Control en mujeres.

Figura 11. Relación alfa-sinucleína vs hemoglobina en plasma.

Figura 12. Relación alfa-sinucleína vs hemoglobina en líquido cefalorraquídeo.

Figura 13. Relación alfa-sinucleína vs hemoglobina en saliva.

Figura 14. Niveles proteínas  $\beta$ 1-42 amiloide, tau total y tau fosforilada en líquido cefalorraquídeo.

Figura 15. Niveles proteína  $\beta$ 1-42 amiloide en líquido cefalorraquídeo.

Figura 16. Niveles proteína  $\beta$ 1-42 amiloide en líquido cefalorraquídeo en hombres. .

Figura 17. Niveles proteína  $\beta$ 1-42 amiloide en líquido cefalorraquídeo en mujeres.

Figura 18. Niveles proteína tau en líquido cefalorraquídeo.

Figura 19. Niveles proteína tau en líquido cefalorraquídeo en hombres.

Figura 20. Niveles proteína tau en líquido cefalorraquídeo en mujeres.

Figura 21. Niveles proteína ptau en líquido cefalorraquídeo.

Figura 22. Niveles proteína ptau en líquido cefalorraquídeo en hombres.

Figura 23. Niveles proteína ptau en líquido cefalorraquídeo en mujeres.

Figura 24. Niveles EGF entre géneros en plasma.

Figura 25. Niveles ApoA1 entre géneros en plasma.

Figura 26. Niveles DJ-1 entre géneros en plasma.

Figura 27. Niveles BDNF entre géneros en plasma.

Figura 28. Niveles GFAP entre géneros en plasma.

Figura 29. Niveles Frac entre géneros en plasma.

Figura 30. Niveles NAMPT entre géneros en plasma.

Figura 31. Niveles PINK1 entre géneros en plasma.

Figura 32. Niveles LRRK entre géneros en plasma.



# 1. Introducción

## 1.1 Contexto y justificación del Trabajo

La enfermedad de Párkinson (PD) es una enfermedad neurodegenerativa progresiva que se describió clínicamente por James Parkinson hace unos 200 años (Chen, 2014). Es la segunda enfermedad degenerativa más común en la población después de la enfermedad de Alzheimer (AD). La enfermedad de Párkinson presenta particularidades clínicas heterogéneas ya que se caracteriza por diversos síntomas motores y no motores (Kim et al, 2014).

La enfermedad de Párkinson se diagnostica a un 1% de los individuos mayores de 65 años (Miller & O'Callagan, 2015). Hasta 2005, >4 millones de personas en el mundo sufrieron esta enfermedad (Chen, 2014). La prevalencia de la enfermedad es de 0.3% del total, pero incrementa con la edad a >3% por individuos mayores de 80 años (Poewe et al, 2017). La mortalidad aumenta el doble comparado con la población general a partir de la década después del diagnóstico de la enfermedad (Poewe et al, 2017) ya que no es una enfermedad que cause directamente la muerte, pero sí que los síntomas comportan una disminución importante de la calidad de vida del individuo afectado (Kim et al, 2014).

Es una enfermedad en la que se han encontrado causas genéticas mendelianas y variantes genéticas comunes como factores de riesgo, pero mayoritariamente, la enfermedad de Párkinson se considera esporádica o idiopática que se diagnostica en fundamentos clínicos (Chen, 2014). Ciertos estudios muestran que la enfermedad de Párkinson es más común en hombres que en mujeres en algunas poblaciones. Esto puede ser debido al efecto de protección que realizan las hormonas sexuales femeninas, a un mecanismo genético-sexual asociado o a diferencias específicas entre sexos en la exposición de factores de riesgo ambientales (Poewe et al, 2017).

Esta neuropatología se caracteriza por la pérdida de neuronas dopaminérgicas en áreas específicas de la sustancia negra pars compacta y la presencia de cuerpos de Lewis, que contienen mayoritariamente la proteína alfa-sinucleína ( $\alpha$ -syn). A pesar de esto, se ha podido comprobar que solamente coinciden un 70-80% de los diagnósticos clínicos realizados durante la vida del paciente con los neuropatológicos en autopsias. Además, cuando se diagnostica la enfermedad el estado de la fisiopatología puede ser avanzado. Concretamente, se estima que, en el momento del diagnóstico con la aparición de los primeros síntomas motores, entre un 50-60% de las neuronas dopaminérgicas pueden haberse degradado (Chen, 2014 y Miller et al, 2015).

Debido a todo esto, es necesario una detección temprana de la enfermedad y es aquí donde los biomarcadores tienen un papel muy

destacado ya que pueden ser útiles para diagnosticar la enfermedad en un estadio inicial cuando la prevención es posible.

En 2001, el Instituto Nacional de Salud (*National Institutes of Health*, NIH) de Estados Unidos definió biomarcador como “una característica que es objetivamente medida y evaluada como un indicador de un proceso biológico normal, un proceso patogénico o una respuesta farmacológica a una intervención terapéutica” (Chen, 2014 y Delenclos et al, 2016). En otras palabras, un biomarcador es una molécula, un síntoma o una particularidad que proporciona información sobre el estado de salud de un individuo. Un biomarcador ideal tendría que ser sensible, reproducible, económico, no invasivo y estar muy validado (Delenclos et al, 2016).

Hay diferentes tipos de biomarcadores relacionados con PD, pero a grandes rasgos, se pueden clasificar en (Chen, 2014, Delenclos et al, 2016; Le et al, 2017; Lotankar et al, 2017; Miller & O'Callagan, 2015, Sharma et al, 2013, Tropea and Chen, 2018 y Kim et al 2014):

- Clínicos: corresponderían los síntomas motores como la bradicinesia, la rigidez muscular, la inestabilidad postural y el temblor del descanso y los síntomas no motores como los trastornos del estado de ánimo, la disfunción cognitiva, la disfunción anatómica y las anomalías del sueño.
- Bioquímicos: se obtienen de tejidos y fluidos corporales: alfa-sinucleína, apoproteína A1, urato, DJ-1,  $\beta$ 1-42 amiloide, tau total, tau fosforilada, entre otros.
- Genéticos: mutaciones en los genes: LRRK2, SNCA, PRKN, PINK1, DJ-1, entre otros.
- Imagen: imágenes obtenidas de técnicas como SPECT, PET, MRI, entre otros.

Hoy en día los más definidos son los clínicos ya que es a través de ellos por los que se diagnostica la enfermedad de Párkinson a un paciente. Este trabajo se centra en la búsqueda de posibles biomarcadores bioquímicos que pueden estar presentes en diferentes tipos de muestras biológicas de fácil obtención.

Debido a la falta de biomarcadores en la enfermedad de Párkinson en los últimos años han ido surgiendo estudios multicentro y biorepositorios a nivel internacional con el fin de poder avanzar en el conocimiento de biomarcadores que puedan reflejar la progresión de la enfermedad. Uno de ellos es el proyecto BioFIND (*Fox Investigation for New Discovery of Biomarkers in Parkinson's Disease*). BioFIND es un estudio observacional y transversal financiado por *The Michael J. Fox Foundation for Parkinson's Research* diseñado para descubrir y verificar biomarcadores en la enfermedad de Párkinson. En este estudio se obtuvieron datos que contienen información clínica e información obtenida mediante muestras de sangre, líquido cefalorraquídeo y saliva de pacientes con una enfermedad de Párkinson moderadamente avanzada y voluntarios sanos

([www.michaeljfox.org](http://www.michaeljfox.org), Kang et al, 2016). Los datos utilizados en este trabajo proceden de este estudio.

## 1.2 Objetivos del Trabajo

Los objetivos del trabajo son los siguientes:

1. Objetivo general: Identificar posibles marcadores bioquímicos en la enfermedad de Párkinson.

1.1. Objetivo específico: Comparar los niveles de  $\alpha$ -sinucleína en líquido cefalorraquídeo, plasma y saliva entre pacientes con la enfermedad de Párkinson y pacientes control y entre diferentes géneros.

1.2. Objetivo específico: Analizar el efecto de la contaminación de sangre en muestras de líquido cefalorraquídeo, plasma y saliva para la detección de  $\alpha$ -sinucleína.

2. Objetivo general: Análisis de implicación de las proteínas tau,  $\beta$ 1-42 amiloide y p-tau en la enfermedad de Párkinson.

2.1. Objetivo específico: Comparar los niveles de las proteínas tau,  $\beta$ 1-42 amiloide y p-tau181 en líquido cefalorraquídeo entre pacientes con la enfermedad de Párkinson y pacientes control y entre diferentes géneros.

3. Objetivo general: Análisis global de las proteínas con mayor relevancia en PD.

3.1. Objetivo específico: Detectar las proteínas más significativas entre diferentes sexos en PD.

## 1.3 Enfoque y método seguido

Tal y como se ha descrito en el apartado de objetivos en este proyecto hay tres objetivos generales, uno de los cuales engloba dos objetivos específicos. Para poder llegar a este fin se tiene que realizar un análisis de datos coherente, ordenado y con sentido. Por esto, se propone el proceso siguiente:

1. Definición del problema y recolección. Desarrollar la descripción general del trabajo centrando el foco de estudio y la recolección de los datos para analizar. Se ha realizado mediante la entrega de la prueba de evaluación continua 0 (PEC0) y la 1 (PEC1).

2. Limpieza de datos. Los datos obtenidos para el estudio (BioFIND) son datos crudos, es decir, observaciones de la realidad, y necesitan ser homogeneizados antes de ser analizados.

3. Exploración de datos. Aplicación de la estadística descriptiva para obtener información de las variables de estudio.
4. Análisis. Aplicación de la estadística inferencial para poder responder a los objetivos específicos del estudio.
5. Conclusión. Interpretación de resultados, obtención de conclusiones y siguientes pasos.

Debido a que es la primera vez que se realiza un análisis de datos de estas características se realiza paso a paso tal y como se ha descrito en el proceso y un objetivo detrás de otro, con el fin de profundizar en cada momento el paso del proceso. Se descarta la opción de realizar el proceso propuesto completo primero con un objetivo y luego los siguientes debido a la falta de experiencia en el análisis de datos.

El análisis de datos se realiza mediante el software libre R. Este es un lenguaje de programación centrado en el análisis estadístico. En este trabajo se ejecuta mediante el entorno RStudio.

Es importante resaltar que la búsqueda bibliográfica no se ha mencionado en el proceso. Este paso es el que se ha realizado al inicio del proyecto y está relacionado con lo descrito en el paso 1 (definición del problema y resolución). La búsqueda bibliográfica y de datos del tema de estudio es un proceso muy amplio que puede realizarse durante un largo período de tiempo gracias a la gran cantidad de artículos científicos disponibles hoy en día. Se considera que la búsqueda bibliográfica finaliza en el paso 1 ya que se estima que hay suficiente material base para la realización de este proyecto. No obstante, a medida que se avanza el trabajo, se continúan realizando búsquedas según la necesidad ya sea a nivel conceptual o a nivel estadístico y de programación en R.

## 1.4 Planificación del Trabajo

### 1.4.1 Tareas

Las tareas referentes al análisis estadístico a realizar se basan en los pasos descritos en el apartado anterior:

1. Objetivo 1.1:
  - 1.1. Limpieza de datos 1.1.1
  - 1.2. Exploración de datos 1.1.2
  - 1.3. Análisis inferencial 1.1.3
  - 1.4. Conclusiones 1.1.4
2. Objetivo 1.2:
  - 2.1. Limpieza de datos 1.2.1
  - 2.2. Exploración de datos 1.2.2

- 2.3. Análisis inferencial 1.2.3
- 2.4. Conclusiones 1.2.4
3. Objetivo 2.1:
  - 3.1. Limpieza de datos 2.1.1
  - 3.2. Exploración de datos 2.1.2
  - 3.3. Análisis inferencial 2.1.3
  - 3.4. Conclusiones 2.1.4
4. Objetivo 3.1:
  - 4.1. Limpieza de datos 3.1.1
  - 4.2. Exploración de datos 3.1.2
  - 4.3. Análisis inferencial 3.1.3
  - 4.4. Conclusiones 3.1.4
5. PEC2. Desarrollo del trabajo-Parte 1
6. PEC3. Desarrollo del trabajo-Parte 2
7. PEC4. Cierre de la memoria
8. PEC5a. Elaboración de la presentación
9. PEC5b. Defensa pública

La limpieza de datos consiste en Generar un *dataset* para cada objetivo unificando el formato de las variables y eliminando la información que no es de interés y los valores vacíos o *missing*.

La exploración de datos consiste en:

- Analizar qué variables forman parte del estudio, de qué tipo son y qué unidades tienen.
- Analizar las variables numéricas mediante la distribución, las características de centralidad y dispersión y los valores atípicos. Realización de histogramas, boxplots, tablas de frecuencias, entre otros.
- Analizar las variables categóricas mediante las proporciones y los porcentajes por grupos. Realización de tablas de contingencia y diagrama de barras, entre otros.
- Analizar los casos multivariantes mediante distribuciones por grupos (ej. histograma de densidad), relaciones entre variables numéricas (ej. correlación) y casos multivariados (ej. diagramas 3D).
- Extraer información de los gráficos realizados y obtener conclusiones de ellos.

El análisis inferencial consiste en definir qué análisis realizar considerando los resultados obtenidos en el paso anterior y las siguientes posibilidades:

- Comparación de distribuciones por grupos
- Comparación de proporciones
- Relación de variables categóricas
- Relación de variables numéricas
- Modelos predictivos

La conclusión consiste en:

- Describir los resultados obtenidos
- Interpretar y sintetizar los resultados
- Reparos de los resultados obtenidos
- Enumeración de los siguientes pasos a realizar

Por otro lado, se encuentran las tareas relacionadas con las entregas de la evaluación continuada del proyecto que se han tenido en cuenta en este apartado ya que la realización de ellas requiere un tiempo de dedicación. Corresponde a las tareas numeradas de la 5 a la 9.

Respecto al tiempo estimado de la realización de una tarea es difícil de definir debido a la poca experiencia en el ámbito de un proyecto de análisis de datos. *A priori*, se considera el objetivo 1 el que más tiempo de dedicación necesita debido a que engloba dos objetivos específicos. También, se considera que las tareas de análisis inferencial son las que requieren más tiempo debido a la poca experiencia en análisis estadísticos y el tiempo que implica escoger qué test se aplica para poder obtener un resultado de la hipótesis de interés.

En referencia al desarrollo de las entregas, la redacción de la memoria es la que se considera que requiere más tiempo debido a que engloba todo el trabajo realizado. La duración de todas las pruebas de evaluación continuada al tener una fecha límite van muy marcadas por el calendario, es decir, que después de la entrega de la memoria sólo hace falta dedicarse a la presentación y después de esta a la defensa pública. Por esto, el factor limitante del tiempo va muy marcado hasta la entrega de la memoria.

La duración se estima en semanas para facilitar la estimación del tiempo. A continuación, se muestra una tabla con la duración de las tareas:

<b>Tarea</b>	<b>Duración</b>
Limpieza de datos 1.1.1	1 semana
Limpieza de datos 1.2.1	
Limpieza de datos 2.1.1	
Limpieza de datos 3.1.1	
Exploración de datos 1.1.1	1 semana
Exploración de datos 1.2.1	
Exploración de datos 2.1.1	1 semana y ½
Exploración de datos 3.1.1	
Análisis inferencial 1.1.1	1 semana y ½
Análisis inferencial 1.2.1	
Análisis inferencial 2.1.1	1 semana

Análisis inferencial 3.1.1	1 semana
Conclusiones 1.1.1	1 semana
Conclusiones 1.2.1	
Conclusiones 2.1.1	
Conclusiones 3.1.1	
PEC2. Desarrollo fase 1	½ semana
PEC3. Desarrollo fase 2	½ semana
PEC4. Memoria	2 semanas
PEC5a. Presentación	½ semana
PEC5b. Defensa pública	½ semana

#### 1.4.2 Calendario

Considerando las tareas mostradas en el apartado anterior y las semanas del año, en la tabla siguiente se presenta la tarea que se planea realizar cada semana:

Número semana	Semana	Número tarea	Tarea
1	18/03/19- 24/03/19	1.1.1, 1.2.1, 2.1.1, 3.1.1	Limpieza de datos
2	25/03/19- 31/03/19	1.1.1 y 1.2.1	Exploración 1
3	01/04/19- 07/04/19	2.1.1 y 3.1.1	Exploración 2 y 3
4	08/04/19- 14/04/19	2.1.1 y 3.1.1	Exploración 2 y 3
		5	PEC2. Fase 1
5	15/04/19- 21/04/19*	-	-
6	22/04/19- 28/04/19	1.1.1 y 1.2.1	Análisis 1
7	29/04/19- 05/05/19	1.1.1 y 1.2.1	Análisis 1
		2.1.1	Análisis 2
8	06/05/19- 12/05/19	2.1.1	Análisis 2
		3.1.1	Análisis 3
9	13/05/19- 19/05/19	3.1.1	Análisis 3
		6	PEC3. Fase 2
10	20/05/19- 26/05/19	1.1.1, 1.2.1, 2.1.1, 3.1.1	Conclusiones
11	27/05/19- 02/06/19	7	PEC4. Memoria
12	03/06/19- 09/06/19	7	PEC4. Memoria
		8	PEC5a. Presentación
13	10/06/19- 16/06/19	8	PEC5a. Presentación
14	17/06/19- 23/06/19	9	PEC5b. Defensa

15	24/06/19- 30/06/19	9	PEC5b. Defensa
----	-----------------------	---	----------------

\*No se dispone de tiempo para dedicar al trabajo

A continuación, se muestra un diagrama de Gantt con la información de programación presentada en la última tabla pero centrado con el número de objetivo general:

Trabajo Final de Máster				Búsqueda de biomarcadores en la enfermedad de Párkinson				
Num	Tarea	Inicio	Final	marzo-19	abril-19	mayo-19	junio-19	julio-19
1	Limpieza de datos objetivo 1	18-3-19	24-3-19	■				
2	Limpieza de datos objetivo 2	18-3-19	24-3-19	■				
3	Limpieza de datos objetivo 3	18-3-19	24-3-19	■				
4	Exploración de datos objetivo 1	25-3-19	31-3-19		■			
5	Exploración de datos objetivo 2	1-4-19	14-4-19		■			
6	Exploración de datos objetivo 3	1-4-19	14-4-19		■			
7	PEC2. Desarrollo Fase 1	8-4-19	14-4-19		■			
1	Análisis inferencial objetivo 1	22-4-19	5-5-19			■		
9	Análisis inferencial objetivo 2	29-4-19	12-5-19			■		
10	Análisis inferencial objetivo 3	6-5-19	19-5-19			■		
11	PEC3. Desarrollo Fase 2	13-5-19	19-5-19			■		
12	Conclusiones objetivo 1	20-5-19	26-5-19			■		
13	Conclusiones objetivo 2	20-5-19	26-5-19			■		
14	Conclusiones objetivo 3	20-5-19	26-5-19			■		
15	PEC4. Memoria	27-5-19	5-6-19				■	
16	PEC5a. Presentación	6-6-19	13-6-19				■	
17	PEC5b. Defensa pública	17-6-19	26-6-19					■
18								

Este diagrama permite observar de manera visual qué tareas se solapan en una misma semana y es más fácil tener en cuenta las fechas de entrega de las PECs.

La duración de cada tarea durante las semanas se ha realizado teniendo en cuenta la disponibilidad de horas 'libres' (aproximadamente 20h) a la semana para realizar el trabajo considerando que se compagina con una jornada laboral de 40h además de otros quehaceres personales.

### 1.4.3 Hitos

Respecto a los hitos, en este trabajo están marcados por las entregas de las PECs y corresponden a:

- PEC2. Desarrollo del trabajo-Parte 1. Entrega 24/04/19.
- PEC3. Desarrollo del trabajo-Parte 2. Entrega 20/05/19.
- PEC4. Cierre de la memoria. Entrega 05/06/19.
- PEC5a. Elaboración de la presentación. Entrega 13/06/19.
- PEC5b. Defensa pública. Entrega 26/06/19.

Tal y cómo se ha mostrado previamente, se puede considerar que se ha programado cada bloque de tarea con una entrega de una PEC. Es decir, que después de la exploración de los datos, cronológicamente se produce una entrega, y así con el resto.



La implementación de estos hitos permite una mejor organización y planificación del desarrollo del trabajo. Por esto, es muy importante tenerlos constantemente en cuenta porque un retraso por algún imprevisto de alguna de las tareas a realizar previas al hito puede generar un retraso de la planificación importante.

#### 1.4.4 Análisis de riesgos

Respecto el análisis de riesgos, los factores más limitantes para la realización de este proyecto son el tiempo y los conocimientos que se tienen para el análisis de datos.

El primero es muy importante de tener en cuenta ya que en el momento que no se haya podido seguir la planificación establecida este va en contra, como si fuera una cuenta atrás. En el caso que esta situación se dé seguramente se tendrá que sacar tiempo de momentos personales para no perder la planificación presentada anteriormente.

El segundo hace referencia al hecho que debido a que es la primera vez que se engloban todos los conocimientos adquiridos durante el máster el proceso de realización de las tareas (limpieza, exploración, análisis y conclusiones) será mucho más lento comparado con una persona con conceptos de análisis de datos muy adquiridos. Además, este factor está relacionado con los posibles problemas que se pueden encontrar en los datos y en la ejecución de su análisis. Si esto sucede implica un retraso considerable de la planificación, pero es importante mantener el orden de tareas establecido para llegar al objetivo final.

Es imprescindible tener en cuenta que estos factores no influyan en la calidad del desarrollo del trabajo ya que este aspecto que es fácil de apartar en situaciones de alta presión donde el tiempo es limitado.

La planificación se ha generado de manera ordenada y lógica con el fin de minimizar los posibles obstáculos pero la realización de un trabajo de este tipo siempre presenta algún impedimento. Es importante destacar un posible punto débil en la planificación referente a la redacción de la memoria. Se estimaron dos semanas pero por *planning* el tiempo final que se ha destinado es de una semana y media aproximadamente. Se espera que la realización de la PEC2 y PEC3 permitan facilitar la redacción de la memoria y así poderla realizar en una semana y media.

Finalmente, es importante destacar que esta planificación se ha realizado al inicio del proyecto. Y como era de esperar, debido a algún obstáculo encontrado y comentado en PECs anteriores se ha modificado el orden de dos tareas, en concreto, la exploración de datos del objetivo 3 se ha realizado después de la redacción de la PEC2.

#### 1.5 Breve sumario de productos obtenidos

Los productos obtenidos son los siguientes:

- Memoria de TFM. Este mismo documento donde se presentan los resultados obtenidos del análisis estadístico.
- Presentación virtual del TFM

#### 1.6 Breve descripción de los otros capítulos de la memoria

Los capítulos siguientes de la memoria que se presentan se han organizado considerando los tres objetivos generales, es decir, por las proteínas de interés de cada objetivo.

Cada parte de cada capítulo se estructura globalmente de la manera siguiente:

- Introducción. Se presentan las proteínas correspondientes y se comentan resultados de los estudios científicos hasta ahora. Además, se explica de donde provienen los datos utilizados en el análisis.
- Resultados obtenidos. No se encuentra un capítulo nombrado como tal ya que se divide en función de las proteínas protagonistas del capítulo. En este apartado, se presentan una parte de los resultados descriptivos e inferenciales. No se pueden presentar todos los resultados obtenidos a lo largo del trabajo debido a la limitación del espacio.

Respecto la parte descriptiva durante el trabajo se realizan diferentes gráficos como diagramas de barra, histogramas, funciones de densidad y boxplots junto con un resumen numérico de los datos (número de muestra, media, mediana y desviación estándar, entre otros)

En referencia a la parte inferencial, se estudia la normalidad de cada grupo de datos mediante un qqplot y un test de Shapiro junto con un test de igualdad de varianzas como el test de Barlett o Levene en función de si los datos siguen una distribución normal o no, respectivamente. Para acabar el estudio inferencial, se realiza un T-test o un test de Wilcoxon en función de si los datos siguen una distribución normal o no y si tienen igualdad de varianzas con el fin de determinar si se encuentran diferencias significativas entre grupos. En todos los casos se ha considerado un nivel de significación del 95%. A parte de esto, en un caso, se realiza un análisis de correlación utilizando el coeficiente de Spearman.

- Conclusiones. Se comentan y analizan los resultados obtenidos en cada apartado.

## 2. Biomarcadores bioquímicos

### 2.1. Alfa-sinucleína

#### 2.1.1. Introducción

La  $\alpha$ -syn es la proteína mayoritaria que se encuentra en las formaciones de cuerpos de Lewis (LB) en pacientes diagnosticados con PD y es por esto por lo que es la proteína más estudiada por su potencialidad como biomarcador bioquímico. Es una proteína constituida por 140 aminoácidos y de aproximadamente 16 KDa. Está asociada a la patología esporádica y a la vez familiar ya que es codificada por el gen *SNCA*. Una mutación de este gen causa la patología monogénica autosómica dominante de PD (Kim et al, 2014).

Se han encontrado diferentes formas de  $\alpha$ -syn (monomérica, oligomérica, fibrilar, total y fosforilada) en varios fluidos y tejidos humanos como el líquido cefalorraquídeo (CSF), sangre, saliva, entre otros (Chen, 2014).

En plasma se han observado algunos estudios donde los niveles de  $\alpha$ -syn en un grupo diagnosticado con PD son elevados respecto al grupo control. No obstante, en otros estudios, se ha visto lo contrario, niveles más bajos de  $\alpha$ -syn en el grupo PD comparado con el grupo control, además de no encontrar diferencias entre género. También se ha encontrado en plasma entre pacientes con PD tratados y no tratados niveles similares de  $\alpha$ -syn, aparte de tener los niveles de la proteína más elevados que el grupo control. Respecto a estudios en plasma de diferentes formas de  $\alpha$ -syn, no se han encontrado diferencias significativas entre las formas oligoméricas, totales y fosforiladas de  $\alpha$ -syn (Malek et al, 2014).

CSF es el fluido con el que se ha estudiado más la  $\alpha$ -syn pero los resultados obtenidos no son consistentes ya que los niveles de  $\alpha$ -syn en CSF en unos estudios encuentran diferencias significativas entre el grupo PD respecto al grupo control mientras que en otros estudios no (Kim et al, 2014). Se ha considerado que estas diferencias podrían haberse dado debido a la variabilidad de los ensayos o la población de estudio, entre otros aspectos (Hong et al, 2010 y Chen, 2014). Además, se ha observado que, mediante el proceso de obtención de CSF, este puede contaminarse de sangre, provocando así, un efecto considerable en los niveles obtenidos de  $\alpha$ -syn ya que los niveles de esta proteína en sangre son más elevados que en CSF (Hong et al, 2010, Shi et al, 2011).

Respecto en saliva, hay menos estudios realizados comparado con los otros dos fluidos comentados anteriormente. Hasta ahora, no se han encontrado diferencias significativas entre pacientes con PD respecto paciente control en los niveles de  $\alpha$ -syn en saliva pero sí que se ha observado que esta proteína en saliva aumenta con la edad, cosa que podría ser interesante para monitorizar el progreso de PD (Kang et al, 2016). Asimismo, se ha detectado  $\alpha$ -syn en el cálculo dental, a pesar de que también se contamina con sangre, pero no se han detectado diferencias considerables entre los pacientes con PD y los controles (Schmid et al, 2018).

Este apartado intenta responder a las preguntas de los objetivos específicos 1.1 y 1.2 con el fin de intentar contribuir a los resultados obtenidos hasta ahora en referencia a la  $\alpha$ -syn en pacientes con PD. La primera parte de este apartado, se analizan los niveles de  $\alpha$ -syn en tres fluidos corporales. Para realizarlo se han utilizado los datos del proyecto número 110 del estudio BioFIND cuyo objetivo es determinar la concentración de  $\alpha$ -syn en muestras de plasma, CSF y saliva mediante un ensayo de ELISA. En la segunda parte, se estudian los efectos de la contaminación de sangre en los mismos fluidos corporales que en la primera parte. Los datos utilizados corresponden al proyecto 111 del estudio BioFIND cuyo objetivo es determinar la concentración de hemoglobina en plasma CSF y saliva mediante un ensayo de ELISA. Estos dos proyectos mencionados, al formar parte de un mismo estudio, corresponden a los mismos pacientes. Para simplificar el análisis se analiza cada tipo de muestra por separado ya que los niveles de  $\alpha$ -syn entre los tres tipos de muestra son diferentes, sobretodo en el caso de plasma (Figura 1).

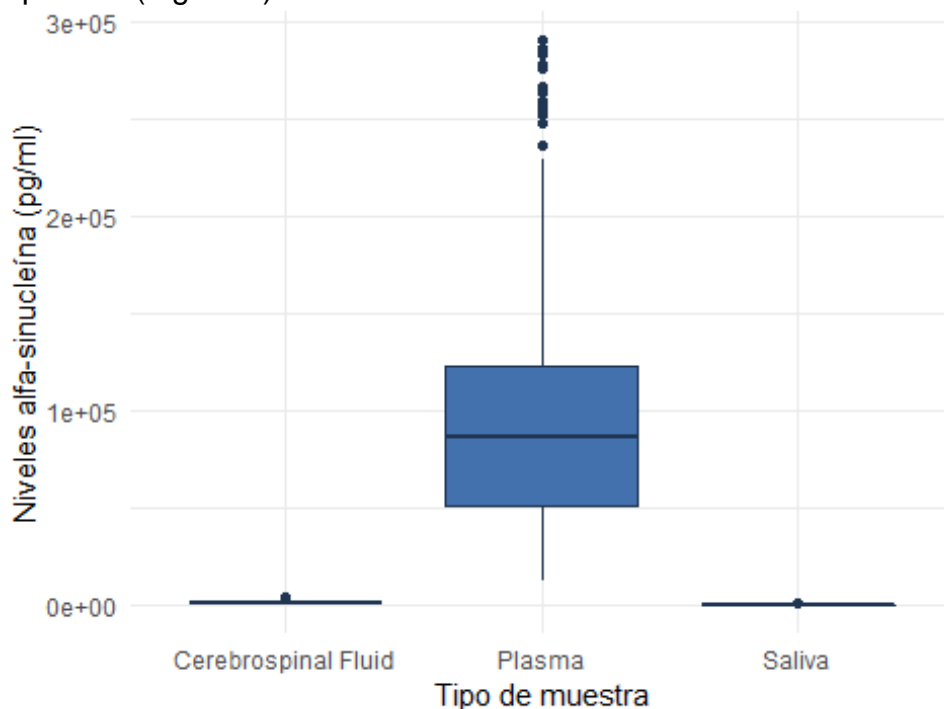


Figura 1. Niveles alfa-sinucleína en muestras biológicas.

### 2.1.2. Niveles alfa-sinucleína en plasma

En plasma, se observan niveles más elevados en el grupo control que en el grupo PD (Figura 2). Se obtiene un p-valor de 0.01, por lo tanto, se observan diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control.

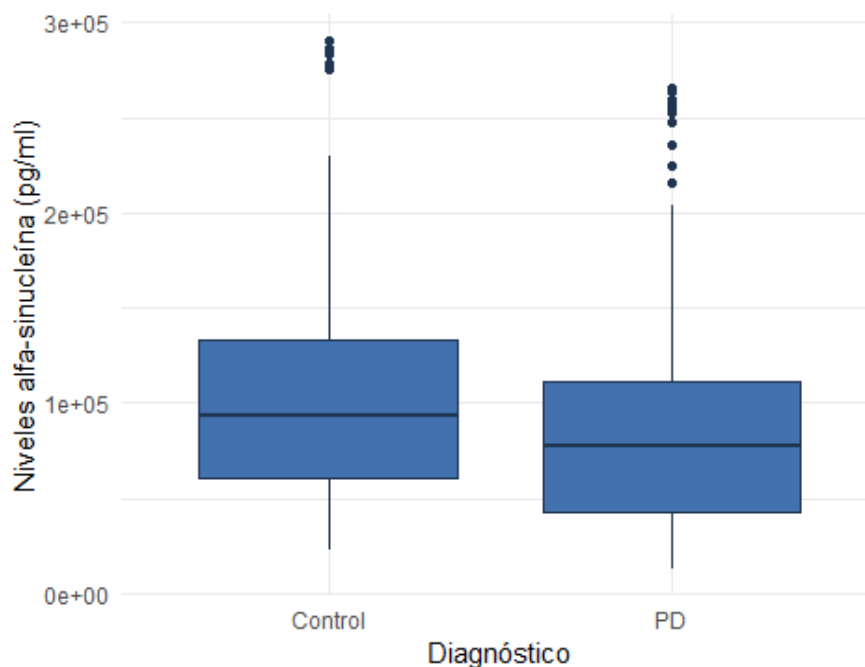


Figura 2. Niveles alfa-sinucleína en plasma entre PD vs Control.

En hombres, se obtienen niveles de  $\alpha$ -syn más elevados en el grupo control respecto al grupo PD (Figura 3). Se obtiene un p-valor de 0.43, entonces se determina que no hay diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control en hombres.

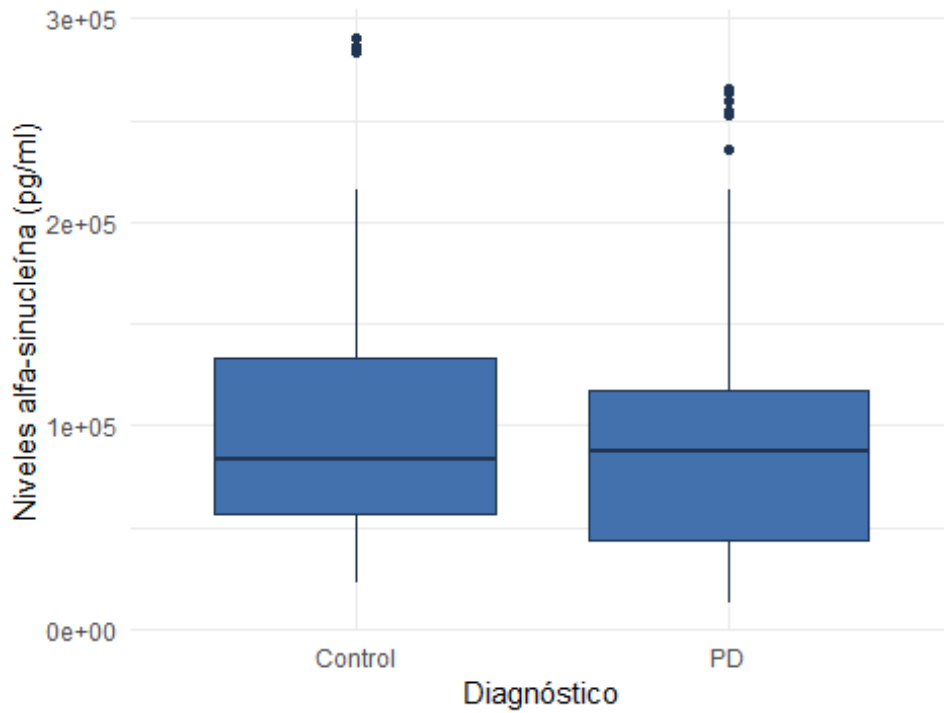


Figura 3. Niveles alfa-sinucleína en plasma entre PD vs Control en hombres.

En mujeres, se obtienen niveles de  $\alpha$ -syn más elevados en el grupo control respecto el grupo PD (Figura 4). Se obtiene un p-valor de 0.00, así se determina que hay diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control en mujeres.

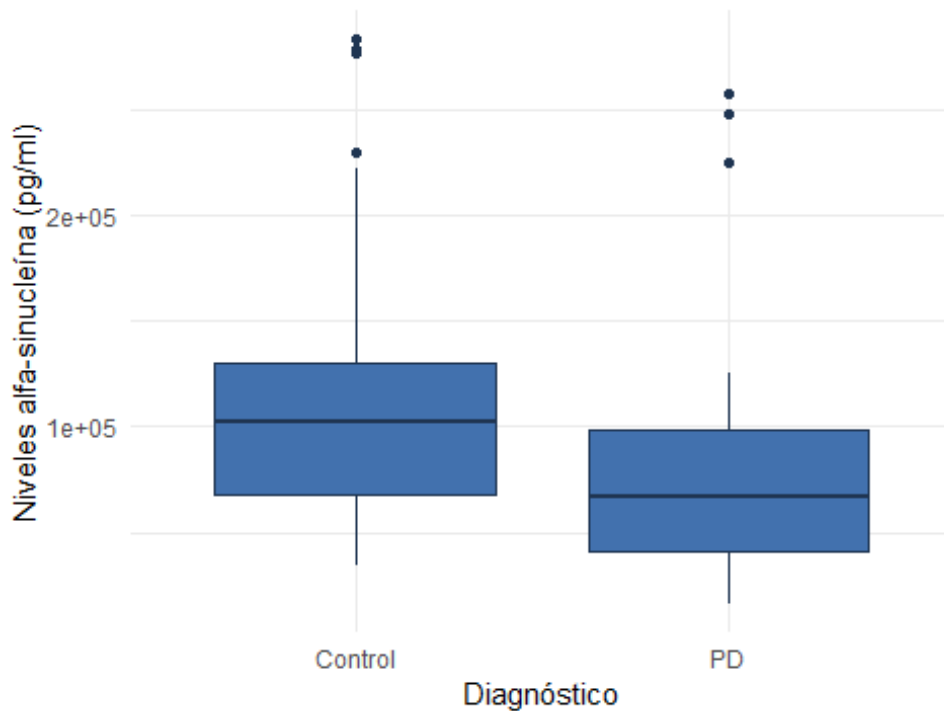


Figura 4. Niveles alfa-sinucleína en plasma entre PD vs Control en mujeres.

### 2.1.3. Niveles alfa-sinucleína en líquido cefaloraquídeo

En CSF, se observan niveles más elevados en el grupo control que en el grupo PD (Figura 5). Se obtiene un p-valor de 0.00, por lo tanto, se observan diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control.

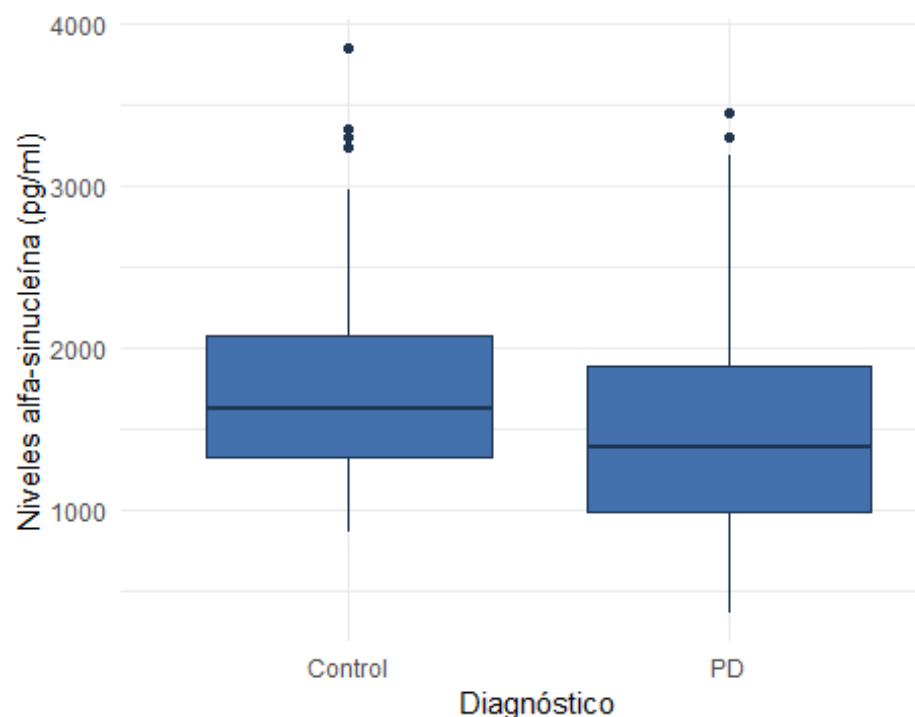


Figura 5. Niveles alfa-sinucleína en líquido cefaloraquídeo entre PD vs Control.

En hombres, se obtienen niveles de  $\alpha$ -syn más elevados en el grupo control respecto al grupo PD (Figura 6). Se obtiene un p-valor de 0.00, entonces se determina que hay diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control en hombres.

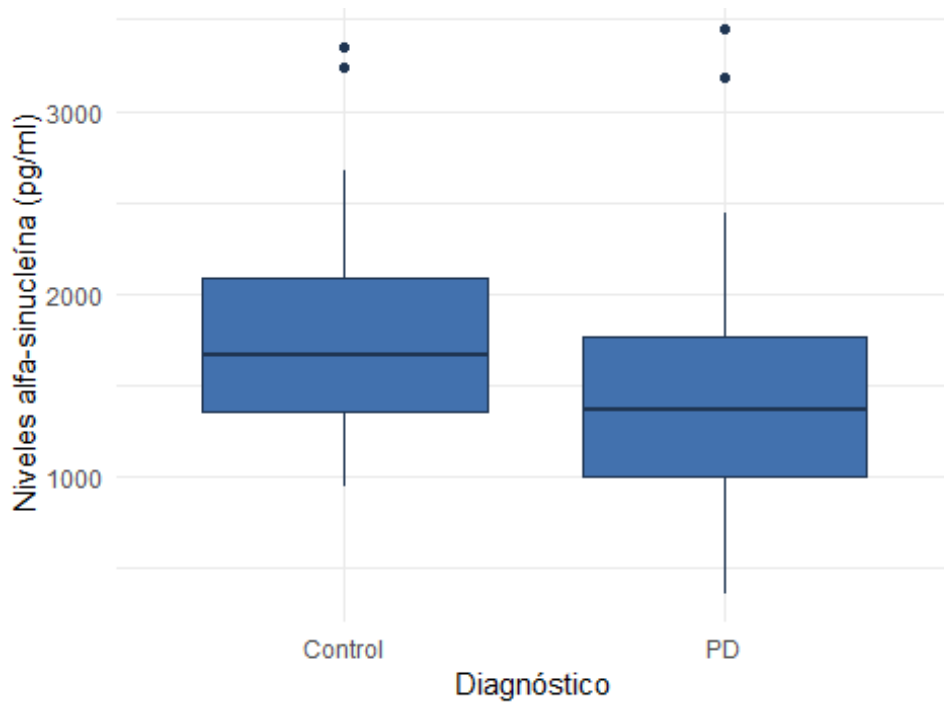


Figura 6. Niveles alfa-sinucleína en líquido cefaloraquídeo entre PD vs Control en hombres.

En mujeres, se obtienen niveles de  $\alpha$ -syn algo más elevados en el grupo control respecto el grupo PD (Figura 7). Se obtiene un p-valor de 0.08, así se determina que no hay diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control en mujeres.

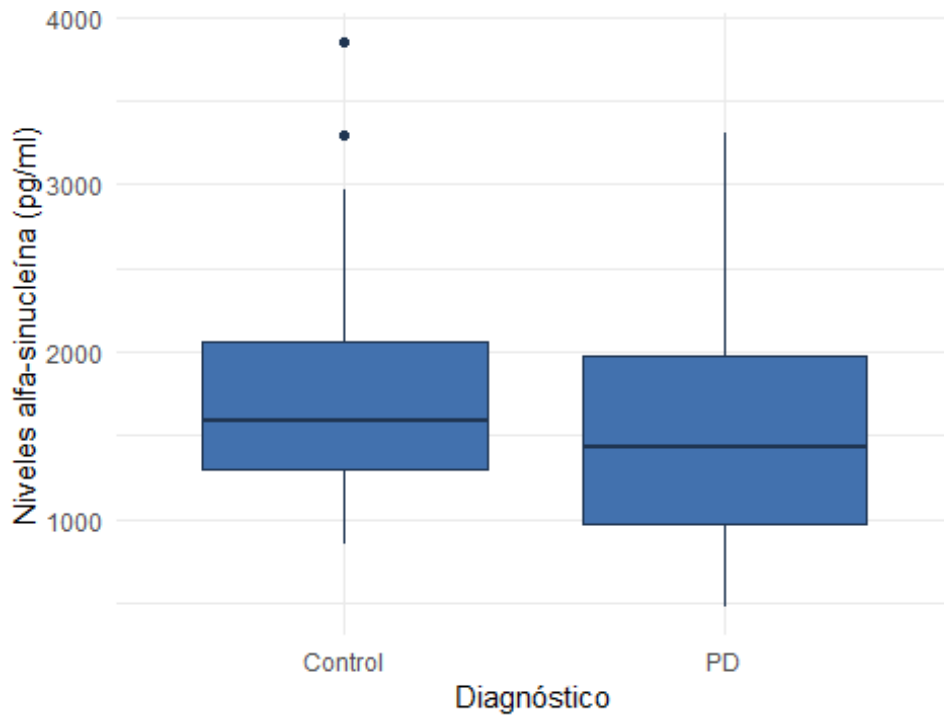


Figura 7. Niveles alfa-sinucleína en líquido cefaloraquídeo entre PD vs Control en mujeres.



#### 2.1.4. Niveles alfa-sinucleína en saliva

En saliva, se observan niveles más elevados en el grupo PD que en el grupo control (Figura 8). Se obtiene un p-valor de 0.04. Es un valor muy cercano a 0.05, así que para no arriesgar se considera que se observan diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control.

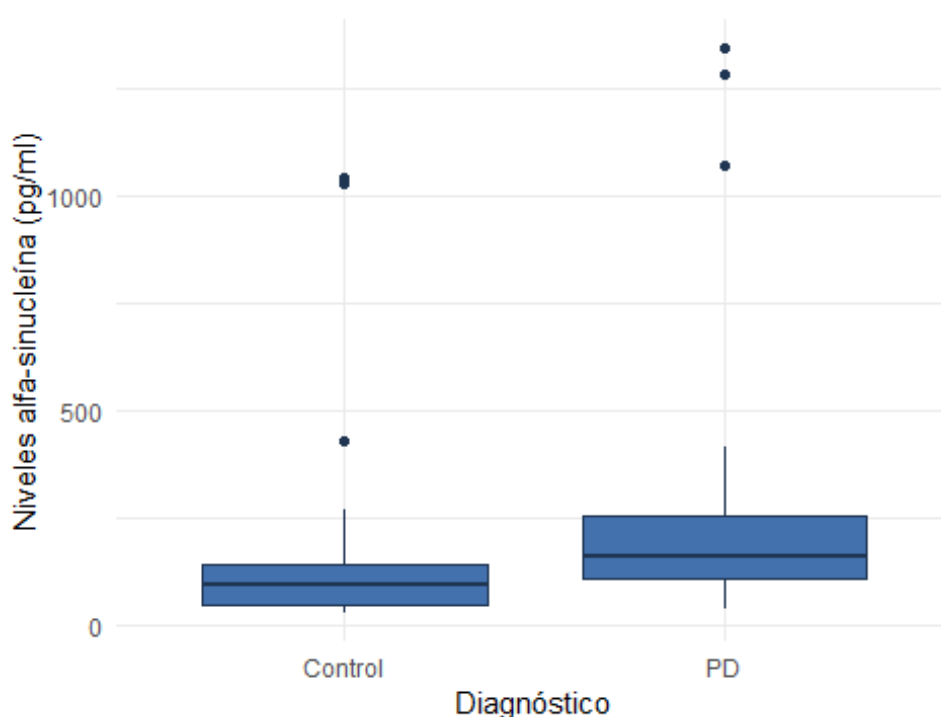


Figura 8. Niveles alfa-sinucleína en saliva entre PD vs Control.

En hombres, se obtienen niveles de  $\alpha$ -syn más elevados en el grupo PD respecto el grupo control (Figura 9). Se obtiene un p-valor de 0.24, con lo que se determina que no hay diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control en hombres.

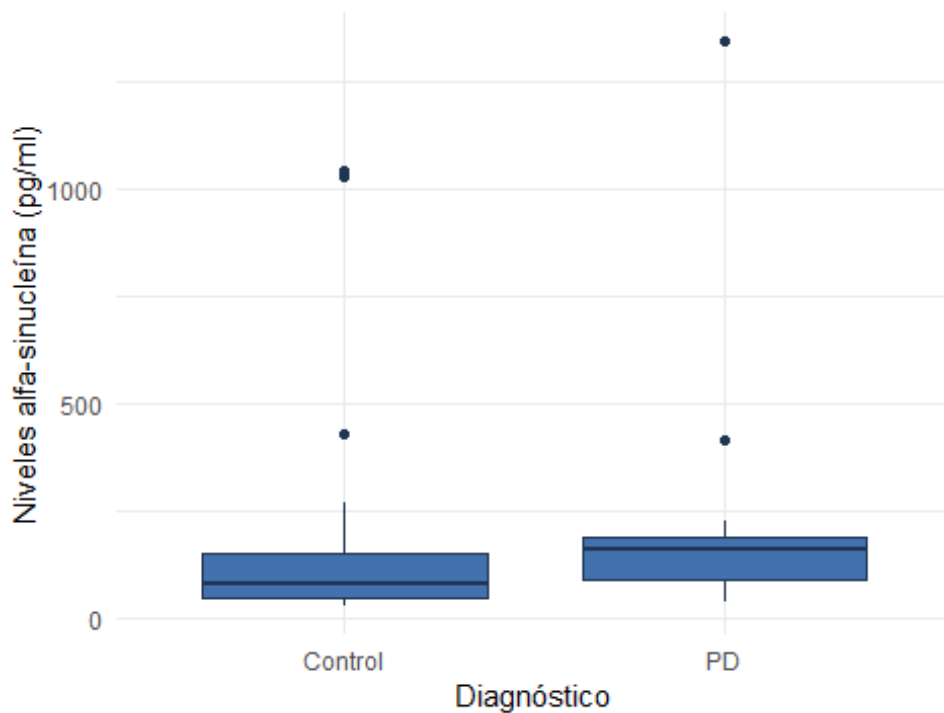


Figura 9. Niveles alfa-sinucleína en saliva entre PD vs Control en hombres.

En mujeres, se obtienen niveles de  $\alpha$ -syn algo más elevados en el grupo PD respecto el grupo control (Figura 10). Se obtiene un p-valor de 0.09, así se determina que no hay diferencias significativas en los niveles de  $\alpha$ -syn entre el grupo PD y el grupo control en mujeres.

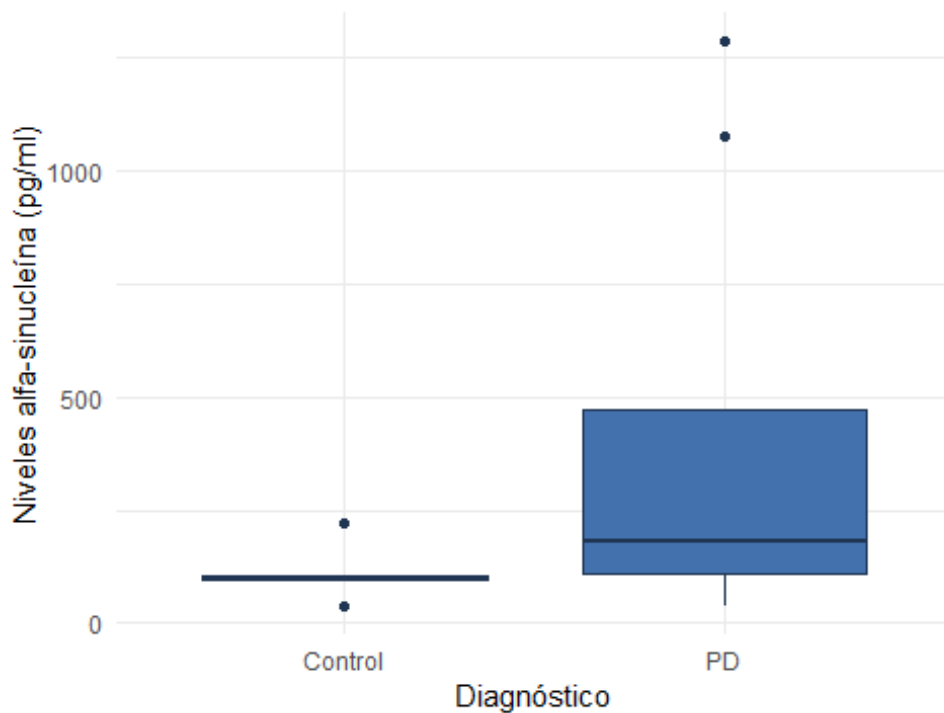


Figura 10. Niveles alfa-sinucleína en saliva entre PD vs Control en mujeres.

### 2.1.5. Efecto de la contaminación en sangre en los niveles de alfa-sinucleína en muestras biológicas

Para analizar el efecto de contaminación en sangre se analiza la relación entre los niveles de hemoglobina detectados en los tres tipos de muestra: plasma, CSF y saliva mediante un análisis de correlación.

En plasma, se observa en el gráfico de la relación entre las dos proteínas de estudio bastante dispersión cosa que indicaría una relación débil (Figura 11). El coeficiente de Spearman obtenido es de 0.02 indicando una relación positiva muy débil debido a que el valor es cercano a 0.

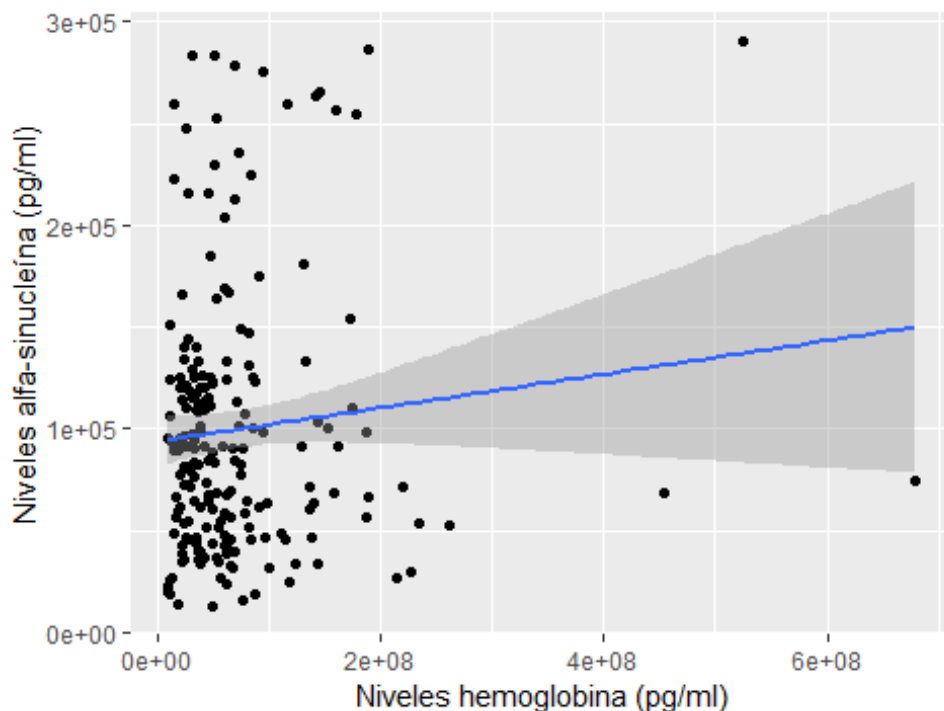


Figura 11. Relación alfa-sinucleína vs hemoglobina en plasma.

En CSF, se observa en el gráfico de la relación entre las dos proteínas de estudio poca dispersión cosa que indicaría una relación entre débil y fuerte (Figura 12). El coeficiente de Spearman obtenido es de 0.10 indicando una relación positiva y débil debido a que el valor es más cercano a 0 que a 1. Según la bibliografía, se considera que una muestra de CSF está contaminada de sangre si la concentración de hemoglobina es igual o superior a 200ng/ml (Hong et al, 2010). Por esto, en los datos de CSF, se descartan las muestras con igual o mayor concentración de 200ng/ml de

hemoglobina y se realiza de nuevo el contraste por si se observa alguna modificación.

En el contraste referente a los niveles de  $\alpha$ -syn entre grupo PD y grupo control se ha obtenido un p-valor de 0.01 determinando diferencias significativas entre ambos grupos. Respecto a los hombres, el p-valor obtenido es de 0.06, confirmando así, las no diferencias significativas entre grupo de hombres PD y grupo hombres control. Se observa lo mismo con las mujeres, se obtiene un p-valor de 0.07 y se determina que no se encuentran diferencias significativas entre el grupo mujer PD y el grupo mujer control.

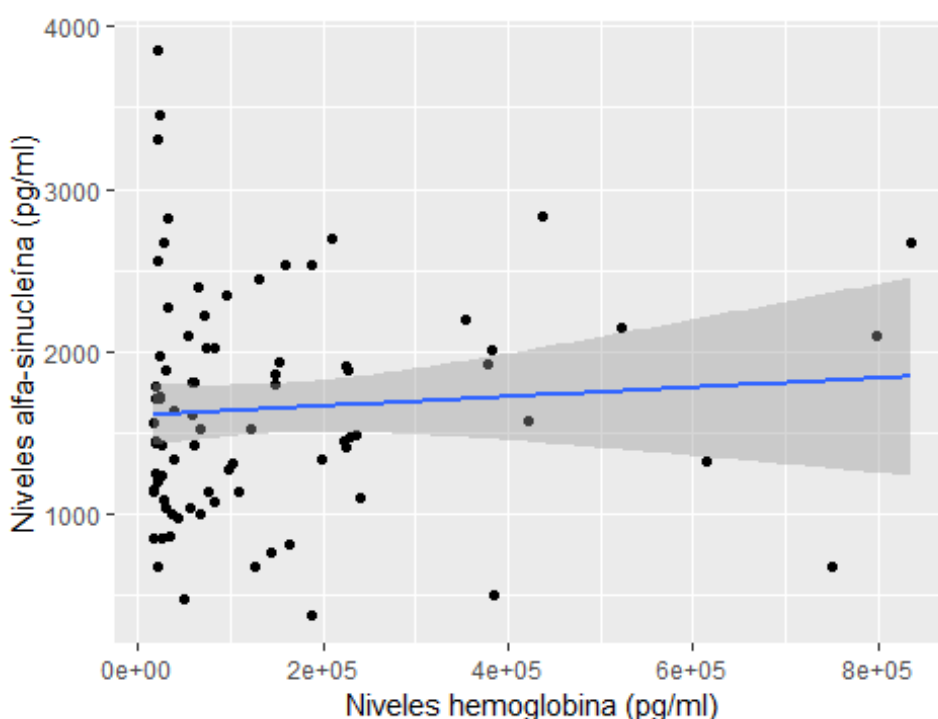


Figura 12. Relación alfa-sinucleína vs hemoglobina en líquido cefalorraquídeo.

En saliva, se observa en el gráfico de la relación entre las dos proteínas de estudio muy poca dispersión cosa que indicaría una relación fuerte (Figura 13). El coeficiente de Spearman obtenido es de 0.40 indicando una relación positiva y moderada. En este caso, en la bibliografía se considera que el umbral de contaminación en saliva es de 4000ng/ml de hemoglobina (Schmid et al, 2018). Se analiza si en los datos se encuentran valores superiores a este umbral de hemoglobina y el resultado obtenido es sin éxito.

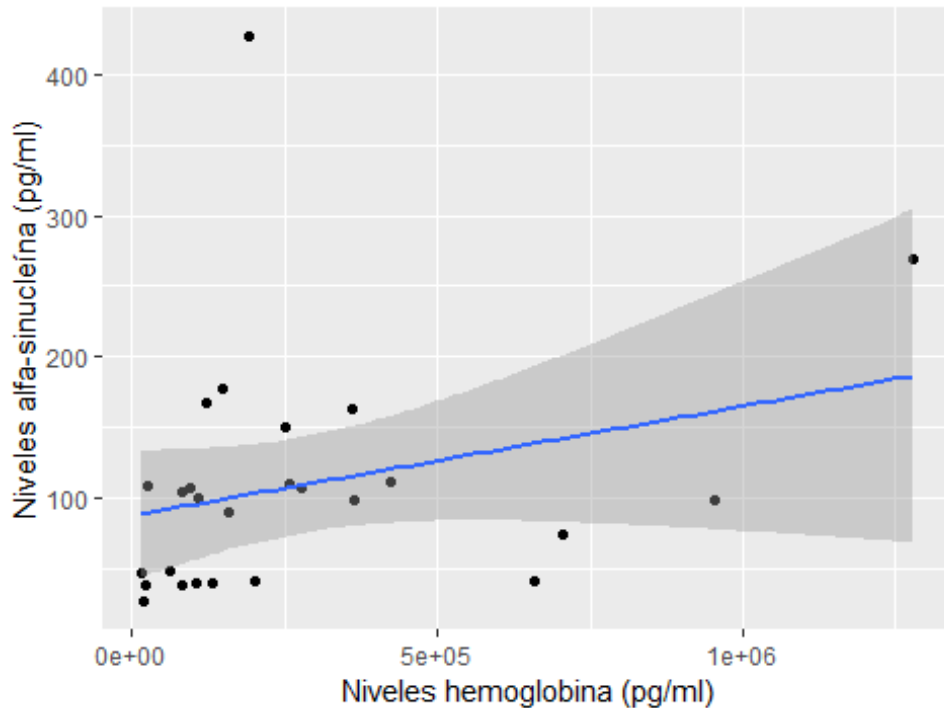


Figura 13. Relación alfa-sinucleína vs hemoglobina en saliva.

#### 2.1.6. Conclusiones

Respecto a los niveles de  $\alpha$ -syn, tanto en plasma como en CSF, los niveles de esta proteína se encuentran menores en el grupo PD que en el grupo control y la diferencia entre ambos grupos es significativa. Contrariamente, en saliva, el grupo PD contiene niveles más elevados respecto al grupo control pero la diferencia entre ambos grupos no es significativa. Es importante tener en cuenta que los tamaños muestrales son más elevados en las muestras de plasma y CSF, 92 y 84 respectivamente, comparado con las muestras de saliva, 22. En referencia al género, en hombres no se ha encontrado diferencia significativa en plasma, pero sí en CSF. Y en las mujeres, se ha observado al revés: diferencia significativa en plasma, pero no en CSF. En ambos géneros, en saliva no se han encontrado diferencias significativas.

Todos estos resultados obtenidos coinciden con algún estudio (Malek et al, 2014) ya publicado y comentado en la introducción de este apartado del trabajo. Estos generan aún más inconsistencia respecto a la potencialidad de la  $\alpha$ -syn como biomarcador de PD ya que en este caso se observan diferencias significativas entre grupo PD y control en plasma y CSF pero no en saliva. En referencia al género, el resultado obtenido es muy difuso y es difícil sacar una conclusión.

En cuanto a la contaminación de sangre de las muestras biológicas en plasma se ha encontrado una relación muy débil. Se considera que tiene su lógica ya que el plasma proviene de la centrifugación de la sangre.

En CSF, se ha detectado también una relación débil entre la  $\alpha$ -syn y la hemoglobina. Se han eliminado las muestras contaminadas con hemoglobina, considerando los datos de la bibliografía (Hong et al, 2010), y no ha variado mucho el contraste final ya que la diferencia entre grupo PD y grupo control en CSF continúa siendo significativo. Solamente, por el género masculino se ha perdido la significancia.

En saliva, la correlación obtenida ha sido elevada. En el estudio mencionado anteriormente, el umbral de 4000ng/ml de hemoglobina se detecta en muestras de cálculo o sarro dental (Schmid et al, 2018). En los datos analizados no se encuentra ningún valor igual o superior a este umbral y esto seguramente se debe al hecho que durante el proceso de extracción del sarro dental se ha obtenido alguna muestra contaminada de sangre. A pesar de esto, la alta correlación obtenida podría relacionarse con la edad (Kang et al, 2016).

Por toda esta variabilidad de resultados obtenidos, es fundamental destacar la importancia del proceso de extracción de las muestras y la de su manipulación con el fin de poder obtener datos procedentes de las muestras que sean comparables. Asimismo, las técnicas mediante las cuales se detectan las proteínas de interés, en este caso,  $\alpha$ -syn, son muy relevantes en el proceso. De esta manera, es posible poder obtener conclusiones significativas.

## 2.2. Proteína $\beta$ 1-42 amiloide, tau total y tau fosforilada

### 2.2.1. Introducción

Las proteínas  $\beta$ 1-42 amiloide (Abeta), tau total (tau) y tau fosforilada (ptau) en CSF son biomarcadores bioquímicos utilizados para diagnosticar la enfermedad de Alzheimer (AD). De manera general, los pacientes que se diagnostican con AD presentan niveles bajos de Abeta combinados con niveles elevados de tau y ptau (Shi et al, 2011) en CSF.

Esas proteínas son de interés para el estudio del diagnóstico de PD ya que pacientes con PD tienen seis veces más riesgo de desarrollar demencia comparado con individuos sin PD diagnosticado. Además, estudios *post-mortem*, han demostrado que el cerebro de un individuo con PD y con demencia contiene un 32-44% de similitud con la patología de AD (Kim et al, 2014).

A nivel molecular, se ha observado cómo se produce una interacción entre la  $\alpha$ -syn y estas tres proteínas puesto que se ha encontrado evidencia que Abeta puede inducir la agregación de  $\alpha$ -syn, y viceversa. Encima,  $\alpha$ -syn aumenta la fosforilación de tau y su inclusión en las neuronas (Kim et al, 2014).

Los resultados obtenidos con pacientes con PD comparado con individuos control centrados en estas tres proteínas no son muy sólidos ya que se

han encontrado niveles de Abeta, tau y ptau más bajos significativamente en CSF en el grupo PD comparado con el grupo control pero no todos los datos estudiados confirman estas diferencias (Kang et al, 2013). También, se ha encontrado que Abeta en CSF tiene niveles menores en pacientes con PD que presentan un deterioro cognitivo más rápido comparado con pacientes con niveles normales de Abeta en CSF (Chen, 2014 y Kim et al, 2014). Así, en otros estudios, no se ha detectado diferencia de esta proteína entre grupo diagnóstico, pero sí que se ha demostrado una disminución de Abeta en pacientes con PD con una rápida degradación cognitiva (Chen, 2014).

Otro enfoque interesante que se ha considerado es la combinación de posibles biomarcadores comunes de AD y PD con el fin de poder diferenciar el diagnóstico entre estas dos enfermedades. Esto se ha estudiado relacionado Abeta, tau, ptau,  $\alpha$ -syn y otras proteínas entre ellas (Shi et al 2011).

Este apartado intenta responder a la pregunta del objetivo específico 2.1 con el fin de intentar contribuir a los resultados obtenidos hasta ahora en referencia a Abeta, tau y ptau en pacientes con PD. Para realizarlo se han utilizado los datos del proyecto número 109 del estudio BioFIND cuyo objetivo es analizar las tres proteínas en CSF mediante la tecnología Luminex xMAP.

Con estos datos se ha detectado como los niveles de Abeta son más elevados que los de tau o ptau (Figura 14).

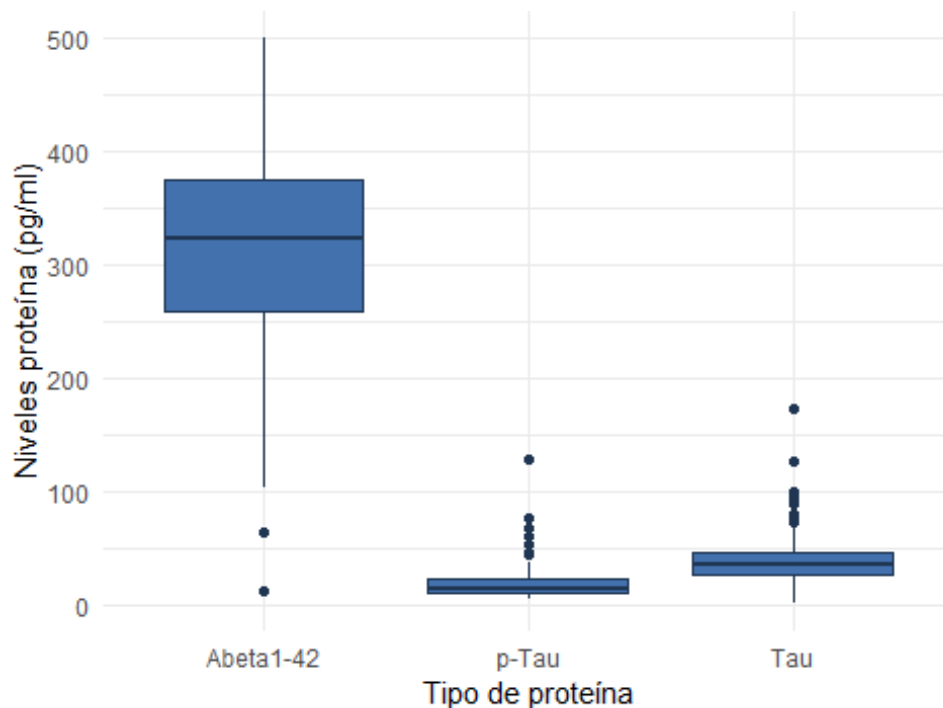


Figura 14. Niveles proteínas  $\beta$ 1-42 amiloide, tau total y tau fosforilada en líquido cefalorraquídeo.

### 2.2.2. $\beta$ 1-42 amiloide

Se observan niveles más elevados de Abeta en el grupo control que en el grupo PD (Figura 15). Se obtiene un p-valor de 0.00, por lo tanto, se detectan diferencias significativas en los niveles de Abeta entre el grupo PD y el grupo control.

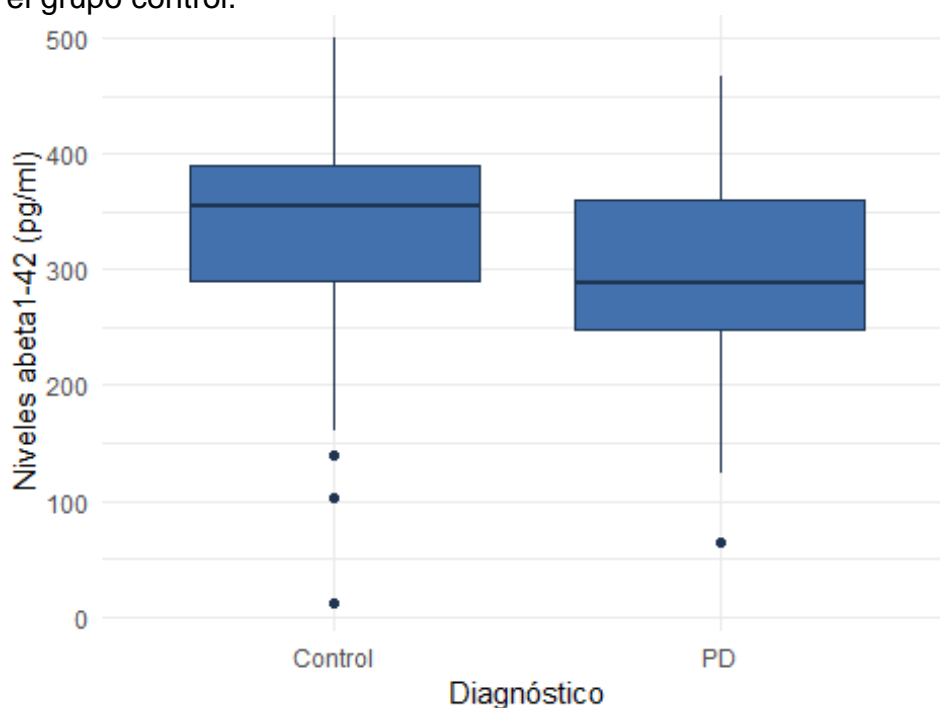


Figura 15. Niveles proteína  $\beta$ 1-42 amiloide en líquido cefalorraquídeo.

En hombres, se observan niveles más elevados de Abeta en el grupo control que en el grupo PD (Figura 16). Se obtiene un p-valor de 0.00, por lo tanto, se detectan diferencias significativas en los niveles de Abeta entre el grupo PD y el grupo control.



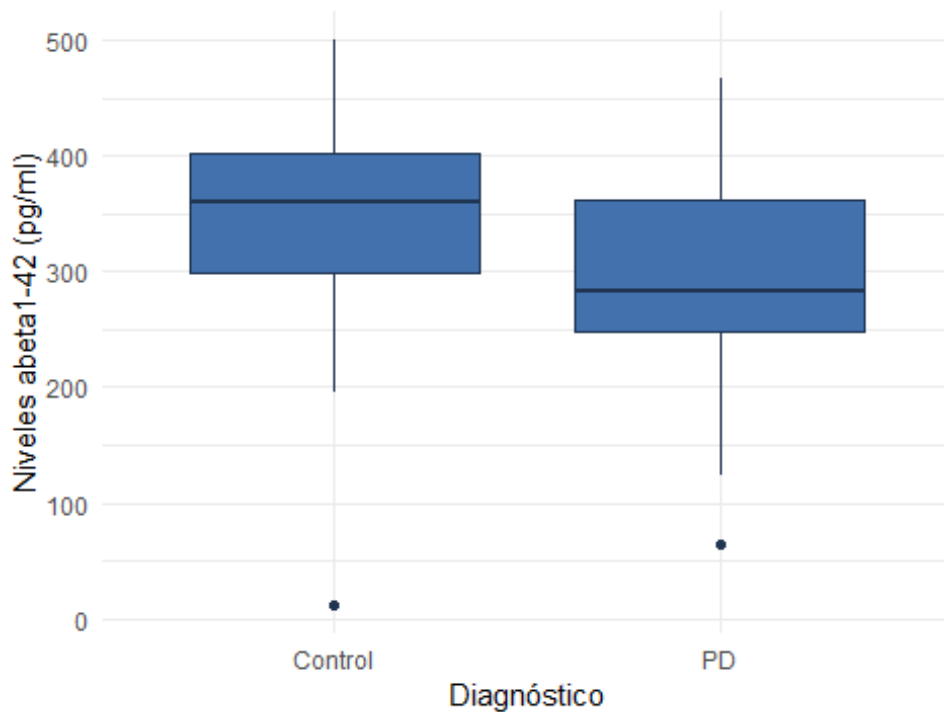


Figura 16. Niveles proteína  $\beta$ 1-42 amiloide en líquido cefalorraquídeo en hombres. .

En mujeres, se observan niveles algo más elevados de Abeta en el grupo control que en el grupo PD (Figura 17). Se obtiene un p-valor de 0.18, por lo tanto, no se detectan diferencias significativas en los niveles de Abeta entre el grupo PD y el grupo control.

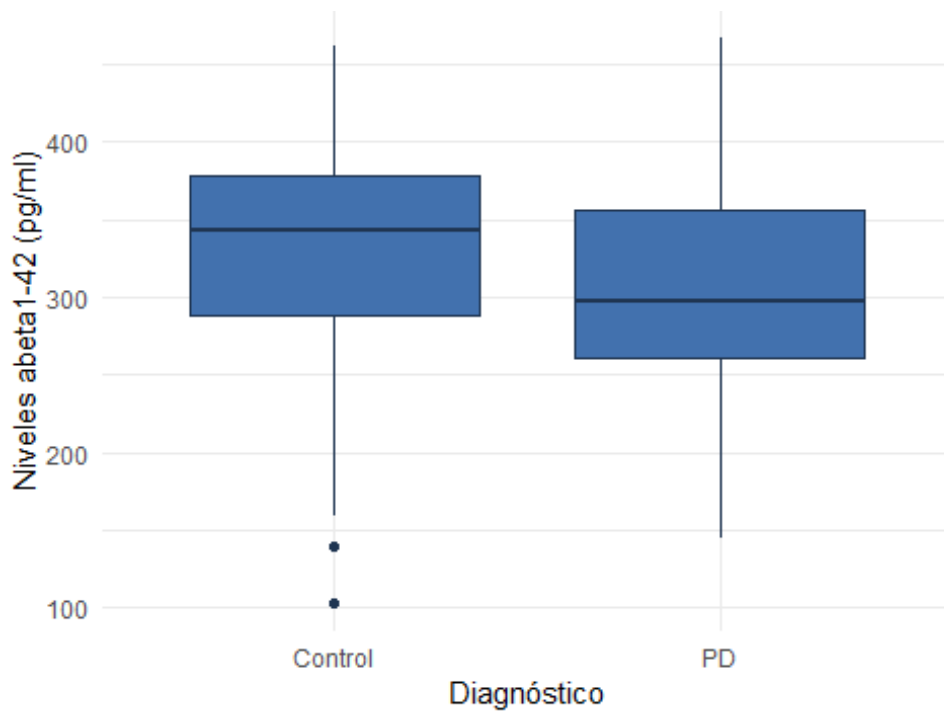


Figura 17. Niveles proteína  $\beta$ 1-42 amiloide en líquido cefalorraquídeo en mujeres.

### 2.2.3. Tau total

Se observan niveles de tau bastante iguales en los dos de grupos diagnóstico: control y PD (Figura 18). Se obtiene un p-valor de 0.06, por lo tanto, no se detectan diferencias significativas en los niveles de tau entre el grupo PD y el grupo control.

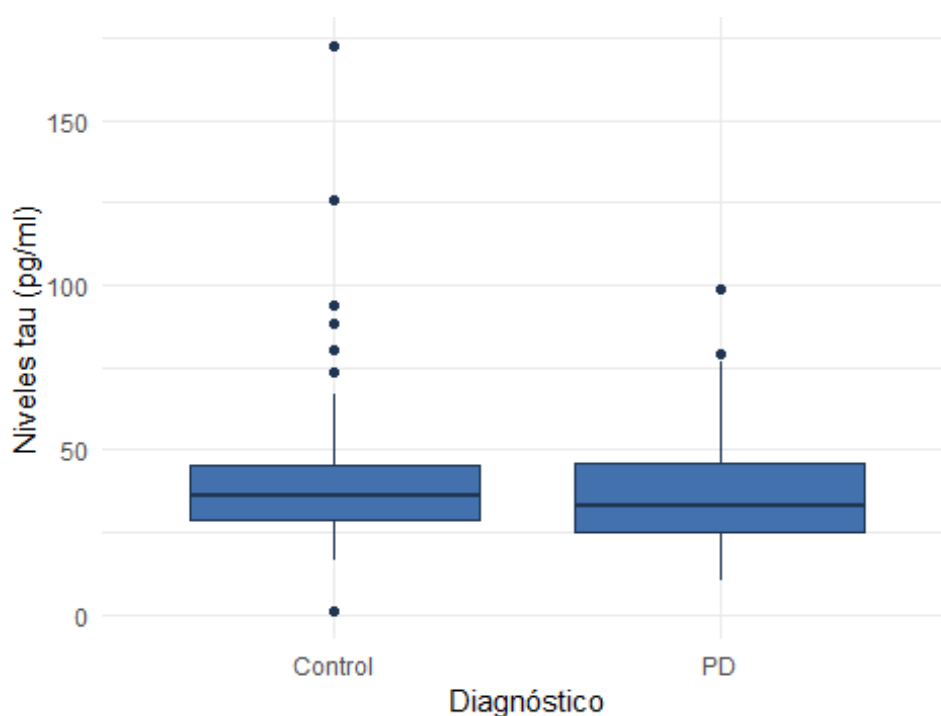


Figura 18. Niveles proteína tau en líquido cefalorraquídeo.

En hombres, también se observan niveles parecidos de proteína tau en los dos grupos (Figura 19). Se obtiene un p-valor de 0.28, por lo tanto, no se detectan diferencias significativas en los niveles de tau entre el grupo PD y el grupo control.

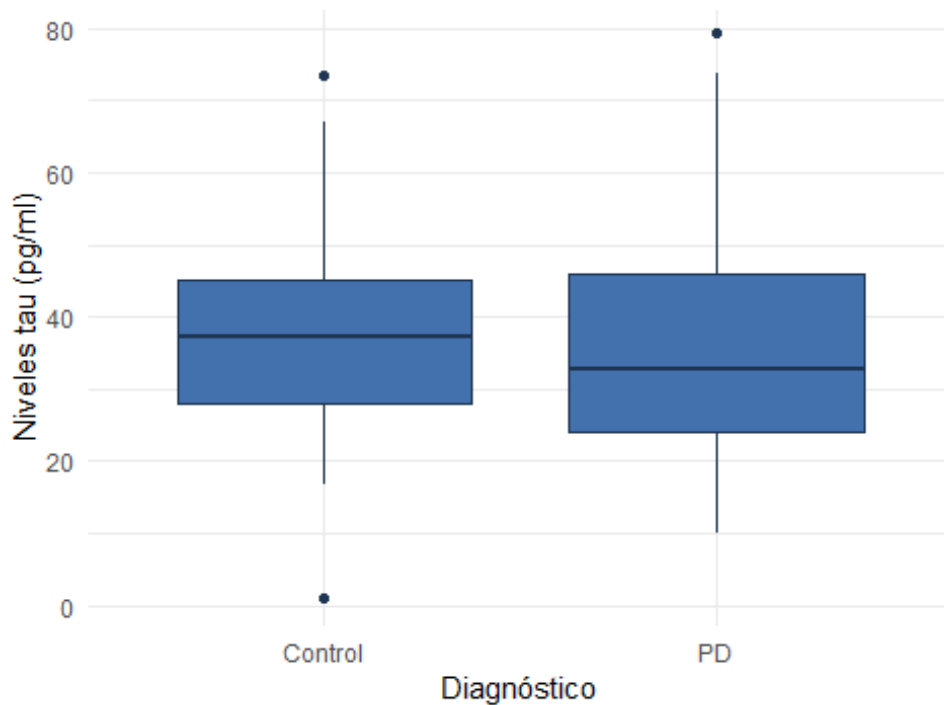


Figura 19. Niveles proteína tau en líquido cefalorraquídeo en hombres.

En mujeres, se observan niveles de tau algo superiores en el grupo control respecto al PD (Figura 20). Se obtiene un p-valor de 0.17, por lo tanto, no se detectan diferencias significativas en los niveles de tau entre el grupo PD y el grupo control.

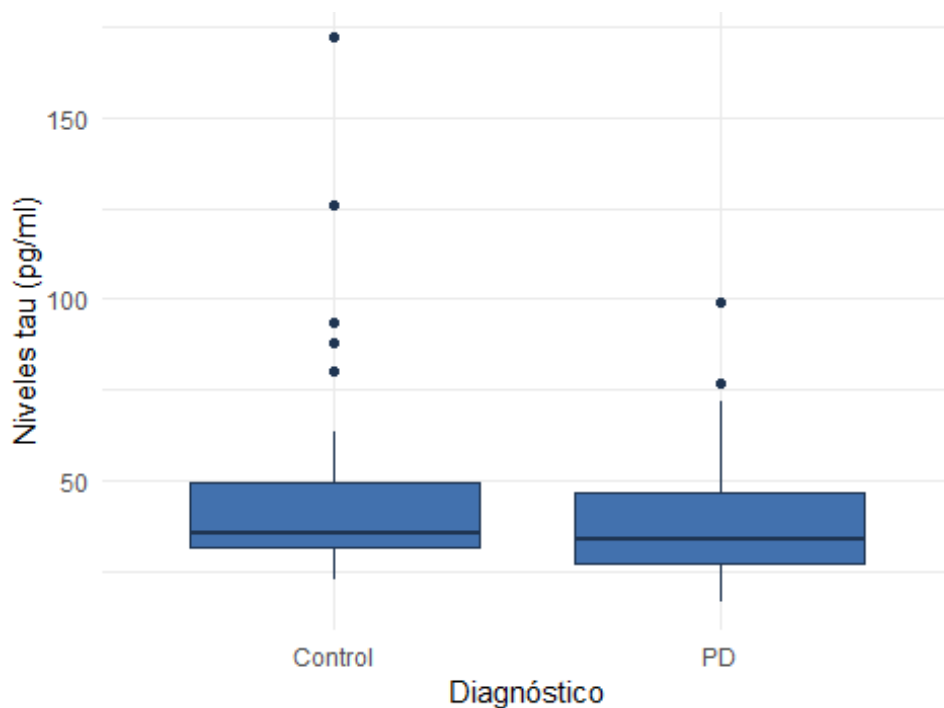


Figura 20. Niveles proteína tau en líquido cefalorraquídeo en mujeres.

#### 2.2.4. Tau fosforilada

Se observan niveles de ptau bastante iguales en los dos grupos diagnóstico: control y PD (Figura 21). Se obtiene un p-valor de 0.03, por lo que se detectan diferencias significativas en los niveles de ptau entre el grupo PD y el grupo control.

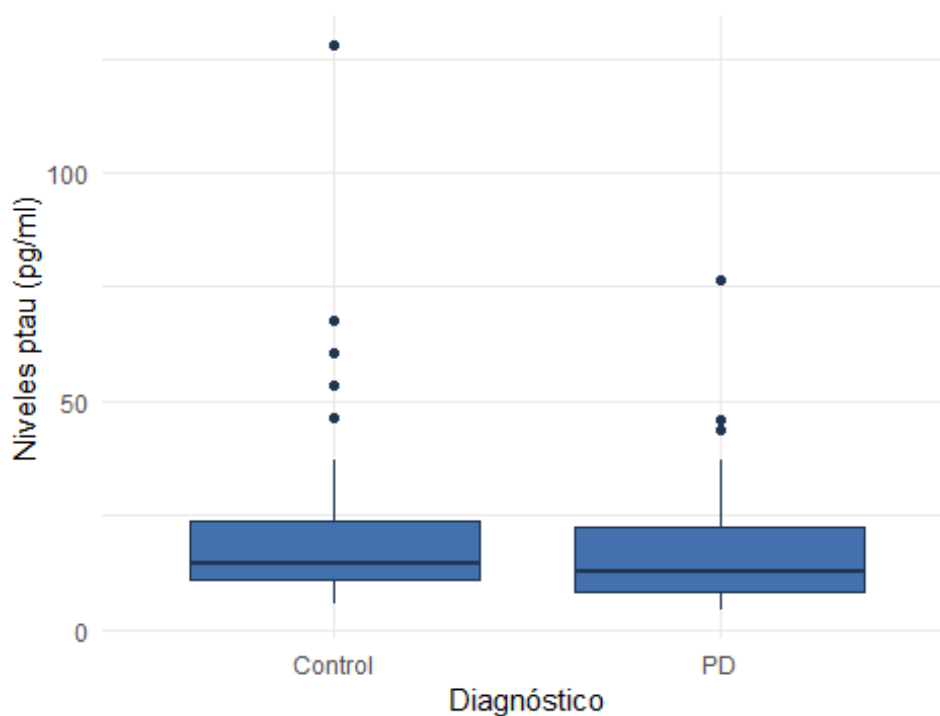


Figura 21. Niveles proteína ptau en líquido cefalorraquídeo.

En hombres, se observan niveles de ptau un poco más elevados elevados en el grupo control respecto al grupo PD (Figura 22). Se obtiene un p-valor de 0.30, por lo tanto, no se detectan diferencias significativas en los niveles de tau entre el grupo PD y el grupo control.

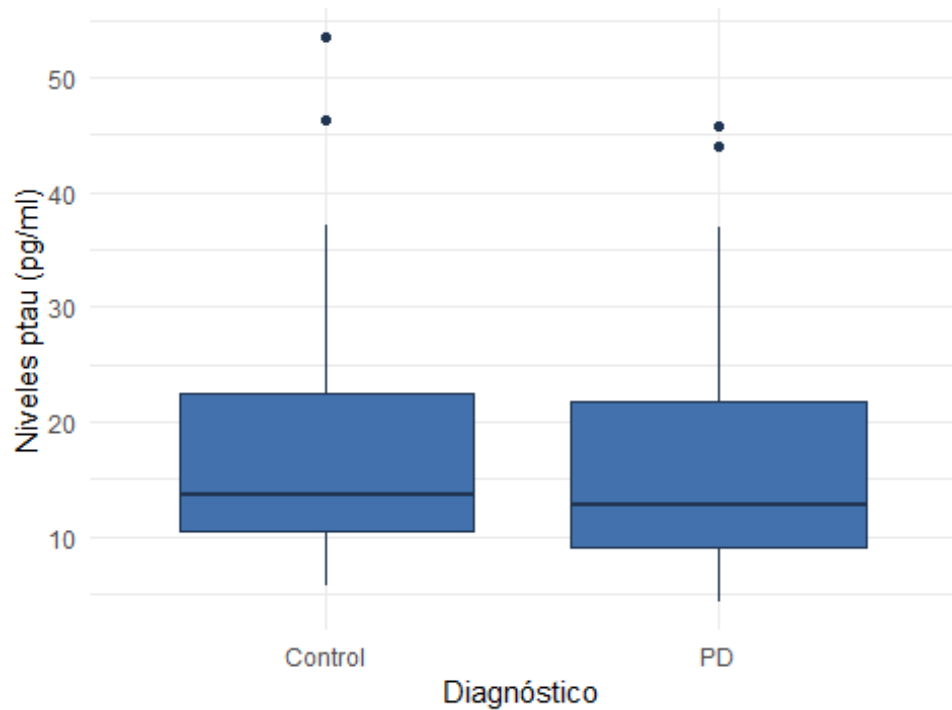


Figura 22. Niveles proteína ptau en líquido cefalorraquídeo en hombres.

En mujeres, se observan niveles de ptau parecidos en los dos grupos (Figura 23). Se obtiene un p-valor de 0.07, por lo tanto, no se detectan diferencias significativas en los niveles de tau entre el grupo PD y el grupo control.

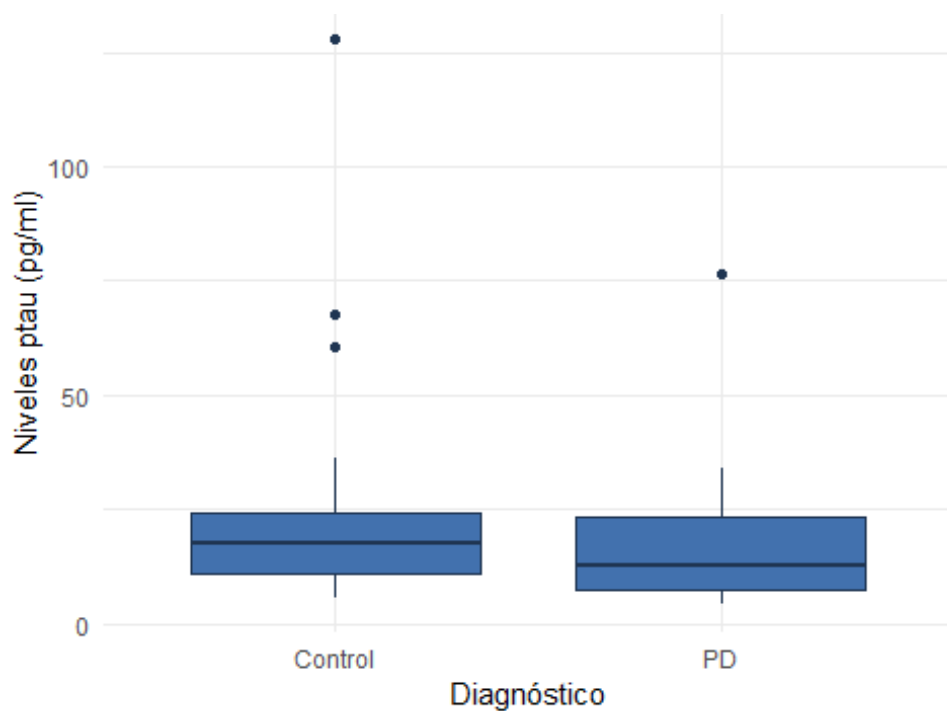


Figura 23. Niveles proteína ptau en líquido cefalorraquídeo en mujeres.

### 2.2.5. Conclusiones

Claramente se ha observado que los niveles de Abeta son superiores a las dos formas de proteína tau, pero este perfil proteico no coincide con el perfil de paciente de AD comentado en la introducción donde los niveles de Abeta son más bajos que los de tau y ptau (Shi et al, 2011).

A pesar de esto, en Abeta, en los tres contrastes realizados se han observado niveles del grupo control superiores respecto al PD con diferencias significativas coincidiendo con estudios comentados previamente (Kang et al, 2013). Respecto al género se han obtenido diferencias significativas entre hombres pero no entre mujeres.

Por lo que respecta a tau total, los niveles en ambos grupos se han hallado bastante iguales con lo que no se han detectado diferencias significativas entre el grupo PD *versus* el grupo control, ni en el conjunto de datos ni separado por género.

En referencia a tau fosforilada, en los niveles observados también se han obtenido niveles parecidos como en tau, pero en este caso, se han detectado diferencias significativas en el conjunto de datos total. Por los grupos de género, hombre y mujer, no se han hallado diferencias significativas.

Estos resultados obtenidos de las dos formas de proteínas tau difieren de lo detectado con Abeta pero tal y como se ha mencionado previamente hay estudios que se obtienen niveles de PD más bajos y otros que no (Kang et al, 2013).

De todos estos datos, sería interesante conocer los niveles de  $\alpha$ -syn para poder deducir sobre el hecho de que Abeta puede promover la degradación de  $\alpha$ -syn y esta el aumento de fosforilación de tau (Kim et al, 2014). Además de conocer el grado de degradación cognitiva del paciente o si poseen o no AD.

## 2.3. Proteínas relevantes

### 2.3.1. Introducción

Este apartado corresponde al tercer objetivo del trabajo cuya descripción es intentar detectar las proteínas más expresadas en la enfermedad de Párkinson entre los dos géneros en plasma.

Los datos de este análisis intentan responder al objetivo 3 y provienen del proyecto 121 del estudio BioFIND. En este proyecto se realizó un screening de proteínas en una población de 10 hombres y 10 mujeres con PD. En el resultado obtenido se detectan 1305 proteínas en cada individuo. Debido a este número tan elevado de proteínas para este trabajo se han seleccionado 9 considerando la bibliografía científica. Estas proteínas son: factor de crecimiento epidérmico (*epidermal growth*

*factor*, EGF), apolipoproteína A1 (*apolipoprotein A1*, ApoA1), proteína desglucasa DJ-1 (*protein deglycase DJ-1*, DJ-1), factor neurotrófico derivado del cerebro (*brain-derived neurotrophic factor*, BDNF), proteína acida fibrilar glial (*glial fibrillary acidic protein*, GFAP), fractalquina (*fractalkine*, Frac), nicotinamida fosforribosiltransferasa (*nicotinamide phosphoribosyltransferase*, NAMPT), quinasa putativa 1 inducida por PTEN (*PTEN induced kinase 1*, PINK1) y leucina ricos repetir las quinasas 2 (*leucine-rich repeat kinase 2*, LRRK). A continuación, se describe la información encontrada en la bibliografía referente a estas proteínas:

- EGF. Puede considerarse un predictor potencial del deterioro cognitivo ya que se ha relacionado con el rendimiento cognitivo y se ha encontrado en niveles bajos en pacientes con PD (Chen et al, 2011 y Pellecchia et al, 2013).
- ApoA1. Niveles altos de la apolipoproteína A1 en plasma se han relacionado con una mejor integridad del sistema dopaminérgico en pacientes con PD, sugiriendo un efecto de protección (Qiang et al, 2013).
- DJ-1. Se ha encontrado que esta proteína forma parte de la PD genética autosómica recesiva por la mutación del gen PARK7 y del PD idiopático. Hay más estudios realizados en CSF que en plasma, pero se han encontrado resultados poco consistentes ya que en función del tipo de técnica utilizada los niveles de DJ-1 en PD son más altos o más bajos comparado con el control (Kim et al, 2014). A pesar de esto, esta proteína está muy presente en los eritrocitos de la sangre así que es importante considerar la contaminación de sangre en la muestra (Lontankar et al, 2017).
- BDNF. Este es un biomarcador relacionado con la memoria y la función cognitiva en el envejecimiento. Se han encontrado niveles alterados en suero y plasma de este factor en pacientes con PD (Sharma et al, 2013). Además, en CSF, BDNF se ha relacionado con el desarrollo cognitivo (Lontankar et al, 2017).
- GFAP. Se han encontrado niveles elevados de esta proteína en pacientes con PD en CSF, suero y plasma (Lontankar et al, 2017).
- Frac. Esta citoquina se ha detectado en altos niveles en PD ya que está relacionado con un proceso de inflamación. Estos niveles elevados también están relacionados con un aumento de la gravedad de PD. También, se utiliza para considerar la correlación Abeta/Frac para estudiar el progreso de la enfermedad (Shi et al, 2011).
- NAMPT. Se ha encontrado que su gen está desregulado ya que se observan niveles elevados de esta proteína en

sangre en pacientes con PD comparado con pacientes sanos (Santiago et al, 2016).

- PINK1. Es un a quinasa cuyo gen es responsable de la forma autosómica recesiva en PD que a veces es confundida por un origen idiopático debido a los síntomas (Rango et al, 2013).
- LRRK. Es una quinasa cuyo gen es responsable de la forma autosómica dominante de PD. Se ha observado que la mutación de este gen genera la fosforilación de  $\alpha$ -syn y tau (Kim et al, 2014).

### 2.3.2. Epidermal growth factor (EGF)

En hombres el rango de valores obtenidos es más amplio y más elevados que en mujeres (Figura 24). El p-valor obtenido es de 0.18 indicando que no se encuentran diferencias significativas entre los niveles de EGF entre hombres y mujeres.

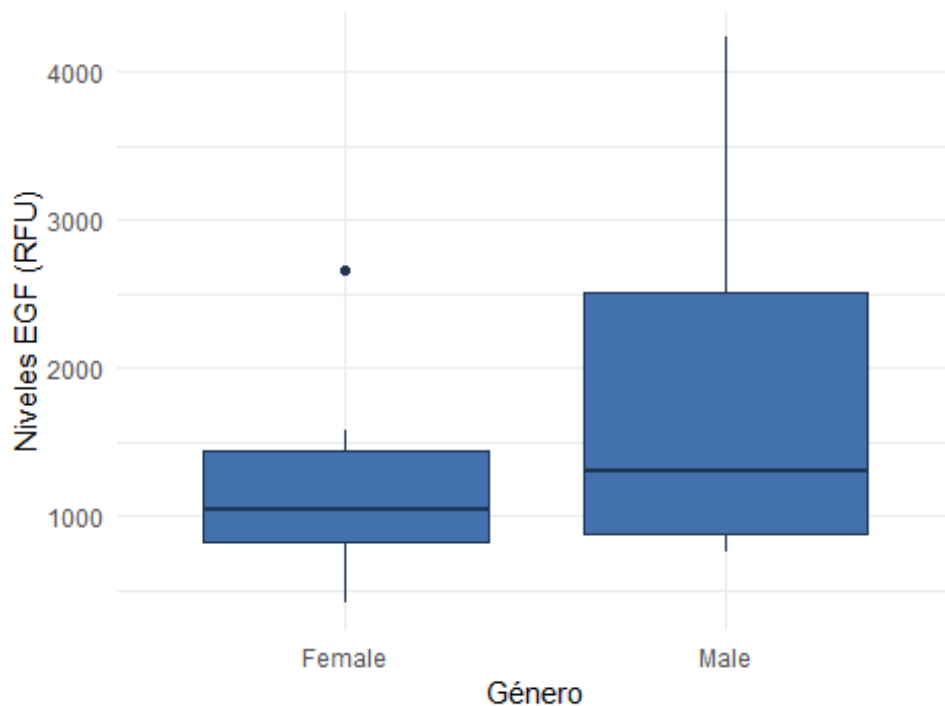


Figura 24. Niveles EGF entre géneros en plasma.



### 2.3.3. Apolipoprotein AI (ApoA1)

Se han encontrado niveles parecidos de ApoA1 entre los dos grupos (Figura 25). El p-valor que se ha obtenido es de 0.73. Se determina que no se encuentran diferencias significativas entre los niveles de ApoA1 entre hombres y mujeres.

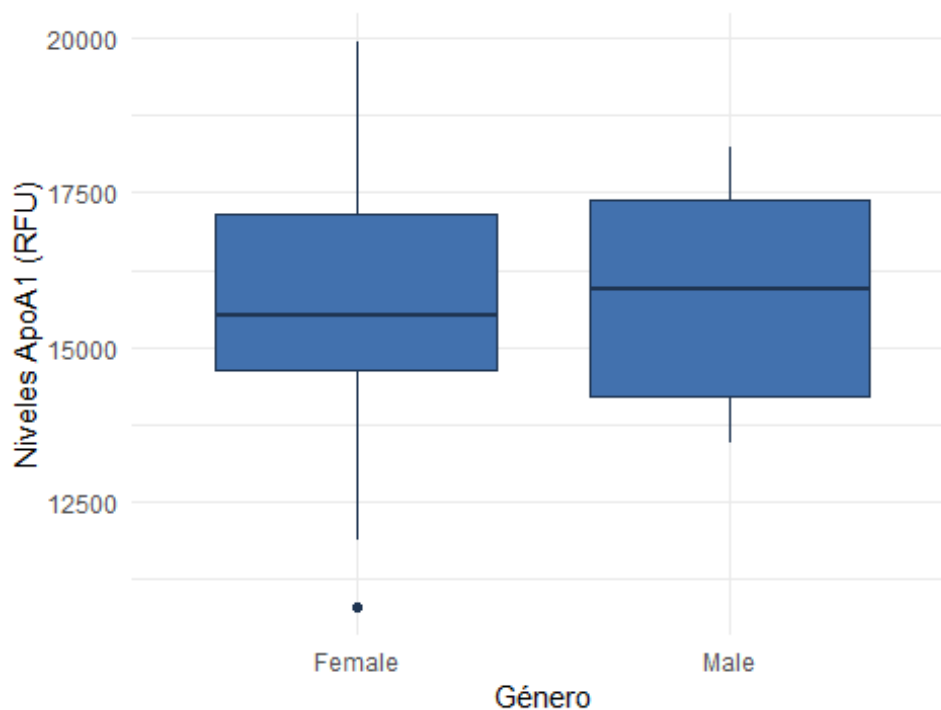


Figura 25. Niveles ApoA1 entre géneros en plasma.

### 2.3.4. Protein deglycase DJ-1 (DJ-1)

Se han encontrado niveles parecidos de DJ-1 entre los dos grupos y hay más rango de valores en el grupo hombre que en el grupo mujer (Figura 26). El p-valor que se ha obtenido es de 0.24. Se determina que no se detectan diferencias significativas entre los niveles de DJ-1 entre hombres y mujeres.

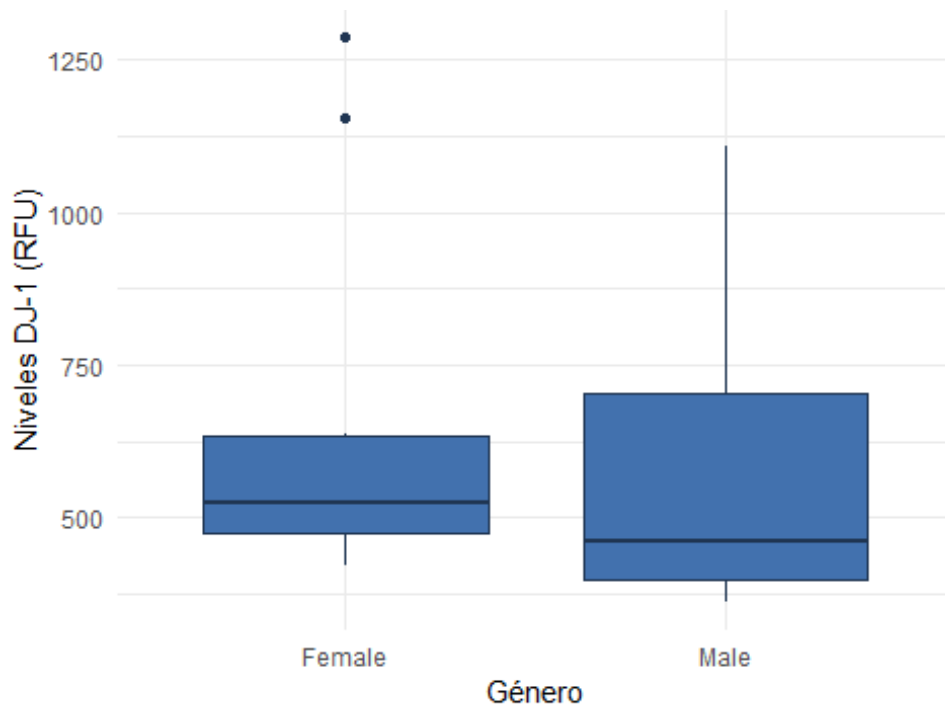


Figura 26. Niveles DJ-1 entre géneros en plasma.

### 2.3.5. Brain-derived neurotrophic factor (BDNF)

Se han encontrado niveles de BDNF en el grupo hombre más elevados que en el grupo mujer (Figura 27). El p-valor que se ha obtenido es cercano a 0.05 con lo que para no arriesgar se acepta la hipótesis nula y se concluye que no se detectan diferencias significativas entre los niveles de BDNF entre hombres y mujeres.

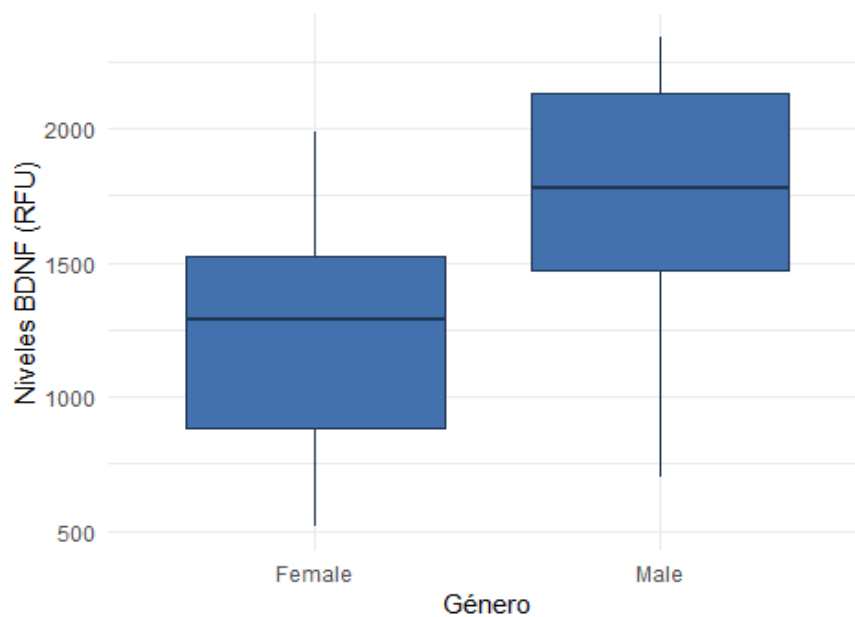


Figura 27. Niveles BDNF entre géneros en plasma.

### 2.3.6. Glial fibrillary acidic protein (GFAP)

En ambos grupos se observan niveles parecidos y el rango de valores obtenido es amplio (Figura 28). El p-valor resultante es de 0.52, con lo que no hay diferencias significativas entre los niveles de GFAP entre hombres y mujeres.

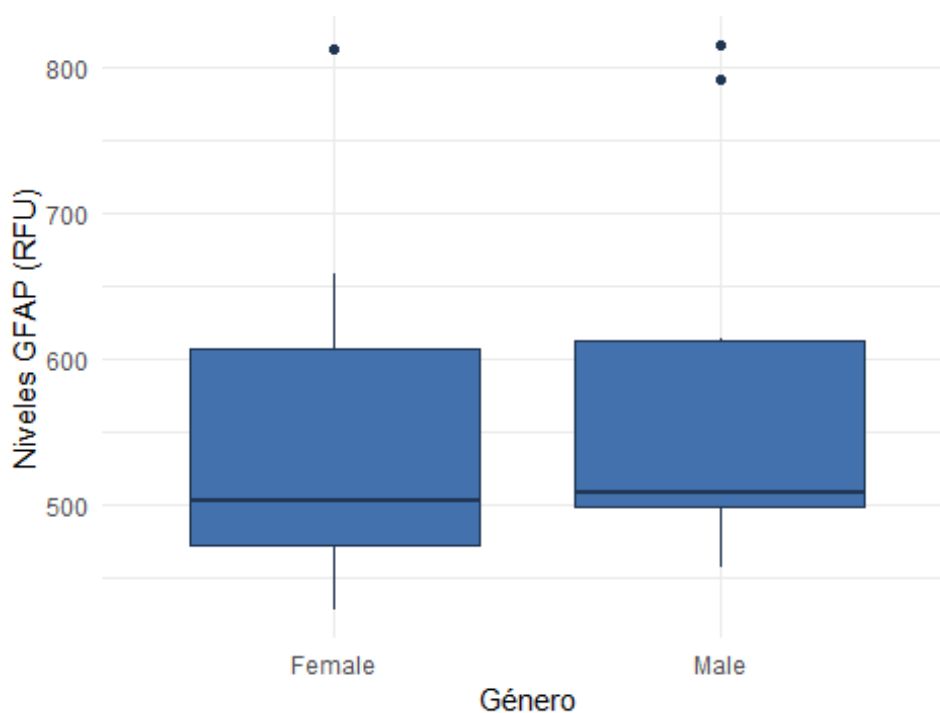


Figura 28. Niveles GFAP entre géneros en plasma.

### 2.3.7. Fractalkine (Frac)

En este caso, se puede observar niveles más elevados de Frac en mujeres que en hombres (Figura 29). El p-valor resultante es de 0.35. Entonces, no se encuentran diferencias significativas entre los niveles de *Fractalkine* entre hombres y mujeres.

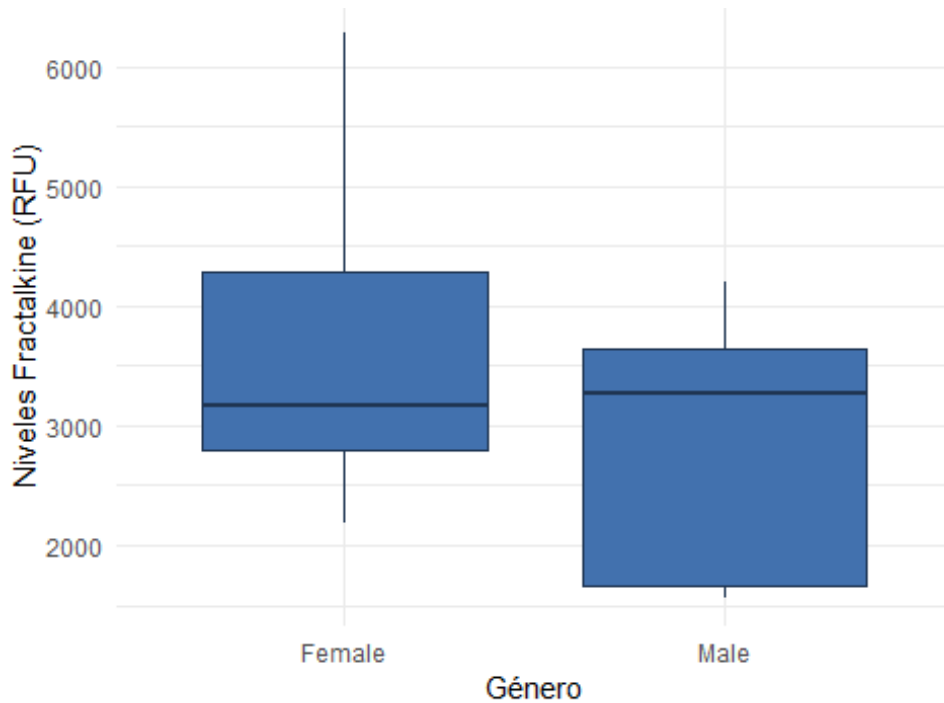


Figura 29. Niveles Frac entre géneros en plasma.

### 2.3.8. Nicotinamide phosphoribosyltransferase (NAMPT)

Los niveles de NAMPT en ambos grupos son parecidos, pero en hombres el rango de valores obtenidos es más amplio (Figura 30). El p-valor es de 0.60. Así, se determina que no hay diferencias significativas entre los niveles de NAMPT entre hombres y mujeres.

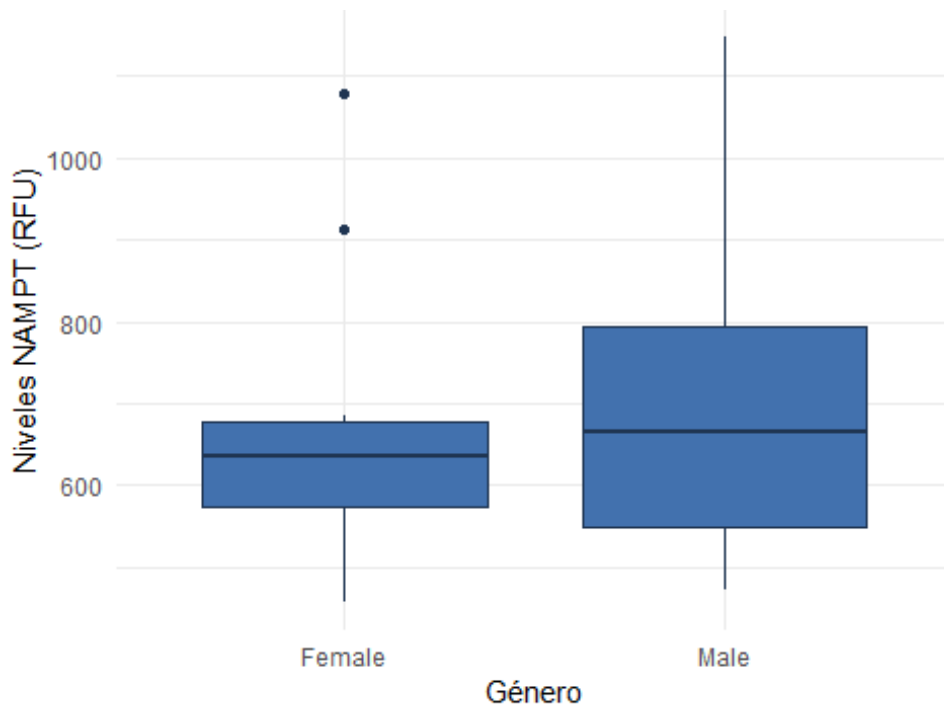


Figura 30. Niveles NAMPT entre géneros en plasma.

### 2.3.9. PTEN induced kinase 1 (PINK1)

Los niveles de PINK1 en mujeres se han encontrado más elevados que en hombres (Figura 31). El p-valor obtenido es de 0.03. Por esto, se determina que hay diferencias significativas entre los niveles de PINK1 entre hombres y mujeres.

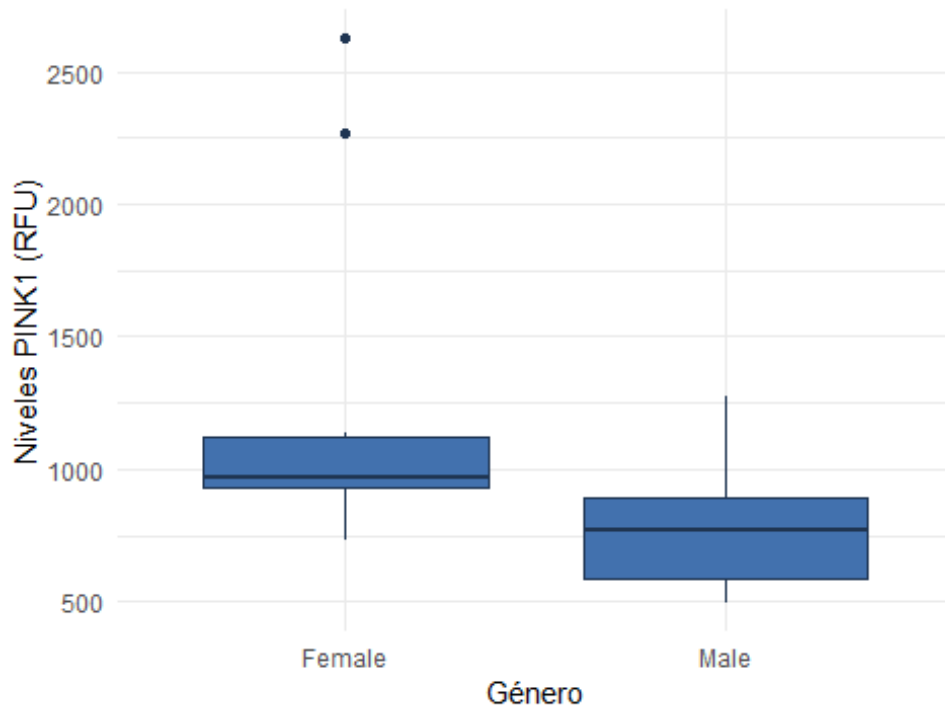


Figura 31. Niveles PINK1 entre géneros en plasma.

### 2.3.10. Leucine-rich repeat kinase 2 (LRRK)

Los valores obtenidos de LRRK en hombres son más variables y alcanzan niveles más elevados en hombres que en mujeres (Figura 32). El p-valor obtenido es muy cercano al nivel de significación, 0.05, por esto, se acepta la hipótesis nula y se determina que no hay diferencias significativas entre los niveles de LRRK entre hombres y mujeres.

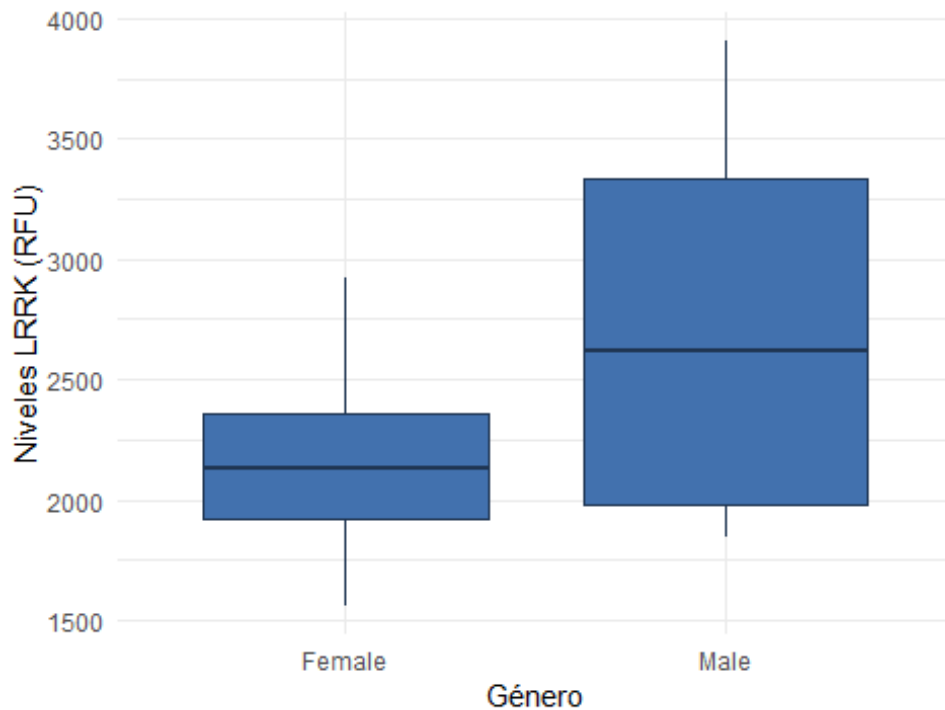


Figura 32. Niveles LRRK entre géneros en plasma.

### 2.3.11. Conclusiones

De todos los resultados obtenidos, solamente se detectan diferencias significativas PINK1. Es complicado determinar el motivo ya que la población de estudio es pequeña pero es interesante considerar que esta quinasa es responsable de la causa de PD autosómica y recesiva (Rango et al, 2013).

En este estudio, hubiera sido interesante tener un grupo control con el fin de poder comparar con la población con PD. Asimismo, información respecto al estado de salud del paciente también se podría haber considerado relevante con el fin de poder contextualizar más los resultados finales.

### 3. Conclusiones

A nivel de contexto, se puede concluir que la búsqueda de biomarcadores engloba diferentes ámbitos y enfoques. Es importante considerar tanto a nivel global como a nivel individual. El primero correspondería a la interacción y relación entre las proteínas y el estado de salud del paciente y el segundo en la función de la proteína de interés ya que, si solamente se considerara esto, la función como biomarcador estaría limitada. En definitiva, una combinación de diferentes biomarcadores potenciales puede ser más útil y permitir realizar un diagnóstico más personalizado.

Asimismo, sería interesante ver en estudios longitudinales la evolución de un conjunto de diferentes biomarcadores potenciales con el fin de ver la evolución de la enfermedad. Es verdad que estos indicadores variarían en función de la genética y el estilo de vida del individuo.

Respecto al trabajo, se considera que hubiera sido interesante utilizar solamente todos los datos de una fuente para poder así realizar comparaciones entre más proteínas. Además, el uso de variables relacionadas con el estado de salud del paciente como la edad, el grado de la enfermedad, entre otros, hubieran permitido poder realizar modelos de predicciones y obtener un resultado final con conclusiones más sólidas.

En referencia a los objetivos del trabajo y su temporización, se puede concluir que el objetivo general 1 y 2 se han podido lograr. Por lo que hace el objetivo 3, al contener los datos tantas proteínas detectadas la selección se ha visto limitada. Además, si los datos hubieran tenido un grupo control, los resultados hubieran sido más firmes. Finalmente, se considera que se ha podido seguir bastante la metodología y planificación realizada en el inicio del proyecto sin muchos obstáculos.

Siguiendo la línea de este trabajo sería interesante estudiar una relación entre la proteína  $\alpha$ -syn y las tres proteínas relacionadas con AD: abeta, tau y ptau. Además, sería atrayente analizar todas las proteínas del tercer apartado, conocer su función entre ellas y poder detectar si hay o no una relación entre ellas.

## 4. Glosario

$\alpha$ -syn: alfa-sinucleína

Abeta:  $\beta$ 1-42 amiloide

AD: Alzheimer's disease (enfermedad de Alzheimer)

ApoA1: apolipoprotein AI (apolipoproteína A1)

BDNF: brain-derived neurotrophic factor (factor neurotrófico derivado del cerebro)

BioFIND: Fox Investigation for New Discovery of Biomarkers in Parkinson's disease

CSF: cerebrospinal fluid (líquido cefalorraquídeo)

DJ-1: protein deglycase DJ-1 (proteína desglucasa DJ-1-)

EGF: epidermal growth factor (factor de crecimiento epidérmico)

ELISA: Enzyme-Linked ImmunoSorbent Assay (ensayo por inmunoabsorción ligado a enzimas)

Frac: fractalkine (fractalquina)

GFAP: glial fibrillary acidic protein (proteína ácida fibrilar glial)

LB: Lewis bodies (cuerpos de Lewis)

LRRK: leucine-rich repeat kinase 2 (leucina ricos repetir las quinasas 2)

MRI: magnetic resonance imaging (imagen por resonancia magnética)

NAMPT: nicotinamide phosphoribosyltransferase (nicotinamida fosforribosiltransferasa)

PD: Parkinson's disease (enfermedad de Parkinson)

PET: positron emission tomography (tomografía de emisión de positrones)

PINK1: PTEN induced kinase 1 (quinasa putativa 1 inducida por PTEN)

Ptau: tau fosforilada

SPECT: single photon emission computed tomography (tomografía computarizada de emisión monofotónica )

Tau: tau total



## 5. Bibliografía

### Artículos:

CHEN, Alice S. et al. (2014). <<Plasma EGF levels predict cognitive decline in Parkinson's disease>>. Annals of neurology (núm. 69(4), pág. 655-663).

CHEN, Alice S. (2014). <<Unbiased approaches to biomarker discovery in neurodegenerative diseases>>. Neuron (núm. 84(3), pág. 594-607).

DELENCÓS, Marion et al. (2016). <<Biomarkers in Parkinson's disease: advances and strategies>>. Hepatology (núm. 63, pág. 965-982).

HONG, Zhen et al. (2010). <<DJ-1 and  $\alpha$ -synuclein in human cerebrospinal fluid as biomarkers for Parkinson's disease>>. Brain (núm. 133, pág. 713-726).

KANG, Ju-Hee et al. (2013). <<Association of cerebrospinal fluid  $\beta$ -amyloid 1-42, T-tau, P-tau181 and  $\alpha$ -synuclein levels with clinical features of drug-naive patients with early Parkinson disease>>. JAMA Neurology (núm. 40, pág. 1277-1287).

KANG, Un Jung et al. (2016). << The BioFIND study: Characteristics of a clinically typical Parkinson's disease biomarker cohort>>. Movement Disorders (núm. 31, pág. 926-932).

KANG, Wenyan et al. (2016). << Salivary total  $\alpha$ -synuclein, oligomeric  $\alpha$ -synuclein and SNCA variants in Parkinson's disease patients >>. Scientific Reports (núm. 6, pág. 28143).

KIM, Dana et al. (2014). << What is the clinical significance of cerebrospinal fluid biomarkers in Parkinson's disease? Is the significance a diagnostic or prognostic? >>Experimental Neurobiology (núm. 23(4), pág. 352-364).

LE, Weidong et al. (2017). <<Can biomarkers help the early diagnosis of Parkinson's disease? >>. Neuroscience Bulletin (núm. 33, pág. 535-542).

LOTANKAR, Sharvari et al. (2017). <<Biomarkers for Parkinson's Disease:recent advancement>>. Neuroscience Bulletin (núm. 33, pág. 585-597).

MALEK, N. et al. (2014). <<Alpha-synuclein in peripheral tissues and body fluids as a biomarker for Parkinson's disease –a systematic review>>. Acta Neurologica Scandinavica (núm. 130, pág. 59-72).

MILLER, Diane B. and O'Callaghan, James P. (2015). <<Biomarkers of Parkinson's disease: Present and future>>. Metabolism (núm. 64, pág. S40-S46).

PELLECCHIA, Maria T. et al. (2013). <<Serum epidermal growth factor predicts cognitive functions in early, drug-naive Parkinson's disease patients>>. Journal of Neurology (núm. 260, pág. 438-444).

POEWE, Werner et al. (2017). <<Parkinson disease>>. Handbook of Clinical Neurology (núm. 159, pág. 173-193).

QIANG JK et al. (2013). <<Plasma apolipoprotein A1 as a biomarker for Parkinson disease>>. Annals of neurology (núm. 74, pág. 119– 27).

RANGO, M. et al. (2013). <<PINK1 parkinsonism and Parkinson disease: Distinguishable brain mitochondrial function and metabolomics>>. Mitochondrion (núm. 13, pág. 59-61).

SANTIAGO, José Antonio et al. (2016). <<Integrative transcriptomic meta-analysis of Parkinson's disease and depression identifies NAMPT as a potential blood biomarker for de novo Parkinson's disease>>. Scientific Reports (núm. 6, pág. 34579).

SCHMID, Sabrina et al. (2018). << Alpha-synuclein is present in dental calculus but not altered in Parkinson's disease patients in comparison to controls>>. Journal of Neurology (núm. 265(6), pág. 1334-1337).

SHARMA, Sushil et al. (2013). <<Biomarkers in Parkinson's disease (recent update)>>. Neurochemistry International (núm. 63, pág. 201-229).

SHI, Min et al. (2011). <<Cerebrospinal fluid biomarkers for Parkinson disease diagnosis and progression>>. Annals of neurology (núm. 69(3), pág. 570-580).

TROPEA, Thomas F. and Chen, Alice (2018). <<Unlocking the mystery of biomarkers: a brief introduction, challenges and opportunities in Parkinson disease>>. Parkinsonism Related Disorders (núm. 46 (Suppl 1), pág. S15-S18).

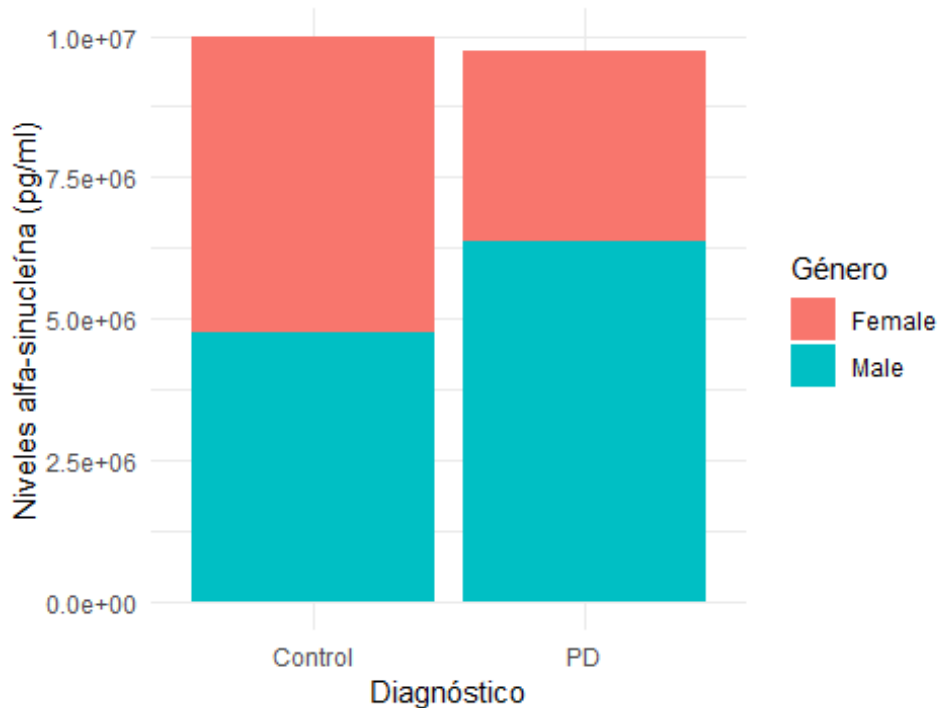
Página web:

The Fox Investigation for New Discovery of Biomarkers (Fecha de consulta: 28/02/19) <<https://www.michaeljfox.org/page.html?biofind-clinical-study>>

## 6. Anexos

En estos anexos se presentan algunos de los gráficos y resultados obtenidos de las proteínas estudiadas con el fin de complementar lo descrito anteriormente.

### 6.1 Alfa-sinucleína en plasma



Niveles alfa-sinucleína en plasma entre diagnóstico y género.

Grupo control:

```
length(alfaplasma.ct$`Alpha-syn`)  
## [1] 92  
summary(alfaplasma.ct$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
## 22836  60539   93805  108502 132885  290308
```

Grupo PD:

```
length(alfaplasma.pd$`Alpha-syn`)  
## [1] 106  
summary(alfaplasma.pd$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
## 12799  42766   77279   91856 110857  265932
```

En hombres grupo control:

```
length(alfaplasma.ct.men$`Alpha-syn`)  
## [1] 45  
summary(alfaplasma.ct.men$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  22836  56704   83673  105620 133215  290308
```

En hombres grupo PD:

```
length(alfaplasma.pd.men$`Alpha-syn`)  
## [1] 64  
summary(alfaplasma.pd.men$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  12799  43491   87111   99801 116911  265932
```

En mujeres grupo control:

```
length(alfaplasma.ct.women$`Alpha-syn`)  
## [1] 47  
summary(alfaplasma.ct.women$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  34303  68264  101827  111261 130306  282970
```

En mujeres grupo PD:

```
length(alfaplasma.pd.women$`Alpha-syn`)  
## [1] 42  
summary(alfaplasma.pd.women$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  16198  41144   66693   79749  98811  257016
```

Comparación niveles de alfa-syn entre PD y control en plasma:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfaplasma.ct$`Alpha-syn`)  
##  
## Shapiro-Wilk normality test  
##  
## data:  alfaplasma.ct$`Alpha-syn`  
## W = 0.87858, p-value = 3.969e-07
```

```
shapiro.test(alfaplasma.pd$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfaplasma.pd$`Alpha-syn`  
## W = 0.84034, p-value = 2.494e-09
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfa.plasma$`Alpha-syn`, group = alfa.plasma$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##           Df F value Pr(>F)  
## group    1  0.0221 0.8821  
##          196
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfa.plasma)
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data:  Alpha-syn by DIAGNOSIS  
## W = 5909, p-value = 0.01024  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación niveles de alfa-syn entre PD y control en hombres en plasma:

Ho: datos siguen normalidad

Ha: datos no siguen normalidad

```
shapiro.test(alfaplasma.ct.men$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfaplasma.ct.men$`Alpha-syn`  
## W = 0.87082, p-value = 0.0001317
```

```
shapiro.test(alfaplasma.pd.men$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfaplasma.pd.men$`Alpha-syn`  
## W = 0.85518, p-value = 2.368e-06
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfaplasma.men$`Alpha-syn`, group = alfaplasma.men$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##           Df F value Pr(>F)  
## group    1  0.0248 0.8753  
##          107
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfaplasma.men)
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data: Alpha-syn by DIAGNOSIS  
## W = 1566, p-value = 0.4399  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación niveles de alfa-syn entre PD y control en mujeres en plasma:

Ho: datos siguen normalidad

Ha: datos no siguen normalidad

```
shapiro.test(alfaplasma.ct.women$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: alfaplasma.ct.women$`Alpha-syn`  
## W = 0.87076, p-value = 9.517e-05
```

```
shapiro.test(alfaplasma.pd.women$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: alfaplasma.pd.women$`Alpha-syn`  
## W = 0.82178, p-value = 1.335e-05
```

Ho: datos tienen igualdad de varianza Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfaplasma.women$`Alpha-syn`, group = alfaplasma.women$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  0.4847 0.4882  
##      87
```

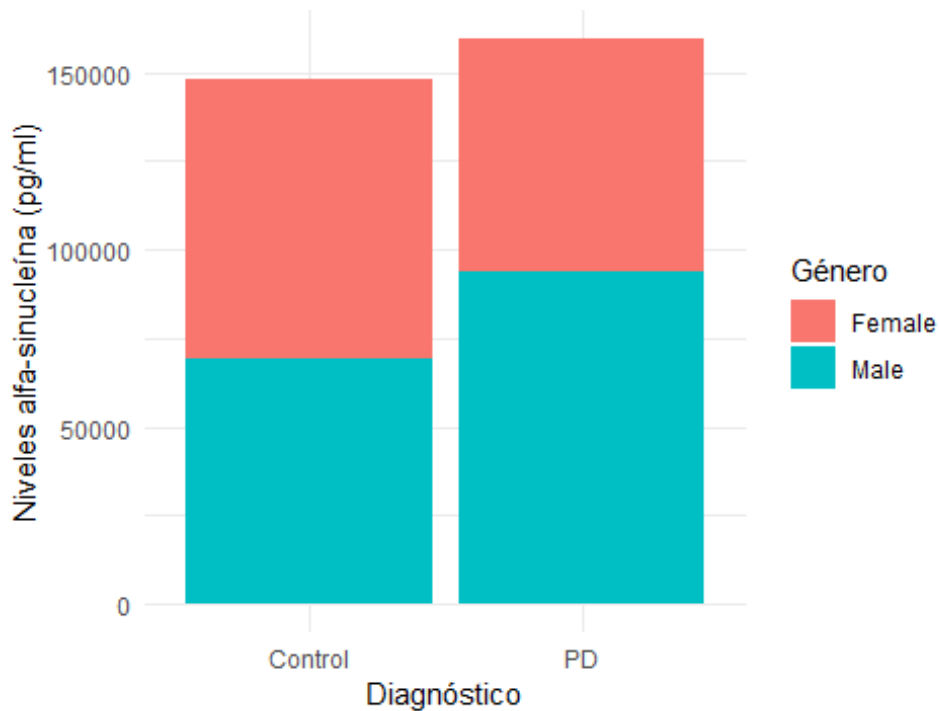
Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfaplasma.women)
```

```
##  
## Wilcoxon rank sum test  
##  
## data: Alpha-syn by DIAGNOSIS  
## W = 1357, p-value = 0.002134  
## alternative hypothesis: true location shift is not equal to
```

## 6.2 Alfa-sinucleína en líquido cefaloraquídeo



Niveles alfa-sinucleína en líquido cefaloraquídeo entre diagnóstico y género.

Grupo control:

```
length(alfacsf.ct$`Alpha-syn`)  
## [1] 84  
summary(alfacsf.ct$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  855.2 1320.0 1617.7 1764.4 2067.3 3852.8
```

Grupo PD:

```
length(alfacsf.pd$`Alpha-syn`)  
## [1] 109  
summary(alfacsf.pd$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  355.8  982.4 1386.9 1463.5 1891.2 3449.3
```

Grupo control hombres:

```
length(alfacsf.ct.men$`Alpha-syn`)  
## [1] 39  
summary(alfacsf.ct.men$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  946.1 1356.0 1661.1 1769.8 2086.0 3352.4
```

Grupo PD hombre:

```
length(alfacsf.pd.men$`Alpha-syn`)  
## [1] 66  
summary(alfacsf.pd.men$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  355.8 1003.6 1359.6 1423.9 1769.6 3449.3
```

Grupo control mujeres:

```
length(alfacsf.ct.women$`Alpha-syn`)  
## [1] 45  
summary(alfacsf.ct.women$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  855.2 1296.2 1591.2 1759.8 2058.4 3852.8
```

Grupo PD mujeres:

```
length(alfacsf.pd.women$`Alpha-syn`)  
## [1] 43  
summary(alfacsf.pd.women$`Alpha-syn`)  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  475.4  977.8 1429.2 1524.2 1974.5 3305.9
```

Comparación niveles de alfa-syn entre PD y control en CSF:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfacsf.ct$`Alpha-syn`)  
##  
## Shapiro-Wilk normality test  
##  
## data:  alfacsf.ct$`Alpha-syn`  
## W = 0.91447, p-value = 3.49e-05  
shapiro.test(alfacsf.pd$`Alpha-syn`)  
##  
## Shapiro-Wilk normality test  
##  
## data:  alfacsf.pd$`Alpha-syn`  
## W = 0.95723, p-value = 0.001475
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfa.csf$`Alpha-syn`, group = alfa.csf$DIAGNOSIS, center = "median")  
## Levene's Test for Homogeneity of Variance (center = "median")  
##           Df F value Pr(>F)  
## group    1  0.7245 0.3957  
##           191
```



Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfa.csf)
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data: Alpha-syn by DIAGNOSIS  
## W = 5856, p-value = 0.0008987  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación niveles de alfa-syn entre PD y control en hombres en CSF:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfacsf.ct.men$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfacsf.ct.men$`Alpha-syn`  
## W = 0.92241, p-value = 0.01033
```

```
shapiro.test(alfacsf.pd.men$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfacsf.pd.men$`Alpha-syn`  
## W = 0.95216, p-value = 0.01263
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfacsf.men$`Alpha-syn`, group = alfacsf.men$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group  1  0.0791 0.7791  
##      103
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfacsf.men)
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data: Alpha-syn by DIAGNOSIS  
## W = 1700, p-value = 0.006226  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación niveles de alfa-syn entre PD y control en mujeres en CSF:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfacsf.ct.women$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfacsf.ct.women$`Alpha-syn`  
## W = 0.89759, p-value = 0.0007976
```

```
shapiro.test(alfacsf.pd.women$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfacsf.pd.women$`Alpha-syn`  
## W = 0.95444, p-value = 0.08667
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfacsf.women$`Alpha-syn`, group = alfacsf.women$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  1.0586 0.3064  
##      86
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfacsf.women)
```

```
##  
## Wilcoxon rank sum test  
##  
## data:  Alpha-syn by DIAGNOSIS  
## W = 1173, p-value = 0.08708  
## alternative hypothesis: true location shift is not equal to 0
```

### 6.3 Alfa-sinucleína en saliva



Niveles alfa-sinucleína en saliva entre diagnóstico y género.

Grupo control:

```
length(alfasaliva.ct$`Alpha-syn` )
## [1] 22
summary(alfasaliva.ct$`Alpha-syn` )
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  26.59  47.56   93.48  191.71 139.42 1044.20
```

Grupo PD:

```
length(alfasaliva.pd$`Alpha-syn` )
## [1] 20
summary(alfasaliva.pd$`Alpha-syn` )
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   34.55 106.20  160.74  311.53  255.66 1346.63
```

Grupo control hombres:

```
length(alfasaliva.ct.men$`Alpha-syn` )
## [1] 17
summary(alfasaliva.ct.men$`Alpha-syn` )
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   26.59  47.34   79.92  214.96  150.08 1044.20
```

Grupo PD hombre:

```
length(alfasaliva.pd.men$`Alpha-syn`)
## [1] 12

summary(alfasaliva.pd.men$`Alpha-syn`)
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  34.55  88.99 160.74  248.71 189.21 1346.63
```

Grupo control mujeres:

```
length(alfasaliva.ct.women$`Alpha-syn`)
## [1] 5

summary(alfasaliva.ct.women$`Alpha-syn`)
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  37.83  97.73 100.05  112.66 107.01  220.67
```

Grupo PD mujeres:

```
length(alfasaliva.pd.women$`Alpha-syn`)
## [1] 8

summary(alfasaliva.pd.women$`Alpha-syn`)
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  38.08 108.51 180.20  405.75 472.85 1284.90
```

Comparación niveles de alfa-syn entre PD y control en saliva:

Ho: datos siguen normalidad

Ha: datos no siguen normalidad

```
shapiro.test(alfasaliva.ct$`Alpha-syn`)
##
## Shapiro-Wilk normality test
##
## data:  alfasaliva.ct$`Alpha-syn`
## W = 0.56297, p-value = 5.43e-07

shapiro.test(alfasaliva.pd$`Alpha-syn`)
##
## Shapiro-Wilk normality test
##
## data:  alfasaliva.pd$`Alpha-syn`
## W = 0.63696, p-value = 7.291e-06
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfa.saliva$`Alpha-syn`, group = alfa.saliva$DIAGNOSIS, center = "median")
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group 1  0.7312 0.3976
##      40
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfa.saliva)
```

```
##  
## Wilcoxon rank sum test  
##  
## data: Alpha-syn by DIAGNOSIS  
## W = 141, p-value = 0.04727  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación niveles de alfa-syn entre PD y control en hombres en saliva:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfasaliva.ct.men$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: alfasaliva.ct.men$`Alpha-syn`  
## W = 0.593, p-value = 8.912e-06
```

```
shapiro.test(alfasaliva.pd.men$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: alfasaliva.pd.men$`Alpha-syn`  
## W = 0.56845, p-value = 5.925e-05
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfasaliva.men$`Alpha-syn`, group = alfasaliva.men$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1 2e-04 0.9895  
##      27
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfasaliva.men)
```

```
##  
## Wilcoxon rank sum test  
##  
## data: Alpha-syn by DIAGNOSIS  
## W = 75, p-value = 0.2449  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación niveles de alfa-syn entre PD y control en mujeres en CSF:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfasaliva.ct.women$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfasaliva.ct.women$`Alpha-syn`  
## W = 0.86636, p-value = 0.252
```

```
shapiro.test(alfasaliva.pd.women$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfasaliva.pd.women$`Alpha-syn`  
## W = 0.72404, p-value = 0.0042
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = alfasaliva.women$`Alpha-syn`, group = alfasaliva.women$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  2.0061 0.1844  
##      11
```

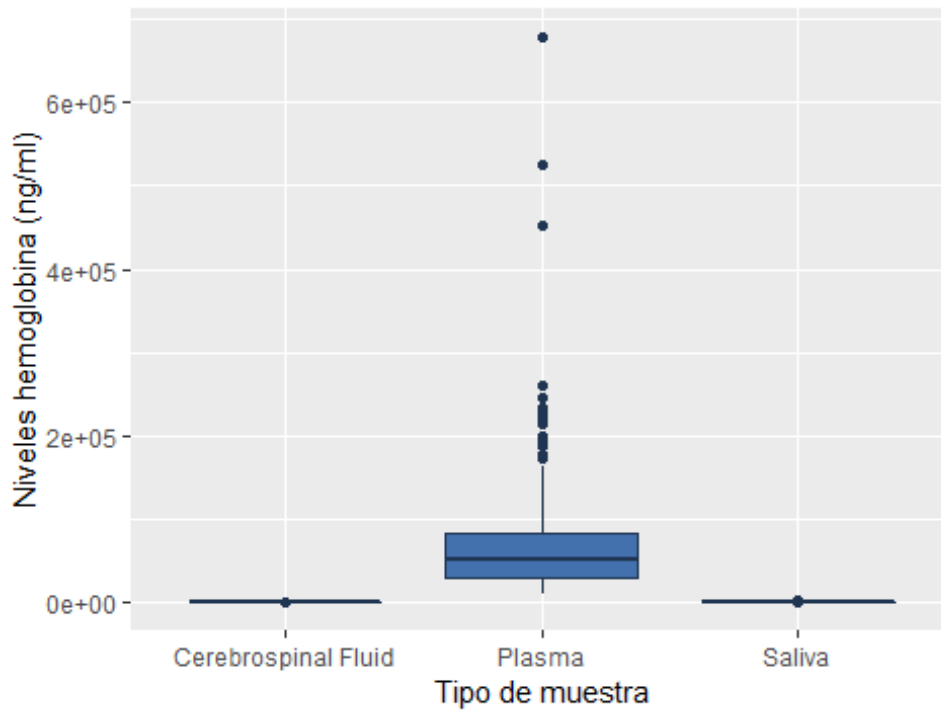
Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(`Alpha-syn`~DIAGNOSIS, data = alfasaliva.women)
```

```
##  
## Wilcoxon rank sum test  
##  
## data:  Alpha-syn by DIAGNOSIS  
## W = 8, p-value = 0.09324  
## alternative hypothesis: true location shift is not equal to 0
```

## 6.4 Efecto contaminación hemoglobina



Niveles hemoglobina entre tipo de muestra.

Análisis correlación en plasma:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfa.hem.plasma$`Alpha-syn`)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfa.hem.plasma$`Alpha-syn`  
## W = 0.87335, p-value = 8.012e-12
```

```
shapiro.test(alfa.hem.plasma$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  alfa.hem.plasma$TESTVALUE  
## W = 0.61412, p-value < 2.2e-16
```

```
cor(x=alfa.hem.plasma$TESTVALUE,y=alfa.hem.plasma$`Alpha-syn`,method = "spearman")
```

```
## [1] 0.02754814
```

Análisis correlación en líquido cefalorraquídeo:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfa.hem.csf$`Alpha-syn`)
```

```
##
## Shapiro-Wilk normality test
##
## data:  alfa.hem.csf$`Alpha-syn`
## W = 0.96906, p-value = 0.0427

shapiro.test(alfa.hem.csf$TESTVALUE)

##
## Shapiro-Wilk normality test
##
## data:  alfa.hem.csf$TESTVALUE
## W = 0.69748, p-value = 8.702e-12

cor(x=alfa.hem.csf$TESTVALUE,y=alfa.hem.csf$`Alpha-syn`,method = "spearman")

## [1] 0.1025576
```

Análisis correlación en saliva:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(alfa.hem.saliva$`Alpha-syn` )

##
## Shapiro-Wilk normality test
##
## data:  alfa.hem.saliva$`Alpha-syn`
## W = 0.75595, p-value = 3.424e-05

shapiro.test(alfa.hem.saliva$TESTVALUE)

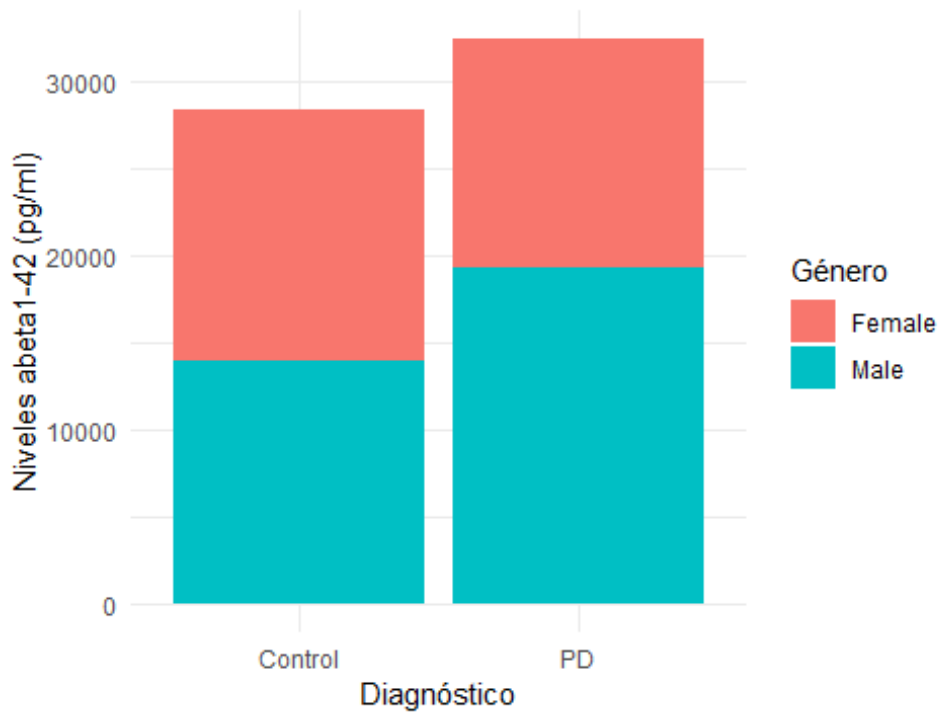
##
## Shapiro-Wilk normality test
##
## data:  alfa.hem.saliva$TESTVALUE
## W = 0.757, p-value = 3.554e-05

cor(x=alfa.hem.saliva$TESTVALUE,y=alfa.hem.saliva$`Alpha-syn`,method = "spearman")

## [1] 0.4092308
```

6.5  $\beta$ 1-42 amiloide





Niveles Abeta1-42 entre diagnóstico y género.

Grupo control:

```
length(ami.ct$TESTVALUE)
```

```
## [1] 85
```

```
summary(ami.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  11.39  290.11  354.88  333.40  389.78  500.70
```

Grupo PD:

```
length(ami.pd$TESTVALUE)
```

```
## [1] 109
```

```
summary(ami.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   64.0  248.9   289.2   297.6  360.6   468.3
```

Grupo control hombre:

```
length(ami.men.ct$TESTVALUE)
```

```
## [1] 40
```

```
summary(ami.men.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   11.39  299.18  361.25  348.48  402.97  500.70
```

Grupo PD hombre:

```
length(ami.men.pd$TESTVALUE)
```

```
## [1] 66
```

```
summary(ami.men.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   64.0  248.3  282.8   292.0  362.7   468.3
```

Grupo control mujer:

```
length(ami.women.ct$TESTVALUE)
```

```
## [1] 45
```

```
summary(ami.women.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   103.2  287.9  343.6   320.0  378.5   461.6
```

Grupo PD mujer:

```
length(ami.women.pd$TESTVALUE)
```

```
## [1] 43
```

```
summary(ami.women.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   144.7  260.3  296.7   306.3  356.7   466.7
```

Comparación grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(ami.ct$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  ami.ct$TESTVALUE
## W = 0.93959, p-value = 0.0006045
```

```
shapiro.test(ami.pd$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  ami.pd$TESTVALUE
## W = 0.98725, p-value = 0.3917
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = ami$TESTVALUE, group = ami$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group  1  0.1193 0.7302
##      192
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = ami)

##
## Wilcoxon rank sum test with continuity correction
##
## data: TESTVALUE by DIAGNOSIS
## W = 5993.5, p-value = 0.0004544
## alternative hypothesis: true location shift is not equal to 0
```

Comparación en hombres grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(ami.men.ct$TESTVALUE)

##
## Shapiro-Wilk normality test
##
## data: ami.men.ct$TESTVALUE
## W = 0.91054, p-value = 0.00393

shapiro.test(ami.men.pd$TESTVALUE)

##
## Shapiro-Wilk normality test
##
## data: ami.men.pd$TESTVALUE
## W = 0.9779, p-value = 0.2867
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = ami.men$TESTVALUE, group = ami.men$DIAGNOSIS, center = "median")

## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group  1  0.0293 0.8644
##      104
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = ami.men)

##
## Wilcoxon rank sum test with continuity correction
##
## data: TESTVALUE by DIAGNOSIS
## W = 1853, p-value = 0.0005191
## alternative hypothesis: true location shift is not equal to 0
```

Comparación en mujeres grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(ami.women.ct$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  ami.women.ct$TESTVALUE  
## W = 0.93048, p-value = 0.009749
```

```
shapiro.test(ami.women.pd$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  ami.women.pd$TESTVALUE  
## W = 0.99189, p-value = 0.9889
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = ami.women$TESTVALUE, group = ami.women$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  0.4932 0.4844  
##      86
```

Ho: no hay diferencias entre los dos grupos

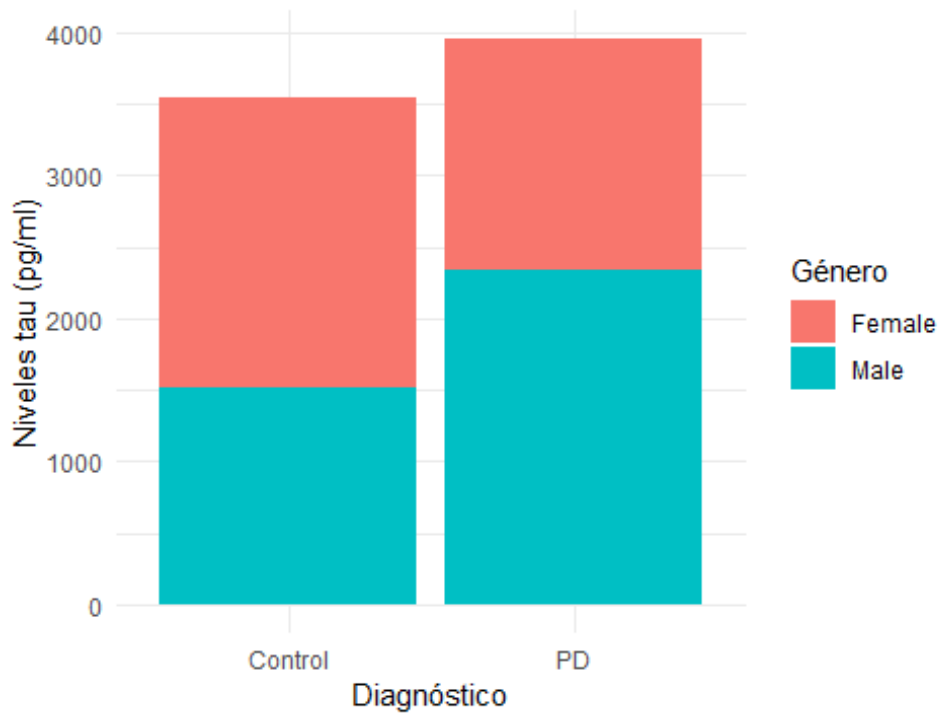
Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = ami.women)
```

```
## Warning in wilcox.test.default(x = c(418.91, 244.6, 159.8, 372.62, 173.6, :  
## cannot compute exact p-value with ties
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data:  TESTVALUE by DIAGNOSIS  
## W = 1127.5, p-value = 0.183  
## alternative hypothesis: true location shift is not equal to 0
```

## 6.6 Tau total



Niveles tau entre diagnóstico y género.

Grupo control:

```
length(tau.ct$TESTVALUE)
```

```
## [1] 85
```

```
summary(tau.ct$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.04   28.98   36.07   41.69  45.25  172.40
```

Grupo PD:

```
length(tau.pd$TESTVALUE)
```

```
## [1] 107
```

```
summary(tau.pd$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  10.00   25.08   33.06   36.95  45.87   98.89
```

Grupo control hombre:

```
length(tau.men.ct$TESTVALUE)
```

```
## [1] 40
```

```
summary(tau.men.ct$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.04   27.97   37.17   37.81  45.13   73.40
```

Grupo PD hombre:

```
length(tau.men.pd$TESTVALUE)
```

```
## [1] 65
```

```
summary(tau.men.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  10.00  24.00   32.66   36.02  45.86   79.32
```

Grupo control mujer:

```
length(tau.women.ct$TESTVALUE)
```

```
## [1] 45
```

```
summary(tau.women.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  22.66  31.39   35.64   45.13  49.54  172.40
```

Grupo PD mujer:

```
length(tau.women.pd$TESTVALUE)
```

```
## [1] 42
```

```
summary(tau.women.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  16.35  26.79   33.55   38.38  46.49   98.89
```

Comparación grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(tau.ct$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  tau.ct$TESTVALUE
## W = 0.71211, p-value = 1.284e-11
```

Ho:datos siguen normalidad Ha:datos no siguen normalidad

```
shapiro.test(tau.pd$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  tau.pd$TESTVALUE
## W = 0.91907, p-value = 6.649e-06
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = tau$TESTVALUE, group = tau$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group 1  0.1194  0.73
##      190
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = tau)
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data: TESTVALUE by DIAGNOSIS  
## W = 5258, p-value = 0.0634  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación en hombres grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(tau.men.ct$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: tau.men.ct$TESTVALUE  
## W = 0.96463, p-value = 0.2404
```

```
shapiro.test(tau.men.pd$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: tau.men.pd$TESTVALUE  
## W = 0.94327, p-value = 0.00497
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = tau.men$TESTVALUE, group = tau.men$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group  1  0.5042 0.4793  
##      103
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = tau.men)
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data: TESTVALUE by DIAGNOSIS  
## W = 1464, p-value = 0.2806  
## alternative hypothesis: true location shift is not equal to 0
```

Comparación en mujeres grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(tau.women.ct$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: tau.women.ct$TESTVALUE  
## W = 0.65499, p-value = 4.996e-09
```

```
shapiro.test(tau.women.pd$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: tau.women.pd$TESTVALUE  
## W = 0.87483, p-value = 0.0002759
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = tau.women$TESTVALUE, group = tau.women$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  0.3558 0.5525  
##      85
```

Ho: no hay diferencias entre los dos grupos

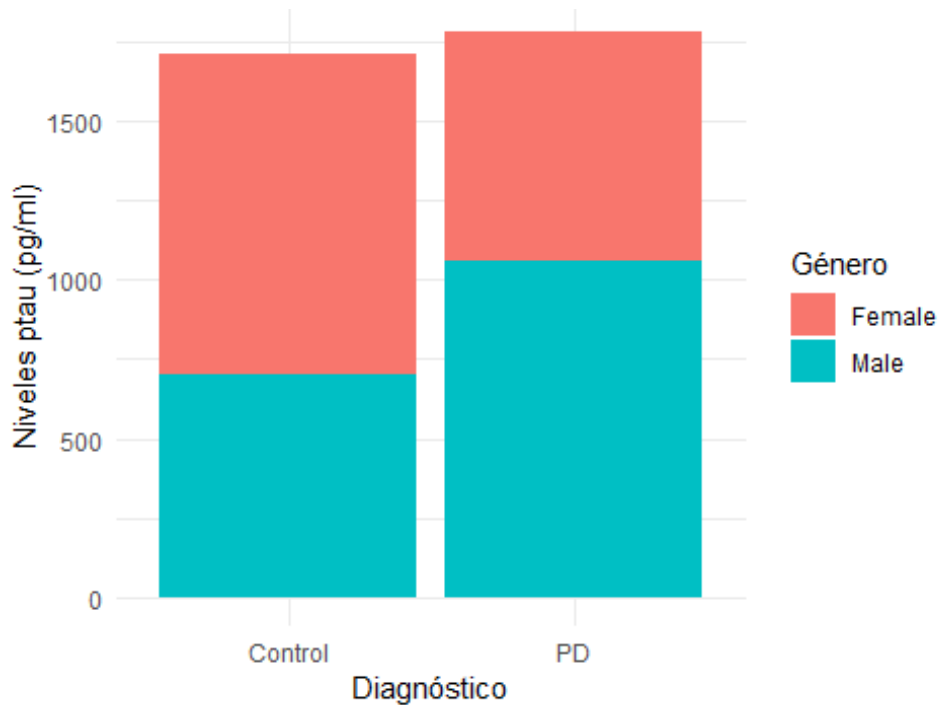
Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = tau.women)
```

```
##  
## Wilcoxon rank sum test  
##  
## data: TESTVALUE by DIAGNOSIS  
## W = 1105, p-value = 0.1764  
## alternative hypothesis: true location shift is not equal to 0
```

## 6.7 Tau fosforilada





Niveles ptau entre diagnóstico y género.

### Grupo control:

```
length(ptau.ct$TESTVALUE)
```

```
## [1] 84
```

```
summary(ptau.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   5.44  10.75   14.62   20.36  23.84  127.90
```

### Grupo PD:

```
length(ptau.pd$TESTVALUE)
```

```
## [1] 107
```

```
summary(ptau.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   4.300  8.355  12.730  16.632  22.275  76.640
```

### Grupo control hombre:

```
length(ptau.men.ct$TESTVALUE)
```

```
## [1] 39
```

```
summary(ptau.men.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   5.78  10.55   13.71   18.06  22.54   53.50
```

### Grupo PD hombre:

```
length(ptau.men.pd$TESTVALUE)
```

```
## [1] 65
```

```
summary(ptau.men.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   4.30   9.07   12.73   16.34   21.76   45.81
```

Grupo control mujer:

```
length(ptau.women.ct$TESTVALUE)
```

```
## [1] 45
```

```
summary(ptau.women.ct$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   5.44  10.93   17.43   22.34   24.27  127.90
```

Grupo PD mujer:

```
length(ptau.women.pd$TESTVALUE)
```

```
## [1] 42
```

```
summary(ptau.women.pd$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   4.30   7.57   12.89   17.09   23.16   76.64
```

Comparación grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(ptau.ct$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  ptau.ct$TESTVALUE
## W = 0.65188, p-value = 8.268e-13
```

```
shapiro.test(ptau.pd$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  ptau.pd$TESTVALUE
## W = 0.83307, p-value = 1.207e-09
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = ptau$TESTVALUE, group = ptau$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group  1  0.8529 0.3569
##      189
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = ptau)

##
## Wilcoxon rank sum test with continuity correction
##
## data: TESTVALUE by DIAGNOSIS
## W = 5274, p-value = 0.03983
## alternative hypothesis: true location shift is not equal to 0
```

Comparación en hombres grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(ptau.men.ct$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data: ptau.men.ct$TESTVALUE
## W = 0.82778, p-value = 3.314e-05
```

```
shapiro.test(ptau.men.pd$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data: ptau.men.pd$TESTVALUE
## W = 0.89348, p-value = 4.039e-05
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = ptau.men$TESTVALUE, group = ptau.men$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group  1  0.0107 0.9177
##      102
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = ptau.men)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: TESTVALUE by DIAGNOSIS
## W = 1421, p-value = 0.3043
## alternative hypothesis: true location shift is not equal to 0
```

Comparación en mujeres grupo PDvscontrol:

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(ptau.women.ct$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
```

```
## data: ptau.women.ct$TESTVALUE
## W = 0.62287, p-value = 1.614e-09
```

```
shapiro.test(ptau.women.pd$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data: ptau.women.pd$TESTVALUE
## W = 0.76421, p-value = 8.306e-07
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
library(car)
leveneTest(y = ptau.women$TESTVALUE, group = ptau.women$DIAGNOSIS, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group 1  0.4753 0.4924
##      85
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~DIAGNOSIS, data = ptau.women)
```

```
## Warning in wilcox.test.default(x = c(13.2, 23.5, 67.8, 29.78, 13.77,
## 12.82, : cannot compute exact p-value with ties
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: TESTVALUE by DIAGNOSIS
## W = 1153, p-value = 0.07798
## alternative hypothesis: true location shift is not equal to 0
```

## 6.8 Epidermal growth factor (EGF)

```
length(egf.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(egf.men$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      766.5  882.0  1306.8  1778.6  2505.4  4234.4
```

```
length(egf.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(egf.women$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      417.5  825.9  1046.9  1196.4  1441.3  2664.5
```

Ho: datos siguen normalidad

Ha: datos no siguen normalidad

```
shapiro.test(egf.men$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  egf.men$TESTVALUE
## W = 0.84965, p-value = 0.05753
```

```
shapiro.test(egf.women$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  egf.women$TESTVALUE
## W = 0.89116, p-value = 0.1747
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
bartlett.test(egf$TESTVALUE ~ egf$GENDER)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  egf$TESTVALUE by egf$GENDER
## Bartlett's K-squared = 2.8832, df = 1, p-value = 0.08951
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
t.test(TESTVALUE~GENDER, data = egf)
```

```
##
## Welch Two Sample t-test
##
## data:  TESTVALUE by GENDER
## t = -1.4063, df = 13.987, p-value = 0.1815
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1470.2251  305.8051
## sample estimates:
## mean in group Female  mean in group Male
##          1196.41          1778.62
```

## 6.9 Apolipoprotein AI (ApoA1)

```
length(apoa1.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(apoa1.men$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 13428 14209   15938 15865 17395 18223
```

```
length(apoa1.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(apoa1.women$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 10783 14636   15498 15507 17166 19943
```

Ho: datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(apoa1.men$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  apoa1.men$TESTVALUE  
## W = 0.90895, p-value = 0.2738
```

```
shapiro.test(apoa1.women$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  apoa1.women$TESTVALUE  
## W = 0.97256, p-value = 0.9135
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
bartlett.test(apoa1$TESTVALUE ~ apoa1$GENDER)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data:  apoa1$TESTVALUE by apoa1$GENDER  
## Bartlett's K-squared = 1.4374, df = 1, p-value = 0.2306
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
t.test(TESTVALUE~GENDER, data = apoa1)
```

```
##  
## Welch Two Sample t-test  
##  
## data:  TESTVALUE by GENDER  
## t = -0.34273, df = 15.583, p-value = 0.7364  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -2577.71 1861.57  
## sample estimates:  
## mean in group Female  mean in group Male  
## 15507.33 15865.40
```

6.10 Protein deglycase DJ-1 (DJ-1)

```
length(dj1.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(dj1.men$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  362.3  399.7   462.5   566.2   703.2  1109.9
```

```
length(dj1.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(dj1.women$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   423.5  475.9   526.0   659.1  635.0  1287.1
```

Ho: datos siguen normalidad

Ha: datos no siguen normalidad

```
shapiro.test(dj1.men$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  dj1.men$TESTVALUE
## W = 0.7933, p-value = 0.01202
```

```
shapiro.test(dj1.women$TESTVALUE)
```

```
##
## Shapiro-Wilk normality test
##
## data:  dj1.women$TESTVALUE
## W = 0.72634, p-value = 0.001839
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = dj1$TESTVALUE, group = dj1$GENDER, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group 1  0.0716  0.792
##      18
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~GENDER, data = dj1)
```

```
##
## Wilcoxon rank sum test
##
## data:  TESTVALUE by GENDER
## W = 66, p-value = 0.2475
## alternative hypothesis: true location shift is not equal to 0
```

## 6.11 Brain-derived neurotrophic factor (BDNF)

```
length(bdnf.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(bdnf.men$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   699.8 1470.7 1779.9 1681.8 2130.9 2340.1
```

```
length(bdnf.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(bdnf.women$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   516.9  879.9 1287.1 1198.2 1524.5 1985.3
```

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(bdnf.men$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  bdnf.men$TESTVALUE  
## W = 0.88023, p-value = 0.1313
```

```
shapiro.test(bdnf.women$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  bdnf.women$TESTVALUE  
## W = 0.93573, p-value = 0.5066
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
bartlett.test(bdnf$TESTVALUE ~ bdnf$GENDER)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data:  bdnf$TESTVALUE by bdnf$GENDER  
## Bartlett's K-squared = 0.25749, df = 1, p-value = 0.6119
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
t.test(TESTVALUE~GENDER, data = bdnf )
```

```
##  
## Welch Two Sample t-test  
##  
## data:  TESTVALUE by GENDER  
## t = -2.0297, df = 17.48, p-value = 0.05789  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -985.44219  18.04219  
## sample estimates:  
## mean in group Female   mean in group Male  
##           1198.15           1681.85
```

## 6.12 Glial fibrally acidic protein (GFAP)

```
length(gfap.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(gfap.men$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  457.3  497.7   508.1   577.4  611.9   814.8
```

```
length(gfap.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(gfap.women$TESTVALUE)
```



```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   427.5  471.4   501.9   552.1  606.2   812.1
```

Ho: datos siguen normalidad

Ha: datos no siguen normalidad

```
shapiro.test(gfap.men$TESTVALUE)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  gfap.men$TESTVALUE
## W = 0.79892, p-value = 0.01407
```

```
shapiro.test(gfap.women$TESTVALUE)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  gfap.women$TESTVALUE
## W = 0.87787, p-value = 0.1233
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = gfap$TESTVALUE, group = gfap$GENDER, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group 1  0.0153 0.9029
##      18
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~GENDER, data = gfap)
```

```
##
##  Wilcoxon rank sum test
##
## data:  TESTVALUE by GENDER
## W = 41, p-value = 0.5288
## alternative hypothesis: true location shift is not equal to 0
```

## 6.13 Fractalkine (Frac)

```
length(frac.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(frac.men$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1560  1666   3276   2844  3642   4198
```

```
length(frac.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(frac.women$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2183  2795   3172   3723  4289   6279
```

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(frac.men$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: frac.men$TESTVALUE  
## W = 0.8271, p-value = 0.03088
```

```
shapiro.test(frac.women$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: frac.women$TESTVALUE  
## W = 0.87707, p-value = 0.1207
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = frac$TESTVALUE, group = frac$GENDER, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  0.0748 0.7876  
##      18
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~GENDER, data = frac)
```

```
##  
## Wilcoxon rank sum test  
##  
## data: TESTVALUE by GENDER  
## W = 63, p-value = 0.3527  
## alternative hypothesis: true location shift is not equal to 0
```

## 6.14 Nicotinamide phosphoribosyltransferase (NAMPT)

```
length(nampt.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(nampt.men$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 471.9  548.1  665.7  724.0  793.1 1147.8
```

```
length(nampt.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(nampt.women$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 457.3  572.8  636.0  674.0  677.8 1079.7
```

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(nampt.men$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  nampt.men$TESTVALUE  
## W = 0.88667, p-value = 0.1555
```

```
shapiro.test(nampt.women$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  nampt.women$TESTVALUE  
## W = 0.86675, p-value = 0.09159
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
bartlett.test(nampt$TESTVALUE ~ nampt$GENDER)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data:  nampt$TESTVALUE by nampt$GENDER  
## Bartlett's K-squared = 0.3989, df = 1, p-value = 0.5277
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
t.test(TESTVALUE~GENDER, data = nampt )
```

```
##  
## Welch Two Sample t-test  
##  
## data:  TESTVALUE by GENDER  
## t = -0.5282, df = 17.213, p-value = 0.6041  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -249.6801 149.6201  
## sample estimates:  
## mean in group Female    mean in group Male  
##           673.97           724.00
```

## 6.15 PTEN induced kinase 1 (PINK1)

```
length(pink1.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(pink1.men$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  494.0  586.6   772.9   795.6  894.9  1275.3
```

```
length(pink1.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(pink1.women$TESTVALUE)
```

```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  731.7  929.1   966.5  1248.2 1118.1  2627.6
```

Ho:datos siguen normalidad Ha:datos no siguen normalidad

```
shapiro.test(pink1.men$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: pink1.men$TESTVALUE  
## W = 0.91229, p-value = 0.2971
```

```
shapiro.test(pink1.women$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: pink1.women$TESTVALUE  
## W = 0.68949, p-value = 0.0006572
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
leveneTest(y = pink1$TESTVALUE, group = pink1$GENDER, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value Pr(>F)  
## group 1  0.819 0.3774  
##      18
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
wilcox.test(TESTVALUE~GENDER, data = pink1)
```

```
##  
## Wilcoxon rank sum test  
##  
## data: TESTVALUE by GENDER  
## W = 78, p-value = 0.03546  
## alternative hypothesis: true location shift is not equal to 0
```

## 6.16 Leucine-rich repeat kinase 2 (LRRK)

```
length(lrrk.men$TESTVALUE)
```

```
## [1] 10
```

```
summary(lrrk.men$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      1843   1981   2620   2712   3336   3906
```

```
length(lrrk.women$TESTVALUE)
```

```
## [1] 10
```

```
summary(lrrk.women$TESTVALUE)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      1558   1923   2131   2134   2360   2919
```

Ho:datos siguen normalidad

Ha:datos no siguen normalidad

```
shapiro.test(lrrk.men$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: lrrk.men$TESTVALUE  
## W = 0.89648, p-value = 0.2003
```

```
shapiro.test(lrrk.women$TESTVALUE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: lrrk.women$TESTVALUE  
## W = 0.96377, p-value = 0.8279
```

Ho: datos tienen igualdad de varianza

Ha: datos no tienen igualdad de varianza

```
bartlett.test(lrrk$TESTVALUE ~ lrrk$GENDER)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data: lrrk$TESTVALUE by lrrk$GENDER  
## Bartlett's K-squared = 3.4977, df = 1, p-value = 0.06146
```

Ho: no hay diferencias entre los dos grupos

Ha: hay diferencias entre los dos grupos

```
t.test(TESTVALUE~GENDER, data = lrrk )
```

```
##  
## Welch Two Sample t-test  
##  
## data: TESTVALUE by GENDER  
## t = -2.112, df = 13.468, p-value = 0.05389  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -1167.00357 11.14357  
## sample estimates:  
## mean in group Female mean in group Male  
## 2133.71 2711.64
```