

Grado en Ingeniería Informática
Ingeniería de Computadores

Trabajo de Fin de Grado

**Análisis del envejecimiento cerebral en
pacientes preclínicos de la enfermedad de
Alzheimer**

Autor

Adrián San Segundo Moya

2022

Grado en Ingeniería Informática
Ingeniería de Computadores

Trabajo de Fin de Grado

**Análisis del envejecimiento cerebral en
pacientes preclínicos de la enfermedad de
Alzheimer**

Autor

Adrián San Segundo Moya

Directores

Ibai Gurrutxaga, Olatz Arbelaitz

Resumen

En este proyecto se estudia las distintas áreas del cerebro que son susceptibles a fallar en relación con el envejecimiento cerebral para poder realizar un diagnóstico a sujetos preclínicos a partir de los 40 años.

El trabajo tratará de detectar las diferencias del envejecimiento cerebral entre sujetos sanos y sujetos en estados iniciales de la enfermedad de Alzheimer.

Para esto se ha trabajado con CITA-alzheimer, fundación que ha proporcionado una base de datos de dos proyectos diferentes: DEBA y PGA1. Será necesario un preprocesado de las bases de datos: normalización, fusión, etc. Se tratará de detectar los cambios cerebrales más significativos que se producen con la edad en base a los sujetos del estudio.

El objetivo principal de este proyecto es comprobar si la hipótesis de CITA es cierta. Para ello se han generado scatterplots con relación edad-volumen, añadiendo sus respectivas curvas. Se ha visto que se confirma en las variables (áreas cerebrales o combinaciones entre ellas) que tienen pendientes fuertes en dichas curvas, teniendo en cuenta a todos los sujetos. Sin embargo, en las variables con pendientes fuertes en base a sujetos cuyo test de la punción lumbar sea negativo o positivo, la hipótesis difiere mucho.

Índice general

Resumen	I
Índice general	III
Índice de figuras	VII
Índice de tablas	IX
1. Introducción	1
1.1. Alzheimer	1
1.2. Sistema nervioso	2
1.3. Objetivo del proyecto	3
1.3.1. Relación entre edad y volumen cerebral. Hipótesis.	4
2. Planificación	5
2.1. Horas estimadas	5
2.2. Horas realizadas	5
2.3. Diagrama de Gantt	7
2.4. Herramientas para el desarrollo del proyecto	7
2.5. Análisis de riesgos	9

3. Fundamentos	11
3.1. Envejecimiento cerebral	11
3.1.1. Envejecimiento	11
3.1.2. El cerebro	12
3.1.3. Características del cerebro humano	14
3.1.4. Las neuronas	14
3.1.5. Funcionamiento del cerebro	15
3.1.6. Punción lumbar	16
3.1.7. Resonancia magnética	16
3.1.8. El Alzheimer	17
3.2. Herramientas utilizadas	18
3.2.1. Estandarización	18
3.2.2. Algoritmo de suavizado	18
4. Trabajo realizado	21
4.1. Introducción a los datos	21
4.2. Análisis visual de los datos	24
4.2.1. Análisis de datos por franjas de edad	27
4.3. Análisis para confirmar la hipótesis inicial de CITA	32
4.3.1. Ratios entre variables	32
4.3.2. Normalización	33
4.3.3. Suavizado	33
4.3.4. Estandarizado	33
4.3.5. Cálculo de pendientes	33
4.3.6. Variables con mayor pendiente moderada y fuerte	36
4.4. Análisis comparativo entre el total de sujetos y sujetos en base al test	43

5. Conclusiones y trabajo futuro	49
---	-----------

Anexos

A. Apéndice	53
--------------------	-----------

Bibliografía	55
---------------------	-----------

Índice de figuras

1.1. Cortex cerebral y el hipocampo.	4
1.2. Hipótesis CITA Alzheimer	4
2.1. Organización del proyecto	8
2.2. IDE de RStudio	9
3.1. Materia gris y blanca del cerebro	13
3.2. Encéfalo humano	13
3.3. Neuronas	15
4.1. Histograma de los sujetos por test	23
4.2. GM_VOL	24
4.3. WM_VOL	25
4.4. CSF_VOL	25
4.5. TIV	26
4.6. Hipocampo. Su volumen es menor en los positivos	26
4.7. GM total por franjas de edad	27
4.8. WM total por franjas de edad	28
4.9. CSF total por franjas de edad	28
4.10. Total por franjas de edad	29

4.11. GM_VOL	30
4.12. WM_VOL	30
4.13. CSF_VOL	31
4.14. Total (GM, WM y CSF)	31
4.15. Histograma de pendientes por tramo	34
4.16. Pendientes moderadas por grupos	35
4.17. Pendientes fuertes por grupos	36
4.18. Patrón 1	39
4.19. Patrón 2	39
4.20. Patrón 3	40
4.21. Patrón 4	40
4.22. Pendientes fuertes curva roja	46
4.23. Pendientes fuertes curva verde	46
4.24. Diferencias entre puntos en base a la curva roja	48
4.25. Diferencias entre puntos en base a la curva verde	48

Índice de tablas

2.1. Tabla con las horas estimadas	6
2.2. Tabla con las horas realizadas	6
2.3. Tabla de las diferencias entre las horas estimadas y las realizadas	7
4.1. Proyectos DEBA y PGA1	22
4.2. Proyecto DEBA	22
4.3. Proyecto PGA1	23
4.4. Mínimo	23
4.5. Máximo	23
4.6. Media	24
4.7. Mediana	24
4.8. Desviación estándar	24
4.9. Nº de sujetos por resultado y edad	27
4.10. Variables con más pendientes fuertes	37
4.11. Variables con más pendientes moderadas	38
4.12. Patrón que siguen las variables con pendientes moderadas	42
4.13. Patrón que siguen las variables con pendientes fuertes	43
4.14. Variables con más pendientes fuertes con la curva roja	44
4.15. Variables con más pendientes fuertes con la curva verde	45

1. CAPÍTULO

Introducción

El envejecimiento es un proceso irreversible que afecta a todos los órganos y sistemas del individuo aunque en cada uno no se manifiesta de la misma manera. Esta fase hace que las células puedan tener como destino la muerte por un deterioro morfofuncional.

Durante la edad mediana temprana muchas funciones del cuerpo comienzan a disminuir gradualmente. Envejecer es la evolución que hace tan característica a la vida.

Es necesario saber que el ser humano, desde su nacimiento posee un número de neuronas limitado y estas no son capaces de reproducirse. A lo largo de la vida, se pierden un cierto número de neuronas por el desarrollo humano e incluso por ciertos episodios patológicos que pueden aparecer.

1.1. Alzheimer

Es la enfermedad neurodegenerativa [3] más común entre las personas mayores. Esta demencia es un trastorno cerebral que afecta gravemente a la capacidad de una persona para poder realizar actividades cotidianas. Es de origen todavía desconocido y frente a la que no se puede ofrecer ningún tratamiento que sea capaz de curarla o prevenirla.

En 1854 nació en Baviera el psiquiatra alemán Alois Alzheimer que publicó en 1907 el caso de una mujer de cincuenta y un años cuyas facultades intelectuales habían desaparecido gradualmente en cuatro años. Una vez realizada la autopsia de su cerebro descubrió dos tipos de anomalías:

1. Placas seniles, que habían sido identificadas como una lesión característica de demencias seniles.
2. Formaciones anormales situadas en el interior de las neuronas, eran desconocidas y las nombró “ovillos neurofibrilares”.

Al ser la paciente tan joven no se permitió este diagnóstico por lo que se aisló el caso como una afección específica recibiendo años más tarde el nombre de “Alzhéimer”.

Hasta los años 60, el Alzhéimer era descrito como una demencia degenerativa presenil que afectaba a personas alrededor de los 50 años. No es lo mismo que las “demencias seniles” que eran mucho más frecuentes y afectaban a personas entre 75 y 80 años.

A mediados de los años 60, se resaltó la similitud entre las lesiones cerebrales que ocasionaba el Alzhéimer y la demencia senil degenerativa, distinguiendo su forma precoz como la enfermedad del Alzhéimer y una forma más tardía como la Demencia Senil tipo Alzhéimer. Hoy en día, entendemos la enfermedad del Alzhéimer como esta última, ya sea en una fase temprana o en una fase más tardía de la enfermedad.

El alzhéimer suele comenzar a partir de los 60 años y el riesgo aumenta a medida que la persona envejece. También es mayor si hay familiares que padecieron la enfermedad.

1.2. Sistema nervioso

El individuo para tener control sobre su cuerpo necesita el cerebro y el sistema nervioso, que no solo se encargan de controlar movimientos, los sentidos, pensamientos y recuerdos... sino también el funcionamiento de órganos como el corazón.

Las órdenes se ejecutan mediante señales que se envían al cerebro mediante los nervios. La médula espinal es un manojo de nervios que va desde el cerebro hasta el centro de la espalda, y los nervios se extienden desde la médula espinal a cada parte del cuerpo.

Dicho esto y en relación con el envejecimiento, el cerebro y el sistema nervioso sufren cambios naturales. Tanto el cerebro como la médula espinal pierden peso y neuronas. Por esto y a medida que el tiempo pasa, las neuronas transmiten los mensajes de manera más lenta. En otras palabras, una menor eficacia en el pensamiento, la memoria y la capacidad cognitiva es una parte normal del envejecimiento, aunque esta reducción no es la misma en todas las personas.

En las personas mayores, la demencia y una pérdida importante de la memoria no es normal en el envejecimiento y puede ser síntoma de enfermedades neurodegenerativas.

1.3. Objetivo del proyecto

El tamaño del cerebro, como se ha visto, se va reduciendo conforme van avanzando los años. Al ser un análisis del envejecimiento cerebral en pacientes preclínicos, se va a estudiar:

- Las diferentes áreas cerebrales.
- Qué relación pueden llegar a tener entre ellas.
- Las zonas cerebrales que más relación tienen con el alzhéimer.
- Análisis de las áreas cerebrales de los sujetos sanos y sujetos en estados preclínicos del alzhéimer, por separado (se verá más adelante).
- Diferencias entre áreas cerebrales en base a pacientes con resultado negativo y positivo.
- Acotar la zona de búsqueda en cuanto a las áreas que pueden tener mayor relevancia para un posterior desarrollo de la enfermedad.

Se parte de una base de datos de pacientes en las que no hay evidencia científica de que el paciente que se esté analizando sea sano o pueda ser asintomático y esté desarrollando la enfermedad.

Las patologías preclínicas que aporta CITA son las siguientes:

- Alzhéimer preclínico. Biomarcadores positivos en personas asintomáticas en aproximadamente el 20-40% de los pacientes a partir de los 40 años.
- Patología vascular cerebral. En la resonancia magnética se detecta hiperintensidades de la sustancia blanca; infartos; microhemorragias. Aproximadamente el 30-50% de los pacientes a partir de los 60 años.

1.3.1. Relación entre edad y volumen cerebral. Hipótesis.

¿Qué es lo normal en el envejecimiento normal? Efectos del envejecimiento, el amiloide y la enfermedad de alzhéimer en la corteza cerebral y el hipocampo.

CITA Alzhéimer adjunta las siguientes gráficas. Ambas indican las trayectorias de vida de las reducciones volumétricas. Son estimaciones transversales de la duración de la vida adulta en relación con el volumen total de la corteza cerebral y del volumen total del hipocampo. El volumen se expresa en unidades de desviación estándar. Son las siguientes (ver figura 1.1):

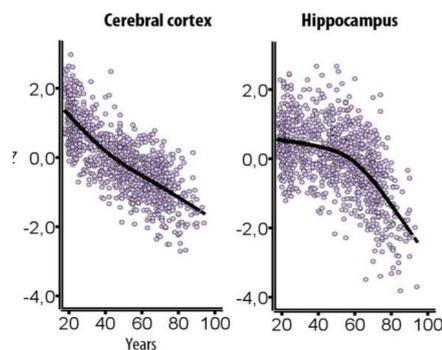


Figura 1.1: Corteza cerebral y el hipocampo.

Y la hipótesis de CITA Alzhéimer es la trayectoria excluyendo patologías preclínicas (ver figura 1.2).

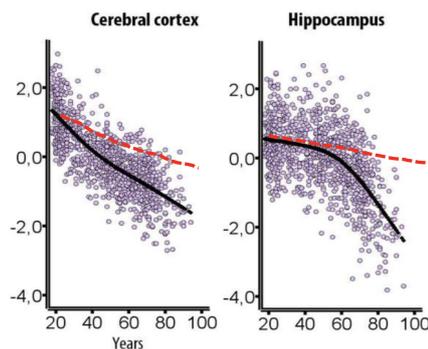


Figura 1.2: Hipótesis CITA Alzhéimer

El objetivo principal que se busca en este proyecto es comprobar si la hipótesis de CITA Alzhéimer está en lo cierto.

2. CAPÍTULO

Planificación

En este capítulo se introduce a continuación, la planificación y la dedicación de horas que se ha llevado a cabo en el proyecto.

Ha sido un proyecto en el que, las tareas que se han ido ejecutando, han sido tanto peticiones de los médicos de CITA como de análisis interno para poder obtener conclusiones de los resultados. Por tanto, no ha habido una planificación de principio a fin. Sin embargo, se han ido estimando las tareas que se han ido requiriendo para una mejor gestión del proyecto.

2.1. Horas estimadas

Como se puede observar en la tabla [2.1](#), se estiman las tareas más costosas aquellas que se tienen que analizar los datos. Esto es debido, a no solo lo que requiere analizar y conseguir resultados en base a lo que se quiere llegar a obtener, sino que el tiempo de ejecución por cada “script” con la gran cantidad de datos que se controlan, puede ser de elevado.

2.2. Horas realizadas

Tras haber finalizado el proyecto, las horas resultantes en cada tarea a lo largo del proyecto se pueden ver en la tabla [2.2](#).

Tareas	Horas estimadas
Recopilación de información sobre el Alzheimer	20
Recopilación de información sobre el cerebro y su funcionamiento	35
Introducción y análisis de la base de datos de CITA	30
Normalización de la base de datos de CITA	30
Análisis visual de los datos	55
Análisis sobre la confirmación de la hipótesis de CITA	70
Análisis comparativo	60
Suma total	300

Tabla 2.1: Tabla con las horas estimadas

Tareas	Horas realizadas
Recopilación de información sobre el Alzheimer	20
Recopilación de información sobre el cerebro y su funcionamiento	30
Introducción y análisis de la base de datos de CITA	25
Normalización de la base de datos de CITA	40
Análisis visual de los datos	55
Análisis sobre la confirmación de la hipótesis de CITA	80
Análisis comparativo	50
Suma total	300

Tabla 2.2: Tabla con las horas realizadas

Se puede comprobar las diferencias entre las horas estimadas y las reales (realizadas) en la tabla 2.3.

Tareas	Horas estimadas	Horas realizadas	Diferencia
Recopilación de información sobre el Alzheimer	20	20	0
Recopilación de información sobre el cerebro y su funcionamiento	35	30	-5
Introducción y análisis de la base de datos de CITA	30	25	-5
Normalización de la base de datos de CITA	30	40	10
Análisis visual de los datos	55	55	0
Análisis sobre la confirmación de la hipótesis de CITA	70	80	10
Análisis comparativo	60	50	-10
Suma total	300	300	0

Tabla 2.3: Tabla de las diferencias entre las horas estimadas y las realizadas

2.3. Diagrama de Gantt

En la figura 2.1 se muestra la organización del proyecto. Se divide en 3 secciones: conocimientos, desarrollo del proyecto y la memoria del trabajo de fin de grado.

2.4. Herramientas para el desarrollo del proyecto

Este proyecto consiste gran parte en el análisis de datos, manipulación de datos y generación de gráficas, entre otros. Al basarse en la estadística se ha decidido usar “R” como lenguaje de programación.

Es comúnmente utilizado para la computación estadística y gráfica, ya que dispone de una amplia variedad de técnicas estadísticas (modelos lineales y no lineales, pruebas estadísticas clásicas, análisis de series de tiempo, clasificación, agrupamiento, etc.) y gráficas.

Algunas características de este lenguaje son:

- Manejo y almacenamiento efectivo de los datos.
- Conjunto de operadores para la realización de cálculos con matrices.
- Gran colección de herramientas para el análisis de datos.
- Utilidades gráficas para la visualización de datos.

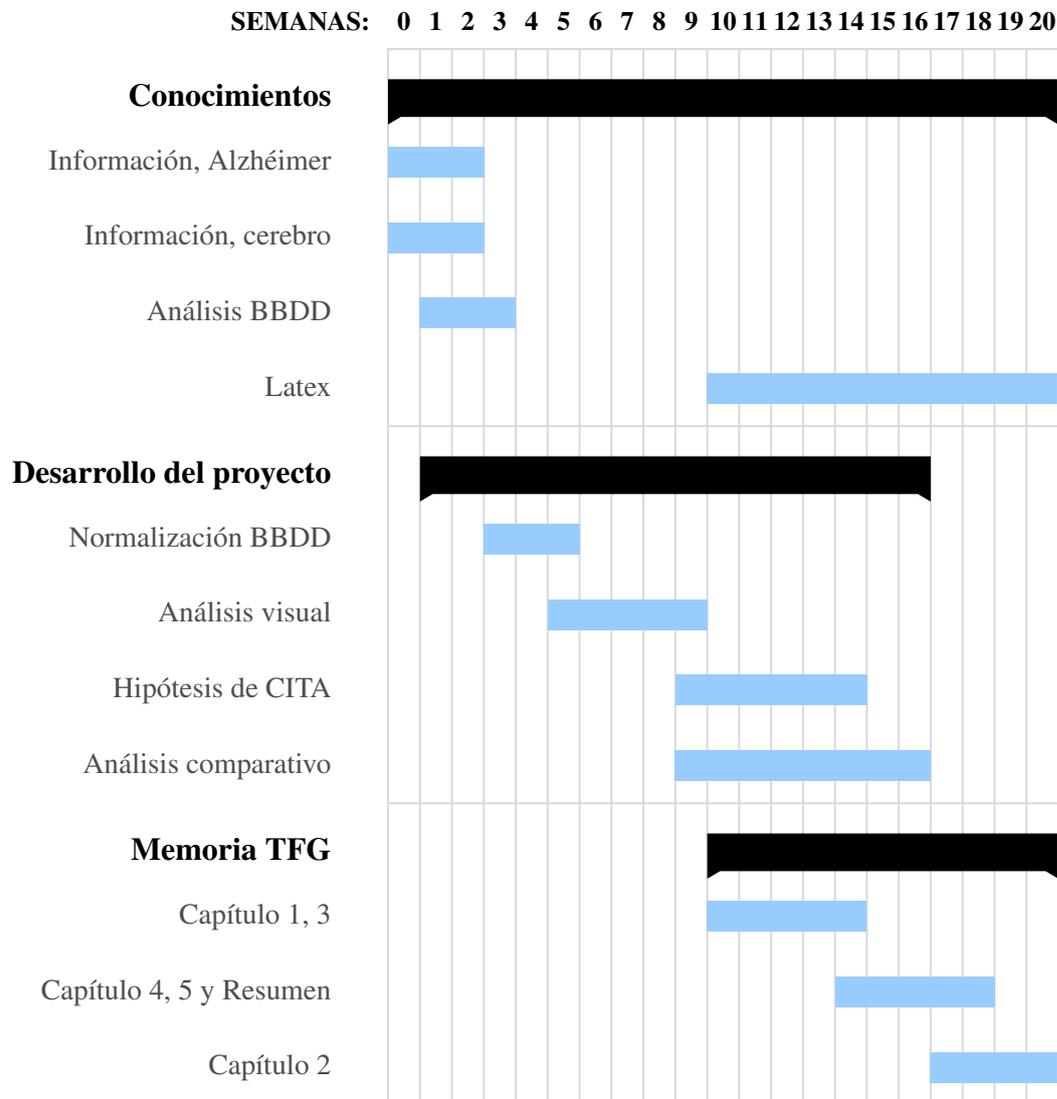


Figura 2.1: Organización del proyecto

El entorno de desarrollo que se ha usado ha sido RStudio. Gracias a R y a este IDE, se ha obtenido un gran ahorro de tiempo. Por ejemplo, RStudio almacena de manera interna las variables junto con sus valores a medida que se van generando. De este modo, se ahorra la implementación de más código y de ejecuciones. Esto es una gran ventaja, ya que en distintos scripts se pueden usar una variable ya generada anteriormente. En la figura 2.2 se puede ver un ejemplo de RStudio.

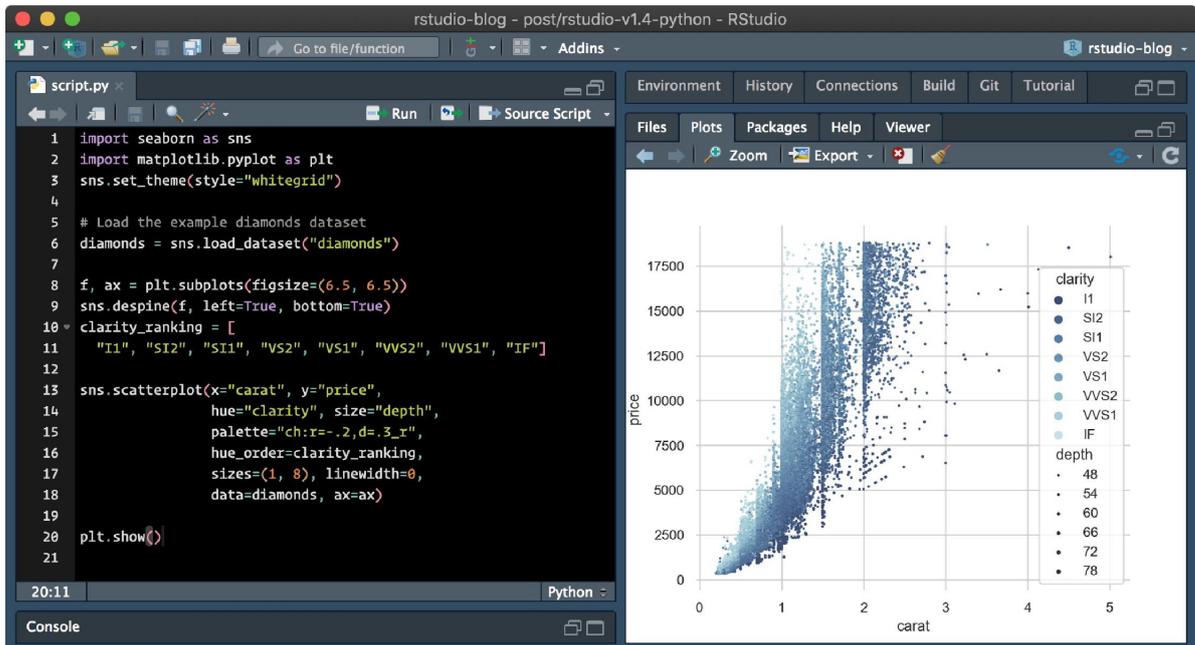


Figura 2.2: IDE de RStudio

2.5. Análisis de riesgos

Antes de dar comienzo al proyecto, se deben valorar los riesgos y sus posibles soluciones a estos problemas. Se mencionan los más relevantes y las soluciones que se dan a estos riesgos.

- **Errores en procesos.** Es posible que pueda ocurrir un error en la manipulación de datos en el “scripting” y arrastrar el error que se haya producido, debido a que hay una gran cantidad de datos y es complicado, a primera vista, encontrar si ha habido algún error. La solución que se plantea es testear y analizar con detenimiento los resultados obtenidos en cada resultado para minimizar al máximo cualquier margen de error.
- **Incumplimiento en las estimaciones de las tareas.** Existen factores que pueden retrasar la fecha de finalización de las tareas propuestas. Para ello, las estimaciones se han llevado a cabo teniendo en cuenta posibles retrasos en tareas.
- **Pérdida de datos.** Para esto, es necesario realizar un “backup” de los datos y resultados que se van obteniendo, porque de no ser así, si es necesario recuperar los datos perdidos las horas que se van a realizar van a aumentar considerablemente.

- **Falta de comprensión o capacidad para la realización de tareas.** En la misma situación, ya sea falta de comprensión en la tarea a realizar o una falta de “habilidad” para poder llevarla a cabo, se hablará con el tutor en busca de una solución para poder continuar con las estimaciones mencionadas y no retrasar el estado del proyecto.

3. CAPÍTULO

Fundamentos

3.1. Envejecimiento cerebral

3.1.1. Envejecimiento

El proceso de senectud no comienza hasta los 60 años, sin embargo, las funciones cerebrales pueden permanecer intactas hasta la muerte. Por ello, se puede afirmar que el envejecimiento cerebral no es proporcional con el paso del tiempo. Por esta razón, una persona anciana sin ninguna enfermedad no tiene deterioro cognitivo, así pues, la pérdida de memoria no es un fenómeno normal. Es una enfermedad.

A pesar de que el cerebro durante la vida sufre una serie de alteraciones tanto micro como macroscópicas y bioquímicas como disminución del volumen cerebral, muerte neuronal, etc., un cerebro bioquímicamente viejo puede ser funcionalmente joven. Las claves que producen esta fase involutiva son de carácter genético y ambiental. No solo depende de los genes de la persona, también de la alimentación que uno lleve, deporte e incluso el lugar en el que se viva.

Si hay algo que caracteriza al cerebro es que sus células neuronales no se reproducen como el resto de las células del cuerpo. Sin embargo, el sistema nervioso tiene una facultad denominada plasticidad neuronal. Es la capacidad que tienen las neuronas para generar nuevas dendritas y sinapsis que es lo que hace capaz de mantener la eficiencia de los circuitos neuronales aunque existan disfunciones por lesiones en áreas neuronales.

Teniendo esto en cuenta, es posible que la aparición de enfermedades neuronales como el Alzheimer sea debido a un deterioro de la plasticidad neuronal.

Las neuronas pueden morir por la detención de las propias funciones y desintegración rápida de la célula (apoptosis) o por apoptosis. Esta última es el conocido suicidio celular, una muerte programada en el que convergen en pequeñas vesículas que son captadas por células vecinas. Se dice que las enfermedades neurodegenerativas se activan de forma anormal por estos procesos de apoptosis.

Con el envejecimiento [8] se debe tener en cuenta también las hormonas en las que se destacan las sexuales, tiroideas, suprarrenales y la melatonina. Estas no solo son importantes por el control de la vida o muerte de las neuronas, sino también porque cualquier cambio hormonal tiene repercusión tanto positiva como negativa en el cerebro.

3.1.2. El cerebro

El cerebro [7] se define como un órgano complejo que se ubica dentro del cráneo y se encarga de gestionar la actividad del sistema nervioso. Se encuentra situado en la parte anterior y superior de la cavidad craneal.

En el interior del cráneo, el cerebro está flotando en un líquido transparente al que se le conoce como líquido cefalorraquídeo que cumple con funciones de protección, tanto las físicas como las inmunológicas. Está formado por millones de neuronas que están conectadas por axones y dendritas que permiten gestionar todas y cada una de las actividades que realiza el cuerpo o incluso el propio cerebro.

Además del líquido cefalorraquídeo mencionado, también existen otros dos tipos de materia en el cerebro: la materia blanca y la materia gris. La materia blanca está compuesta por axones de las neuronas y se ve de este color blanquecino por la mielina. La materia gris está compuesta por somas y dendritas de las neuronas (ver figura 3.1).

En lo que a las actividades del cerebro se refiere, se podrían resumir en gestionar la mayoría de las funciones del cuerpo y la mente. No solo se encarga de funciones vitales como respirar, el sueño, hambre... también el razonamiento, atención, emociones, etc. Todo esto es regulado por el encéfalo. En la antigüedad, uno de los médicos más importantes como lo fue Hipócrates ya pensaba que el cerebro humano es una de las creaciones más complejas y enigmáticas. En la actualidad, gracias a avances tecnológicos con la neuroimagen y medicina, neurociencia, psicología, etc., se ha podido averiguar grandes misterios en cuanto a su funcionamiento.

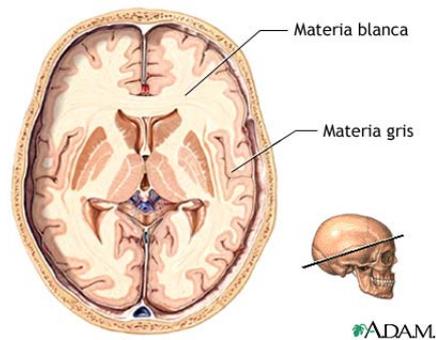


Figura 3.1: Materia gris y blanca del cerebro

Todos los seres vertebrados tienen cerebro, aunque dependiendo de la anatomía de este la forma, el tamaño y ciertas funciones varían. Aún así, se muestra a continuación un encéfalo humano que está constituido por las siguientes partes 3.2:

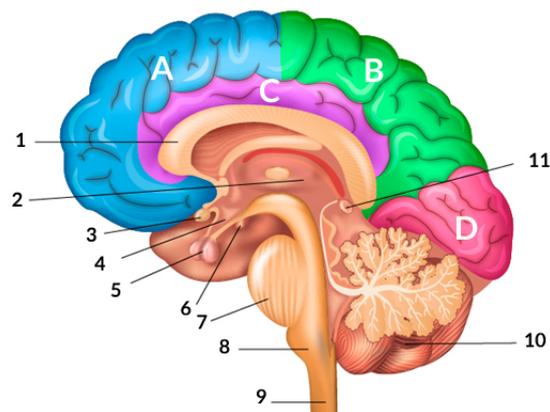


Figura 3.2: Encéfalo humano

- Cerebro. Formado por estructuras corticales y subcorticales que se encuentran bajo la corteza cerebral. Estas se dividen en las siguientes áreas: lóbulo frontal (A), lóbulo parietal (B), corteza cingulada (C), el lóbulo occipital (D), el lóbulo temporal y la corteza insular (estas dos últimas no aparecen en la figura). También encontramos el cuerpo calloso (1), el tálamo (2), los ganglios basales, amígdala, hipocampo y cuerpos mamilares (6).
- Cerebelo (10). Segundo órgano más grande del encéfalo. Involucrado principalmente en el control postural y del movimiento. También realiza algunas funciones cognitivas.

- Hipotálamo (4), glándula pituitaria o hipófisis (5) y glándula pineal (11). La glándula pituitaria y la glándula pineal se comunican con el hipotálamo a través de la liberación de hormonas para regular las funciones viscerales, como la regulación de la temperatura corporal y comportamientos básicos. La glándula pineal tiene un importante papel en la sincronización de la liberación de la hormona de la melatonina, implicada en la regulación de los ciclos de sueño/vigilia, para lo cual se coordina con el quiasma óptico (3).
- Tronco cerebral. Comienza en el extremo superior de la médula espinal (9). Formado por el bulbo raquídeo (8), el puente de Varolio o protuberancia (7) y el mesencéfalo.

3.1.3. Características del cerebro humano

Algunas de las características del cerebro humano son las siguientes:

- El córtex cerebral es uno de los más complejos que existen. Es un tamaño que supera al del resto de especies. Además, también es capaz de doblarse y plegarse sobre sí mismo formando circunvoluciones y surcos que le dan esa apariencia tan característica.
- El encéfalo tiene un peso aproximado de 1.5 kilos y un volumen alrededor de 1130 cm³ en mujeres y 1260 cm³ en hombres.
- Tanto el cerebro como la médula espinal están protegidas por membranas denominadas meninges, que las protegen de golpes contra el cráneo.
- El cerebro también está protegido por el líquido cefalorraquídeo, líquido en el que flota.
- Aproximadamente el cerebro humano está formado por más de 100 billones de células nerviosas, mayoritariamente células gliales y neuronas.

3.1.4. Las neuronas

Las neuronas son células que se encargan de recibir, procesar y enviar la información, intra e intercelular. Esto se hace mediante impulsos nerviosos. A pesar de que las neuronas

tienen los mismos elementos citoplasmáticos y la misma información genética que el resto de las células del organismo, tienen unas características muy especiales. Las neuronas se constituyen por tres partes 3.3:

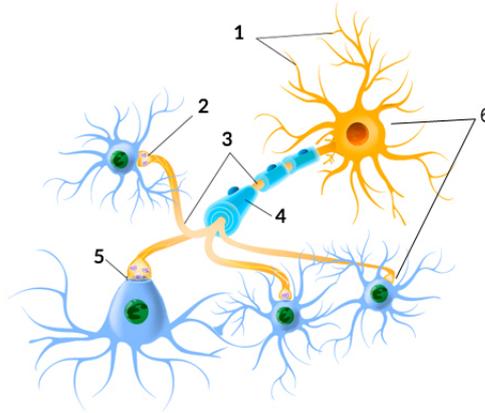


Figura 3.3: Neuronas

- Cuerpo o soma (6). Contiene el núcleo con el ADN. Se realizan la mayoría de funciones metabólicas de la célula. Es por esto que el soma es muy importante para la supervivencia de la célula neuronal.
- Axones (3). Prolongación que sale del soma celular. Comparable a un “cable” en el que el final tiene unos botones terminales (2) que son los puntos donde se produce la sinapsis (5), o lo que es lo mismo, la transmisión de información de una neurona a otra. Algunos axones están recubiertos por una capa de mielina (4) que agiliza y facilita la transmisión de información.
- Dendritas (1). Terminaciones nerviosas que salen del soma celular que se ramifican en forma de árbol. Constituyen el componente principal de recepción de información y posibilitan la comunicación entre dos neuronas.

3.1.5. Funcionamiento del cerebro

El cerebro funciona gracias a la sinapsis, transmisión de información que se produce por impulsos eléctrico-químicos. Neuronas y células se intercambian neurotransmisores que son los encargados de excitar o inhibir la acción de otra célula.

La transmisión de información se establece en milisegundos. De manera paralela se producen cientos de conexiones que nos permiten recibir, entender y responder de manera

adecuada. En cuestión de segundos recibimos miles de “inputs” y generamos miles de “outputs”. Las neuronas efectúan con precisión todo este proceso.

3.1.6. Punción lumbar

Una punción lumbar [2] es una prueba que se utiliza para diagnosticar determinadas afecciones médicas. Se realiza en la región lumbar. Durante la punción, se introduce una aguja en el espacio entre dos vértebras para extraer una muestra de líquido cefalorraquídeo. Este es el líquido que rodea el cerebro y la médula espinal para protegerlos de lesiones.

Se realiza una punción lumbar para poder diagnosticar con más facilidad infecciones graves. En ocasiones, esta se utiliza para inyectar medicamentos anestésicos o de quimioterapia en el líquido cefalorraquídeo.

La información que se obtiene mediante una punción lumbar puede ayudar a diagnosticar lo siguiente:

- Infecciones bacterianas, micóticas y virales graves, como meningitis, encefalitis y sífilis.
- Determinados tipos de cáncer que afectan al cerebro o a la médula espinal.
- Enfermedad de Alzheimer y otras formas de demencia.

3.1.7. Resonancia magnética

Se suele prescribir una resonancia magnética [6] [5] para estudiar y analizar una parte determinada del cuerpo humano. Este estudio permite detectar y diagnosticar lesiones y enfermedades, y hacer el seguimiento de algunos tratamientos. En la actualidad esta técnica es muy utilizada y está disponible en la mayoría de los hospitales.

Durante el transcurso de la resonancia magnética, el paciente permanece acostado en el interior de la máquina (lo habitual es que sea similar a un tubo largo y estrecho, aproximadamente de un metro de diámetro y abierto por sus extremos). Dentro de ella, la persona está rodeada por un campo magnético que reacciona con los elementos magnéticos dentro del cuerpo y transmite una débil señal de radio.

Esta señal se procesa por ordenador y posteriormente se visualiza en la pantalla. finalmente, el paciente debe de permanecer inmóvil durante toda la prueba para que los resultados sean lo más claros posibles.

Por lo tanto y con relación al cerebro, una resonancia magnética de la cabeza es un examen que utiliza imanes y ondas de radio potentes para poder crear imágenes del cerebro y de los tejidos nerviosos en el que no se emplea radiación.

3.1.8. El Alzheimer

El alzhéimer [4] es una enfermedad neurodegenerativa que puede ser provocada por la presencia de placas seniles y/o ovillos neurofibrilares. Esta hace que el cerebro se encoja y que las neuronas cerebrales mueran.

Es un proceso lento que afecta a zonas cerebrales que se encargan de controlar el pensamiento, la memoria y el lenguaje. Las personas que padecen la enfermedad comienzan teniendo dificultades para poder recordar eventos que ocurrieron de forma reciente o los nombres de personas que conocen. Un problema relacionado es el deterioro cognitivo leve [1], que es una afección que produce pérdidas de memoria más comunes que otras personas de la misma edad. Sin embargo, algunos olvidos pueden ser normales en el proceso del envejecimiento. A pesar de poder desarrollar esta afección, las personas que la padecen son capaces de desarrollar actividades del día a día y cuidarse a sí mismas. En un porcentaje alto, muchos individuos que padecen el deterioro cognitivo leve acabarán desarrollando alzhéimer.

Es la causa más común de demencia en la que existe un deterioro continuo en el pensamiento y comportamiento que afecta al individuo en la capacidad de vivir independientemente. A medida que el alzhéimer se desarrolla, el paciente presentará un grave deterioro de la memoria e irá perdiendo la habilidad de realizar sencillas tareas de la vida cotidiana.

Los síntomas más frecuentes que se pueden encontrar en esta enfermedad son: no reconocer a una persona que se acaba de saludar, proceder a realizar una tarea y no saber cuál es, marcar un número de teléfono y no saber a quién se está llamando... entre otros.

Para entender en cierto modo la memoria, es necesario dividirla en tres apartados: memoria sensorial (analítica o visual), memoria a corto plazo y memoria a largo plazo. Se especula que los problemas en la memoria relacionados con la edad se deben a un declive en la memoria a corto plazo. Se le conoce como la habilidad para almacenar y utilizar la información simultáneamente.

La memoria a largo plazo incluye dos subtipos: la memoria no declarativa (no intencional o automática) y la declarativa (episódica y semántica). Ambas son importantes, pero la no declarativa es la que incluye habilidades, aprendizaje perceptivo, motor y las respuestas condicionadas. Esta no padece ninguna modificación en el proceso del envejecimiento normal, así como en las fases precoces del alzhéimer.

Las personas ancianas tienen más dificultades a menudo con la memoria declarativa, ya sea la episódica (suele ser la primera) o la semántica.

3.2. Herramientas utilizadas

3.2.1. Estandarización

La estandarización significa ajustar los valores medidos en diferentes escalas respecto a una escala común, a menudo previo a un proceso de realizar promedios. En casos más complicados, la estandarización puede referirse a ajustes más sofisticados donde la intención es conseguir todas las distribuciones de probabilidad que se ajustan a los valores.

En este proyecto se ha usado la estandarización debido a que a la hora de analizar los datos de los distintos proyectos (DEBA y PGA1), los volúmenes de cada área del cerebro son ciertamente dispares.

Se ha usado la siguiente fórmula:

$$\frac{X - \mu}{\sigma}$$

donde:

- X es un valor.
- μ es la media.
- σ la desviación estándar.

3.2.2. Algoritmo de suavizado

El algoritmo de suavizado que se usa en el proyecto se usa para generar curvas que no son lineales, polinomiales, etc. Se ha decidido usar el algoritmo “Kernel Smoother”. Este

método devuelve una línea de cualquier forma, es decir, no tiene que ser recta, ni polinomial, ni exponencial, que resume el scatterplot de forma más o menos suavizada (existe un parámetro que permite controlar el nivel de suavizado).

El término kernel se utiliza en el análisis estadístico para referirse a una función de ventana.

Kernel smoother es una técnica estadística que estima un valor real como la media ponderada de los datos vecinos observados. La ponderación está definida por el núcleo (kernel), de modo que los puntos más cercanos reciben una ponderación mayor. La función estimada es suave, y el nivel de suavidad se establece mediante un único parámetro.

Sea $K_{h_\lambda}(X_0, X)$ un kernel definido por:

$$K_{h_\lambda}(X_0, X) = D\left(\frac{\|X - X_0\|}{h_\lambda(X_0)}\right)$$

donde:

- $X, X_0 \in R^p$
- $\|\cdot\|$ es la norma euclídea.
- $h_\lambda(X_0)$ es un parámetro (el radio del kernel).
- $D(t)$ es típicamente una función de valor real positivo, cuyo valor es decreciente (o no creciente) para la distancia creciente entre X y X_0 .

Los kernels populares utilizados para el suavizado incluyen los kernels parabólicos (Epanechnikov), Tricube y Gaussiano.

4. CAPÍTULO

Trabajo realizado

4.1. Introducción a los datos

Se ha utilizado para este trabajo una base de datos proporcionada por CITA-Alzhéimer que está compuesta por dos proyectos: DEBA y PGA1. En cada proyecto, se ha realizado una resonancia magnética a un total de 537 personas: 327 sujetos de PGA1 y 210 sujetos de DEBA. En PGA1, la edad de los sujetos está comprendida entre 45 y 80 años, mientras que en el estudio DEBA se ha realizado el estudio a sujetos mayores de 60 años.

La resonancia magnética junto con el software de cuantificación de imágenes de Neurocloud, han sido capaces de obtener el volumen de la materia gris (GM), materia blanca (WM) y el líquido cefalorraquídeo (CSF).

Cada sujeto ha realizado de manera voluntaria una punción lumbar, procedimiento que comúnmente se realiza para obtener muestras de líquido cefalorraquídeo. En este caso, se ha analizado el líquido para comprobar si el sujeto está desarrollando la enfermedad en estado preclínico. En PGA, 202 pacientes tienen el resultado del líquido, mientras que en DEBA son 158 los que tienen resultados del líquido disponibles.

Los datos de volumen cerebral de Neurocloud incluyen 7 medidas globales y las regiones cerebrales, que son 34 por hemisferio cerebral, es decir, 68 en total. Además, por región cerebral y hemisferio, se incluyen los volúmenes de cada materia y líquido. Lo que haría un total de 204 variables solo con las áreas cerebrales.

Las medidas globales son las siguientes:

TIV	Total Intracranial Volume
GM_VOL	Total Gray Matter Volume
WM_VOL	Total White Matter Volume
CSF_VOL	Total Cerebrospinal Fluid Volume
BPF	Brain Parenchymal Fraction
GM_BPF	Gray Matter Brain Parenchymal Fraction
WM_BPF	White Matter Brain Parenchymal Fraction

El TIV no es la suma de la materia gris, blanca y líquido cefalorraquídeo, es mayor que la suma de estas porque TIV implica todo el volumen intracraneal. Las áreas tienen materia blanca, gris y líquido, pero fuera de las áreas también hay líquido. Cuando el cerebro disminuye, el cráneo sigue siendo el mismo por lo que la suma de las áreas disminuye, pero el TIV se mantiene constante.

También está entre las variables el BPF que es el tanto por ciento del líquido cefalorraquídeo, GM_BPF que es el tanto por ciento de la materia gris y WM_BPF que es el tanto por ciento de la materia blanca.

Para el resultado del líquido cefalorraquídeo mencionado, se tiene en cuenta la variable CSF_Result en la que los valores pueden ser "0" si negativo, "1" si positivo o "NA" si el sujeto no se ha realizado la punción lumbar.

Se ha trabajado con los proyectos DEBA y PGA1 unidos, no por separado. Aún así, en una primera instancia era necesario ponerse en contexto en los proyectos con los que se estuvo trabajando y se obtuvo los siguientes resultados:

Sujetos	Negativos	Positivos	NA	Total
Hombres	102	65	65	232
Mujeres	126	67	112	305
Total	228	132	177	537

Tabla 4.1: Proyectos DEBA y PGA1

Sujetos	Negativos	Positivos	NA	Total
Hombres	35	44	20	99
Mujeres	37	42	32	111
Total	72	86	52	210

Tabla 4.2: Proyecto DEBA

Para tener en cuenta y una visión general de todos los sujetos, se ha construido un histo-

Sujetos	Negativos	Positivos	NA	Total
Hombres	67	21	45	133
Mujeres	89	25	80	194
Total	156	46	125	327

Tabla 4.3: Proyecto PGA1

grama (ver figura 4.1) siendo el eje “x” la edad y el eje “y” el número de sujetos. Se ha distribuido dependiendo el test entre negativo, positivo y no realizado.

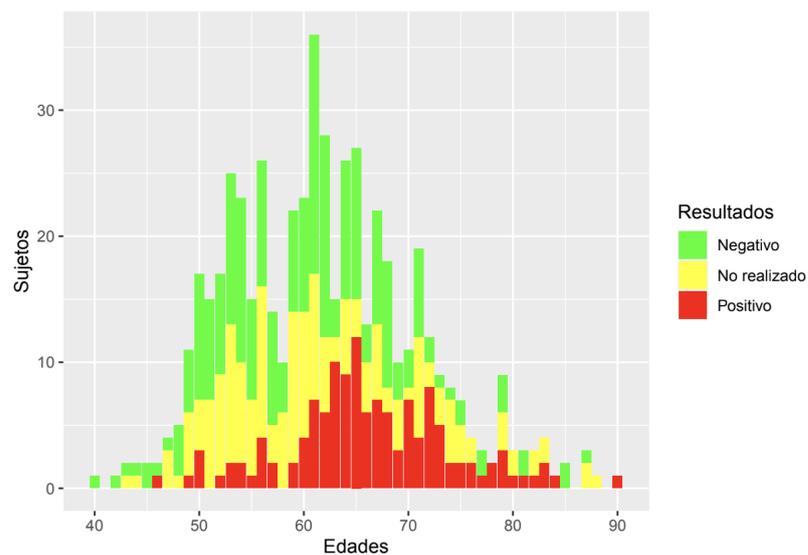


Figura 4.1: Histograma de los sujetos por test

Por otro lado, se obtienen los datos de mínimo, máximo, media, mediana y desviación estándar sobre la base de datos de los dos proyectos.

TIV	GM_VOL	WM_VOL	CSF_VOL
1081.1	410.23	282.67	174.63

Tabla 4.4: Mínimo

TIV	GM_VOL	WM_VOL	CSF_VOL
1828.45	703.64	705.65	608.49

Tabla 4.5: Máximo

TIV	GM_VOL	WM_VOL	CSF_VOL
1375.55	561.31	496.14	318.1

Tabla 4.6: Media

TIV	GM_VOL	WM_VOL	CSF_VOL
1363.22	557.85	491.69	310.87

Tabla 4.7: Mediana

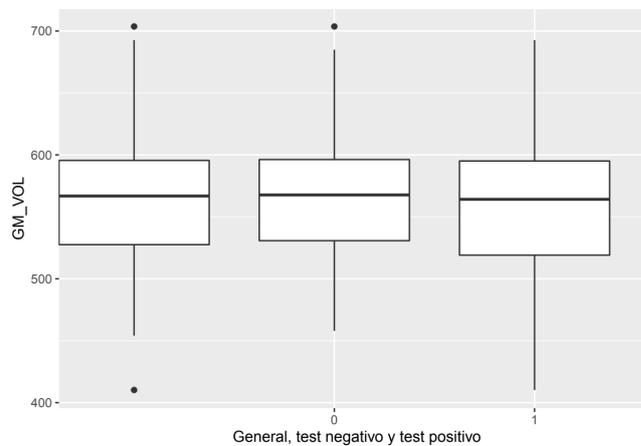
TIV	GM_VOL	WM_VOL	CSF_VOL
128.46	48.97	65.57	60.39

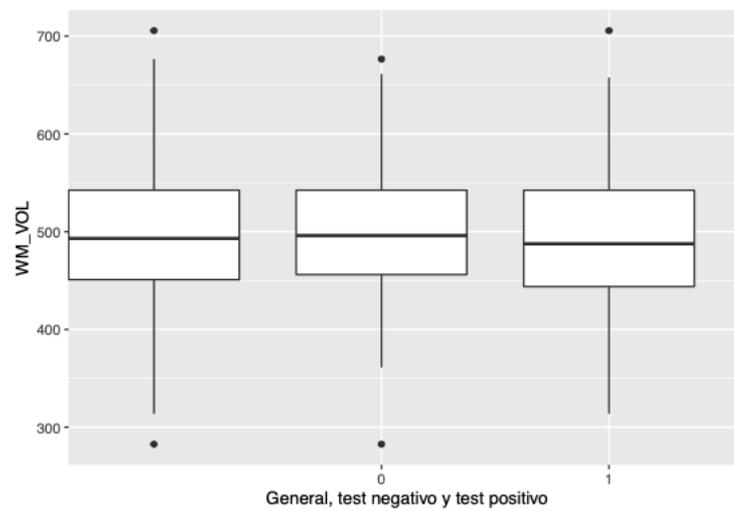
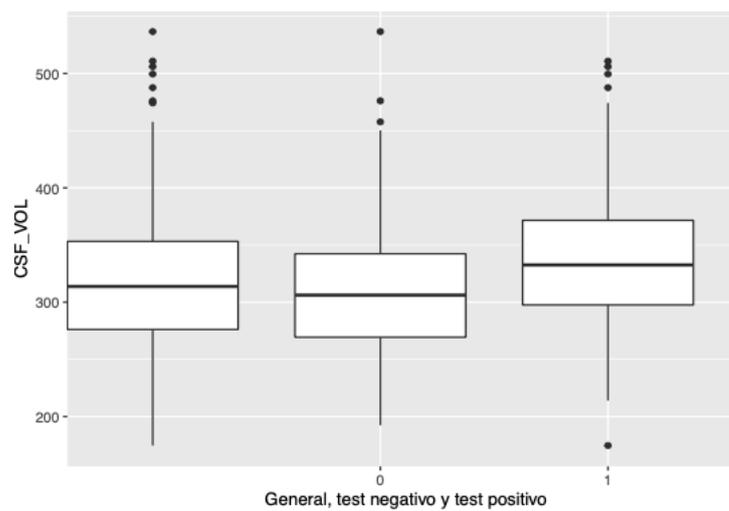
Tabla 4.8: Desviación estándar

4.2. Análisis visual de los datos

A continuación, para entender mejor la base de datos introducida se va a realizar un análisis visual. Para entrar a los datos con un poco de “zoom” se muestran los resultados de las sumas de las áreas por materia y también la suma total.

Algunos de los boxplots obtenidos son los que se muestran en las figuras posteriores, siendo de la caja izquierda a derecha los datos generales, los datos de los resultados negativos y positivos, respectivamente (ver figuras 4.2, 4.3, 4.4 y 4.5). El boxplot de la izquierda representa el total de los sujetos, el boxplot del medio representa los sujetos con resultado negativo del test de la punción lumbar y el boxplot de la derecha representa los sujetos con resultado positivo de dicho test.

**Figura 4.2:** GM_VOL

**Figura 4.3: WM_VOL****Figura 4.4: CSF_VOL**

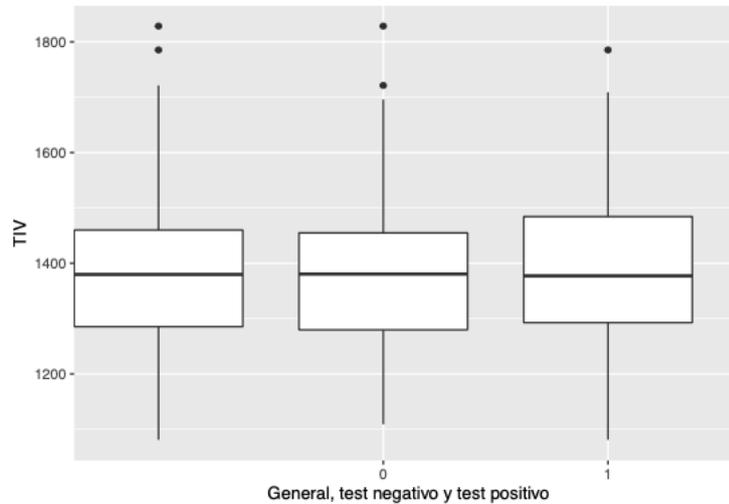


Figura 4.5: TIV

Se ha realizado el mismo análisis para todas las áreas cerebrales y, aunque por espacio no se muestran, se puede ver, por ejemplo, que el CSF es mayor en los positivos o que el hipocampo es menor en los positivos. También el rango de volúmenes en cada área. En áreas más pequeñas se tratan valores alrededor de los 5cm^3 mientras que en otras más grandes se tratan valores alrededor de los 100cm^3 . En cuanto a datos atípicos, no se encuentran outliers muy extremos que puedan denotar posibles errores en los datos.

Hay muchas variables y no es posible mostrar todas en el proyecto (se pueden ver en el apéndice 1), pero se ven características interesantes como puede ser, entre ellas, la siguiente (ver figura 4.6).

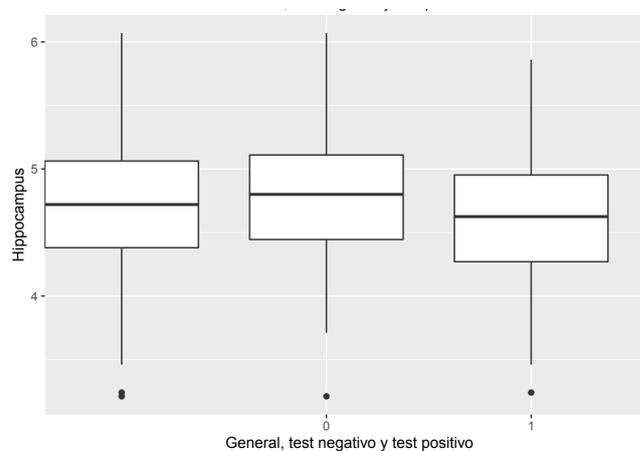


Figura 4.6: Hipocampo. Su volumen es menor en los positivos

4.2.1. Análisis de datos por franjas de edad

En este punto del proyecto y teniendo en cuenta el análisis realizado, es interesante analizar a continuación lo mismo (boxplots) pero por franjas de edad. En este caso, viendo el volumen de datos del que se dispone en la base de datos, se dividen los datos en franjas de edad. Se dividen en pacientes menores de 60 años, pacientes entre 60 y 70 años y pacientes mayores de 70 años (ver tabla 4.9).

Resultado	<60	60-70	>70
Negativos	111	94	23
Positivos	21	72	39
Todos	226	214	97

Tabla 4.9: N° de sujetos por resultado y edad

Estas figuras que se muestran a continuación son un ejemplo de ello (ver figuras 4.7, 4.8, 4.9 y 4.10).

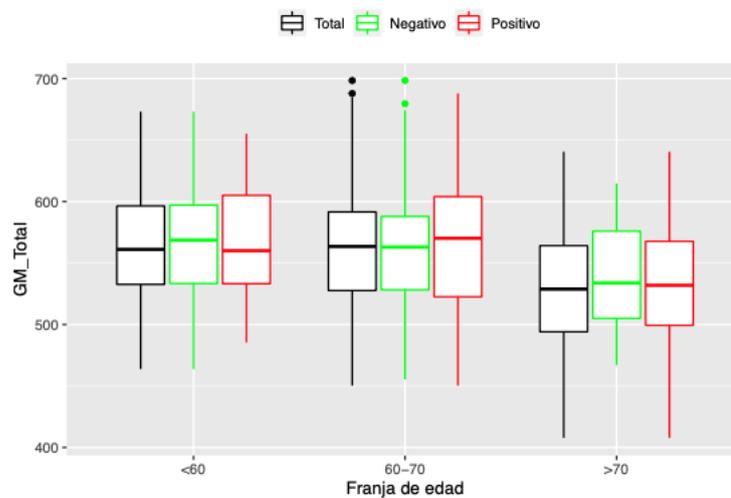


Figura 4.7: GM total por franjas de edad

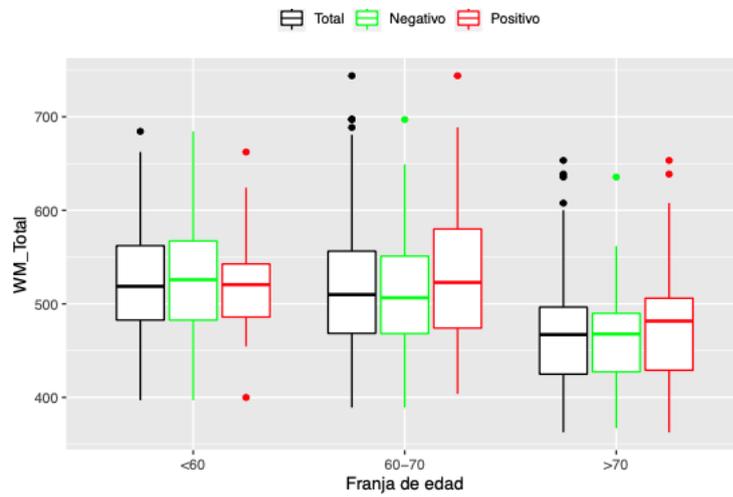


Figura 4.8: WM total por franjas de edad

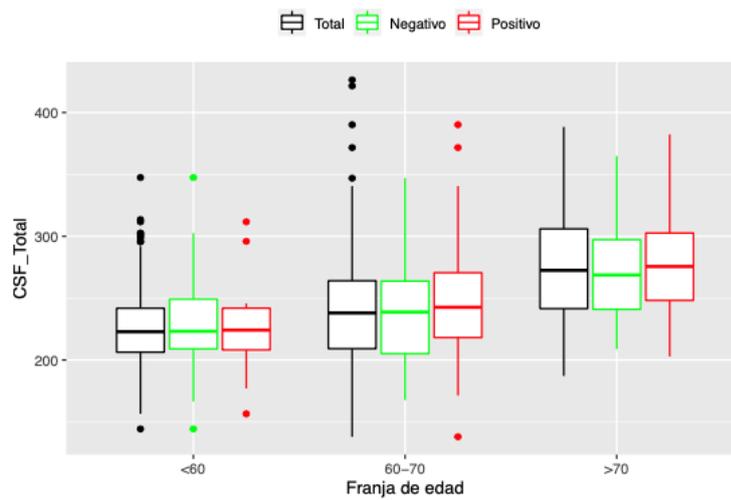


Figura 4.9: CSF total por franjas de edad

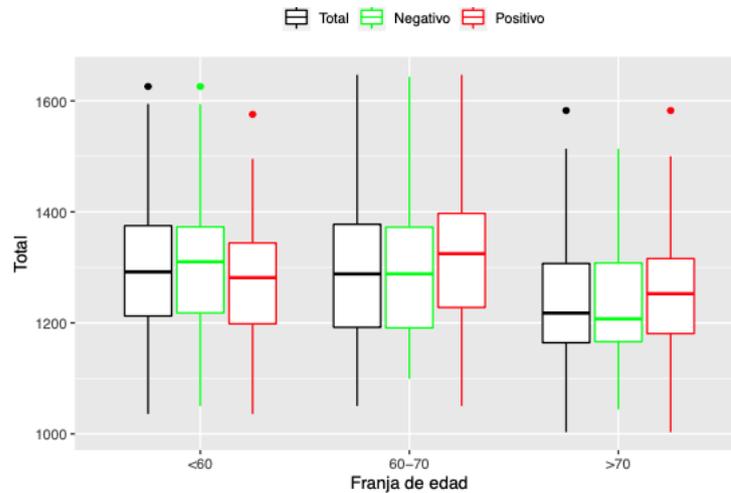


Figura 4.10: Total por franjas de edad

Siendo el boxplot de color negro la suma de los sujetos con resultado negativo, positivo y los que no se han realizado test, en verde los negativos y en rojo los positivos. Se puede observar, cómo con el paso de los años existe una reducción en el volumen cerebral y cómo hay una reducción notable en la franja de los sujetos mayores de 70 años. A su vez, se puede ver cómo en la figura del CSF (ver figura 4.9) el volumen aumenta conforme van pasando los años. El líquido cefalorraquídeo rellena aquellas zonas cuyo volumen se ha visto reducido, por eso se puede ver reflejado un aumento del volumen del CSF. Se pueden ver todas las figuras en el apéndice 2.

El análisis que se está siguiendo hasta ahora, no tiene en cuenta los hemisferios. Suele haber asimetrías entre hemisferios y suelen ser importantes. Al haber realizado el análisis de este modo, es posible que se estén enmascarando efectos interesantes, pero primero era importante tener una idea general antes de seguir profundizando en los datos que se disponen.

El procedimiento siguiente al realizado consiste en calcular los volúmenes de las 34 áreas pero esta vez separadas por materia (GM, WM y CSF). La idea es dibujar un scatterplot (gráfica de puntos) por cada una de las materias más una de la suma total de las materias, es decir, 136 scatterplots. El eje “x” de la gráfica representa la edad y el eje “y” representa el volumen. Los puntos tendrán un color distinto según el resultado del test (positivo: rojo, negativo: verde, no realizado: amarillo).

Estos scatterplots representan la evolución del volumen de cada área y materia con la edad y sirve para tener una primera aproximación visual. Se han añadido líneas que muestran la tendencia para cada grupo de sujetos.

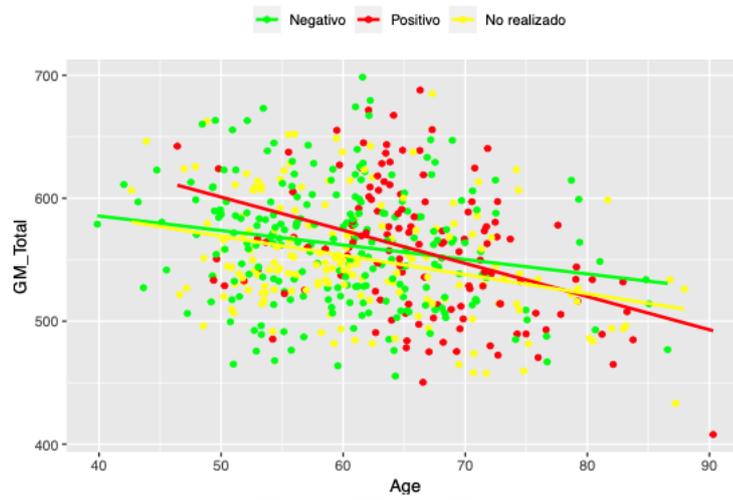


Figura 4.11: GM_VOL

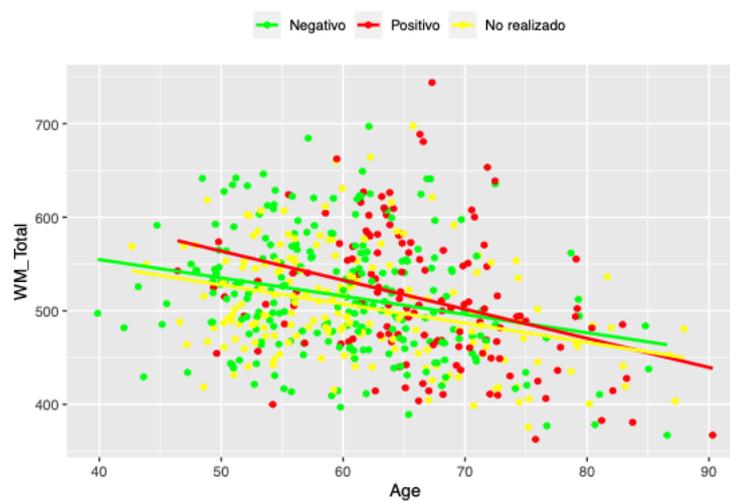
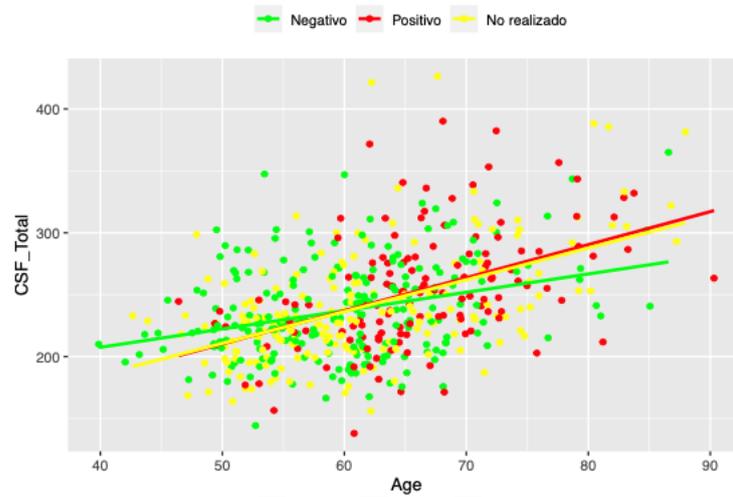
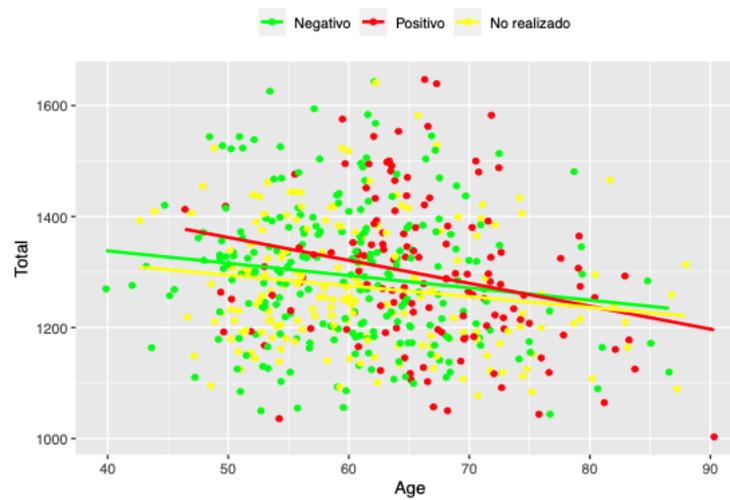


Figura 4.12: WM_VOL

**Figura 4.13:** CSF_VOL**Figura 4.14:** Total (GM, WM y CSF)

Por otro lado, además de las 4 variables que se han mostrado (GM, WM, CSF y total) se ha sumado los volúmenes de las áreas de los lóbulos del cerebro (temporal, occipital, frontal, parietal e interno). A cada área del cerebro, le corresponde un lóbulo. Teniendo esto, se realiza el mismo procedimiento: 4 gráficas por lóbulo (GM, WM, CSF y el total de las tres), que son 20 scatterplots más.

Además de los lóbulos, se ha realizado también las sumas por áreas, que sumadas a las gráficas de los lóbulos dan un total de 156 más (ver en el apéndice 3).

Estos casos concretos pertenecen a la materia gris (figura 4.11), blanca (figura 4.12), líquido cefalorraquídeo (figura 4.13) y total (figura 4.14). Se puede ver como el volumen de la materia gris y blanca se va reduciendo con el paso de los años mientras que el líquido cefalorraquídeo aumenta. Como se ha explicado antes, esto es debido a que el líquido rellena el volumen perdido en las materias gris y blanca.

4.3. Análisis para confirmar la hipótesis inicial de CITA

4.3.1. Ratios entre variables

Para seguir más en detalle en el análisis, es interesante analizar ratios entre distintas áreas porque están conectadas entre sí. Los médicos a día de hoy no saben cómo están conectadas y es interesante que este proyecto revele algunas de estas conexiones (correlaciones entre áreas, es decir, áreas que varían su volumen de manera opuesta). Los ratios que se han calculado son GM/GM, WM/WM, CSF/CSF y GM/WM.

Ahora bien, en las áreas en las que el volumen varía con la edad, lo normal es que GM y WM bajen y CSF suba. Esto hace que en áreas con comportamientos parecidos (varían con la edad de forma similar) el ratio entre GM y WM sea más o menos constante (como varían de forma similar, su ratio no cambia mucho). Sin embargo, el ratio entre GM (o WM) y CSF varía mucho porque uno sube y el otro baja. Por lo que estas relaciones no tienen mucho interés ya que corresponden a áreas que varían de forma similar.

Para solucionarlo, la idea es generar los ratios como se ha mencionado pero sin tener en cuenta las columnas correspondientes a ratios entre GM y CSF, ni WM y CSF. Por lo que únicamente se tratarían los ratios consigo mismos (GM/GM, WM/WM, CSF/CSF) y GM/WM, además de las áreas.

4.3.2. Normalización

También se deja a un lado los volúmenes absolutos y se utilizan volúmenes relativos. Es mejor usar volúmenes relativos que absolutos y dividirlo por materias, por lo que es de gran interés proceder a la normalización de los datos. Se divide ahora el volumen absoluto por TIV (se le va a llamar relativo total o RT) y se divide también el volumen absoluto dependiendo de la materia, es decir, GM_VOL, WM_VOL y CSF_VOL (se le va a llamar relativo materia o RM). Según los médicos, es un planteamiento que tiene lógica aunque no recuerdan haberlo visto así realizado nunca.

4.3.3. Suavizado

Como se ha mencionado en el capítulo previo, se ha decidido usar el algoritmo “Kernel Smoother”. Este método devuelve una línea de cualquier forma, es decir, no tiene que ser recta, ni polinomial, ni exponencial, que resume el scatterplot de forma más o menos suavizada (existe un parámetro que permite controlar el nivel de suavizado). Este parámetro es el “bandwidth” y el valor que se le ha establecido al parámetro es 10, que tras prueba y error se ha visto que es el valor 10 el más indicado para un buen nivel de suavizado. Mediante este método se van a suavizar las curvas. Se ha optado por usar un kernel Gaussiano.

4.3.4. Estandarizado

No solo eso, sino que también se propone realizar un estandarizado de los datos. Esto es importante, debido a que existen áreas con volúmenes de 5cm^3 y otras de 100cm^3 , como ya se ha mencionado. Con la estandarización, se consigue que los volúmenes de todas las áreas cerebrales oscilen entre rangos similares y, por tanto, puedan ser comparables.

4.3.5. Cálculo de pendientes

El siguiente paso va a consistir en calcular las pendientes de las curvas suavizadas que se han conseguido. Una vez obtenidas, las líneas suavizadas se dividen por tramos de un año sin ningún solapamiento entre tramos. Se calcula la diferencia en absoluto entre los extremos de cada tramo, es decir, para cada tramo de un año (salen 50 tramos, ya que las gráficas van de 40 a 90 años) se resta el extremo más bajo del extremo más alto. A este

valor se le va a denominar “pendiente” porque es una forma de medir la pendiente de ese tramo.

Llegados a este punto, se quiere categorizar las pendientes en llanas, moderadas y fuertes. Para ello, se definen dos umbrales: inferior y superior. A la hora de decidir sus valores, se ha usado el histograma de la figura 4.15. Los valores que se han escogido son 0.04 para el umbral inferior y 0.1 para el umbral superior.

Las pendientes llanas son todas aquellas que estén por debajo del umbral inferior, las pendientes moderadas las que estén entre el umbral inferior y superior, y las pendientes fuertes las que estén por encima del umbral superior.

Se ha contado cuántas pendientes hay de cada tipo y se han sumado sus valores.

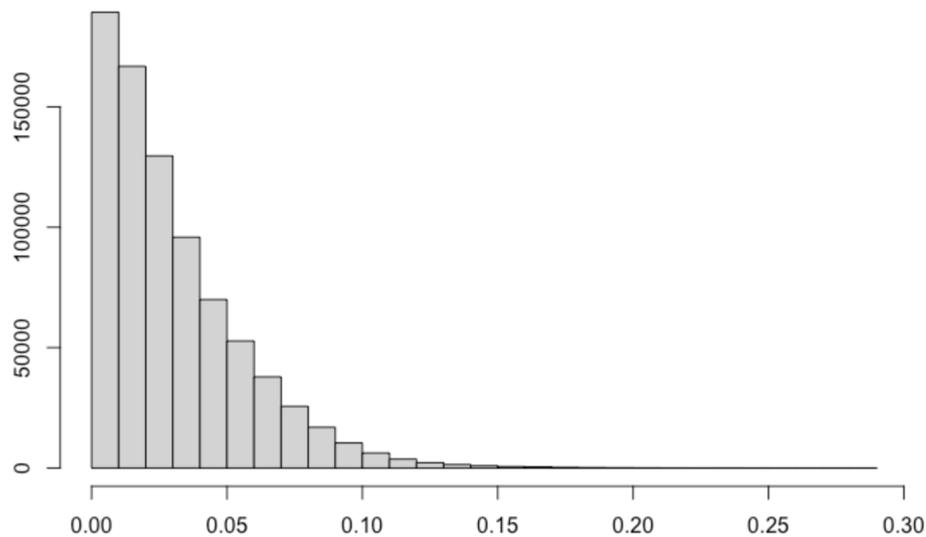


Figura 4.15: Histograma de pendientes por tramo

Habiendo definido las pendientes, se obtiene el siguiente resultado:

- Porcentaje de pendientes llanas: 72.78 %
- Porcentaje de pendientes moderadas: 25.24 %
- Porcentaje de pendientes fuertes: 1.96 %

Es interesante a continuación, teniendo los valores de las pendientes categorizadas en moderadas y fuertes (que son las pendientes en las que ponemos más atención) poder visualizarlo para realizar un análisis. Se ha decidido construir dos scatterplots: uno hace referencia a las pendientes moderadas y el restante a las fuertes. Estos scatterplots van a estar divididos por áreas y ratios.

De este modo, se puede identificar qué variables son las que más pendientes moderadas o fuertes tienen (ver figuras 4.16 y 4.17).

Es importante saber que, en el scatterplot, en el eje horizontal se ha desplazado cada grupo (puntos de colores) hacia la izquierda y derecha. Esto hace que la visualización y localización de los puntos sea más sencilla.

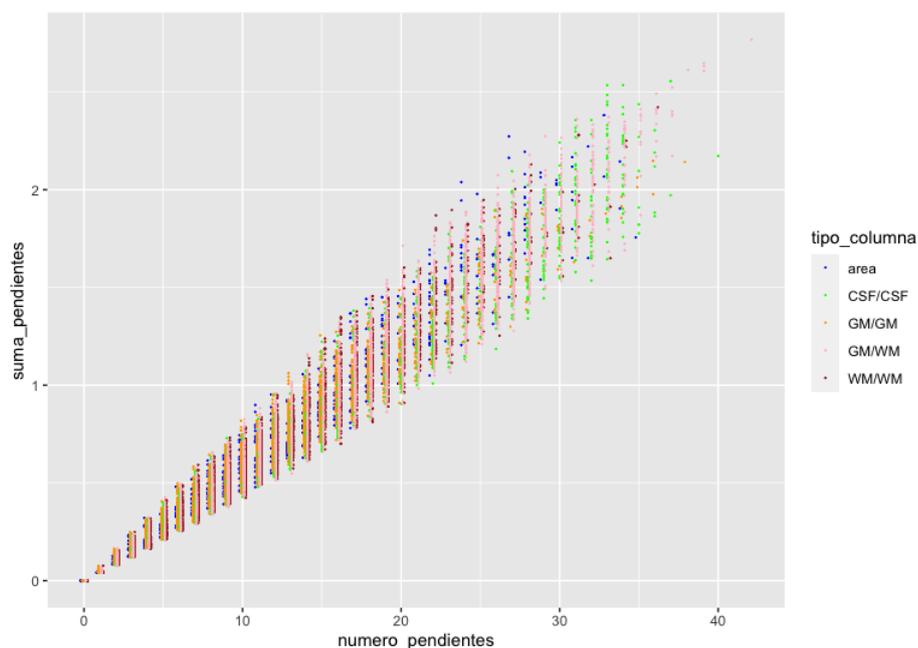


Figura 4.16: Pendientes moderadas por grupos

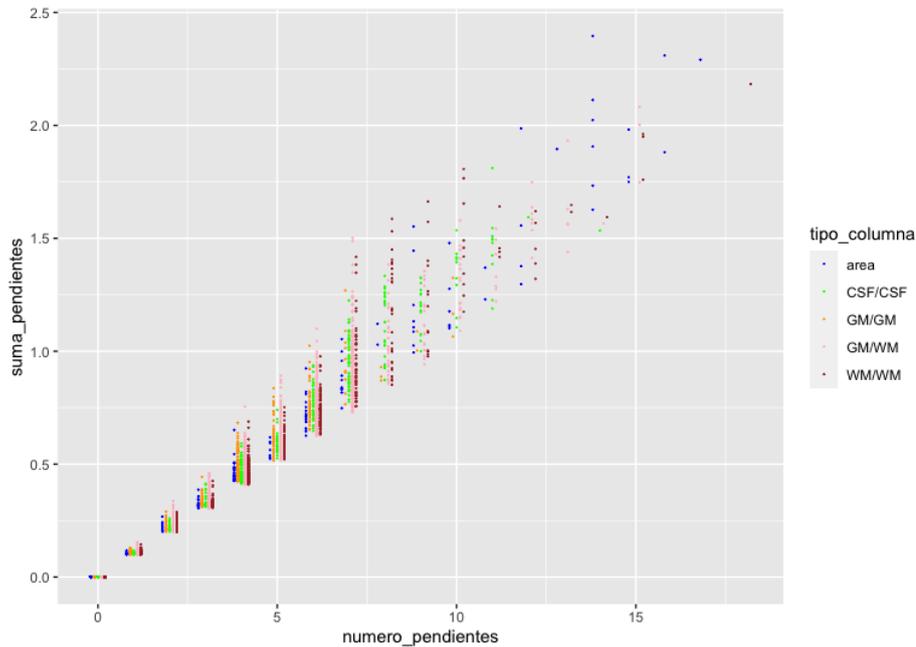


Figura 4.17: Pendientes fuertes por grupos

Los puntos más interesantes son los que tienen mayor pendiente, son los representados en la parte superior derecha de estas gráficas. Para escoger el conjunto más interesante se ha decidido utilizar los siguientes criterios en base a una inspección visual de las figuras (4.16 y 4.17):

- Para las moderadas: se realiza el análisis para aquellas variables cuyo número de pendientes moderadas es mayor que 33 y para aquellas 5 cuya suma de pendientes sean las mayores.
- Para las fuertes: se realiza el análisis para aquellas variables cuyo número de pendientes fuertes es mayor que 14 y para aquellas 5 cuya suma de pendientes sean las mayores.

4.3.6. Variables con mayor pendiente moderada y fuerte

Se muestra en las tablas 4.10 y 4.11 el nombre de la variable, número de pendientes y la suma de pendientes por variable tras haber realizado el análisis anterior.

Área o ratio entre áreas	# pendientes	Σ pendientes
Caudate_R_WM RM/Sup_Front_Gyr_R_WM RM	18	2,18
Hipp_R_CSF RT	17	2,29
Hipp_L_CSF RT	16	2,31
Caudate_R_WM RM	16	1,88
Hipp_L_GM RT/Hipp_L_WM RT	15	2,08
Caudate_L_WM RM/Accumb_L_GM RM	15	2,00
Post_Temp_R_CSF RT	15	1,98
Caudate_R_WM RM/Sup_Front_Gyr_L_WM RM	15	1,96
Insul_L_WM RM/Caudate_L_WM RM	15	1,95
Parahipp_R_CSF RT	14	2,40
Parahipp_L_CSF RT	14	2,11
Post_Temp_L_CSF RT	14	2,02
Lingual_Gyr_L_CSF RT	14	1,91
Sup_Temp_Post_R_CSF RT	14	1,73

Tabla 4.10: Variables con más pendientes fuertes

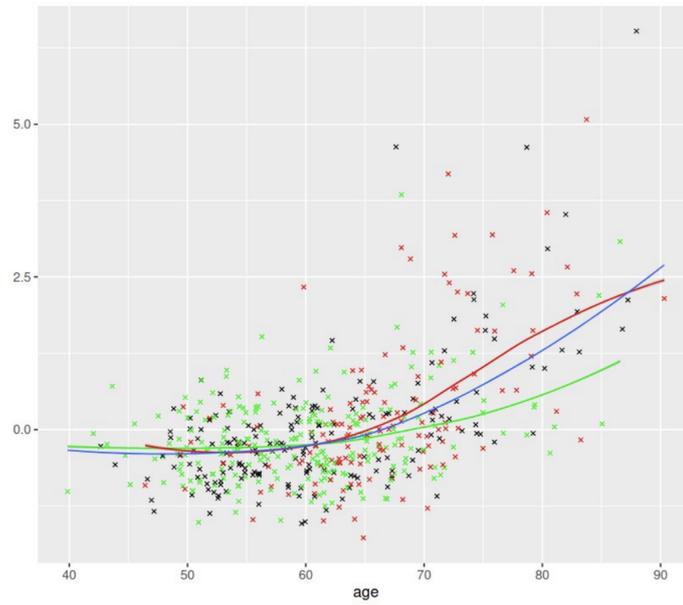
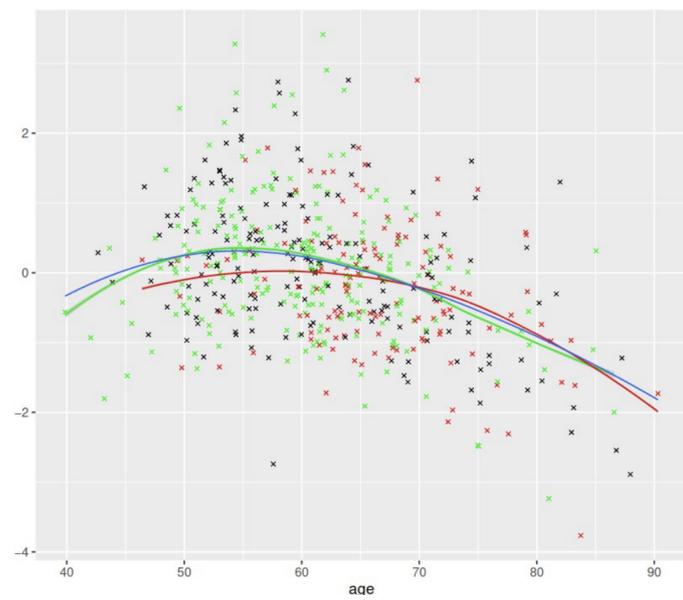
Una vez seleccionados los puntos más interesantes, se han analizado visualmente sus curvas. Este análisis ha permitido observar que entre las curvas hay ciertos patrones que se repiten. En concreto, se han detectado 4 patrones diferentes. Las figuras 4.18, 4.19, 4.20, y 4.21 muestran los ejemplos de estos patrones.

El patrón de la figura 4.18 muestra unas curvas que en edades tempranas tienen valores similares, pero que a partir de unos 65 años empiezan a diferenciarse, incrementando su valor de forma mucho más acusada la curva de sujetos positivos que la curva de sujetos sanos.

En la 4.19 las curvas empiezan con algo de diferencia, pero luego se juntan, en la 4.20 las curvas van juntas y al final (a partir de unos 80 años) se separan y en la 4.21 las curvas son muy similares en todo el recorrido.

Área o ratio entre áreas	# pendientes	Σ pendientes
Parahipp_L_WM RT/Sup_Front_Gyr_R_GM RT	42	2,76
Mid_Front_R_CSF RM/Thal_L_CSF RM	40	2,17
Parahipp_L_WM RT/Precuneus_R_GM RT	39	2,65
Parahipp_L_WM RT/Mid_Front_L_GM RT	39	2,63
Parahipp_L_WM RT/Precuneus_L_GM RT	39	2,61
Parahipp_L_WM RT/Sup_Front_Gyr_L_GM RT	38	2,61
Hipp_L_GM RM/Sup_Front_Gyr_L_GM RM	38	2,14
Sup_Temp_Post_R_CSF RM/Mid_Inf_Temp_R_CSF RM	37	2,56
Parahipp_L_WM RT/Insul_L_GM RT	37	2,52
Parahipp_L_WM RT/Caudate_L_GM RT	37	2,40
Ant_temp_lat_L_WM RT/Sup_Front_Gyr_L_GM RT	37	2,38
Parahipp_L_WM RT/Lat_Occip_L_GM RT	37	2,38
Parahipp_L_WM RT/Caudate_R_GM RT	36	2,49
Parahipp_L_WM RM/Brainstem_R_WM RM	36	2,42
Mid_Inf_Temp_L_WM RT/Sup_Front_Gyr_R_GM RT	36	2,41
Inf_Lat_Pariet_L_WM RT/Precuneus_R_GM RT	36	2,41
Parahipp_L_WM RT/Ant_Orbit_Gyr_R_GM RT	36	2,39
Mid_Inf_Temp_R_CSF RM/Post_Temp_R_CSF RM	35	2,44
Mid_Inf_Temp_L_WM RT/Sup_Front_Gyr_L_GM RT	35	2,41
Parahipp_L_WM RT/Insul_R_GM RT	35	2,38
Mid_Inf_Temp_R_WM RT/Sup_Front_Gyr_L_GM RT	35	2,37
Ant_temp_lat_R_WM RT/Sup_Front_Gyr_L_GM RT	35	2,36
Sup_Temp_Post_R_CSF RM/Ant_Orbit_Gyr_L_CSF RM	34	2,56
Sup_Temp_Post_R_CSF RM/Straight_Gyr_R_CSF RM	34	2,42
Thal_L_CSF RM/Ant_Orbit_Gyr_R_CSF RM	34	2,39
Parahipp_L_WM RT/Post_Centr_Gyr_L_GM RT	34	2,35
Inf_Lat_Pariet_L_WM RT/Precuneus_L_GM RT	34	2,33
Caudate_R_CSF RM/Ant_Orbit_Gyr_R_CSF RM	33	2,53
Caudate_R_CSF RM/Ant_Orbit_Gyr_L_CSF RM	33	2,48
Thal_R_CSF RM/Ant_Orbit_Gyr_R_CSF RM	33	2,45
Caudate_R_CSF RM/Straight_Gyr_R_CSF RM	33	2,44
Sup_Temp_Post_R_CSF RM/Ant_Orbit_Gyr_R_CSF RM	33	2,43

Tabla 4.11: Variables con más pendientes moderadas

**Figura 4.18: Patrón 1****Figura 4.19: Patrón 2**

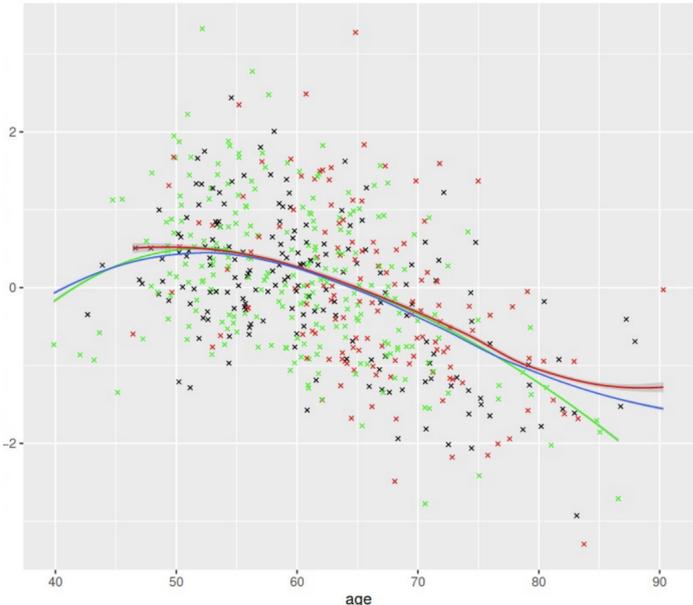


Figura 4.20: Patrón 3

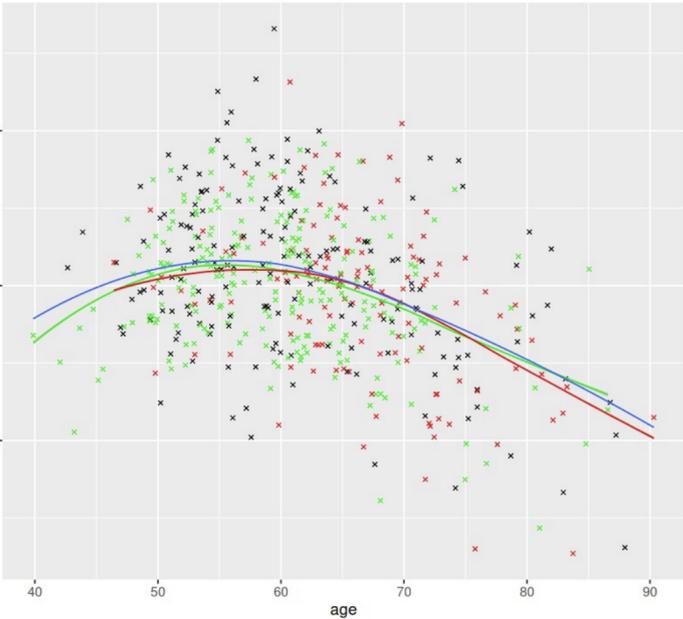


Figura 4.21: Patrón 4

En las tablas [4.12](#) y [4.13](#) se puede ver el tipo de patrón que sigue cada variable.

Pendientes moderadas

Área o ratio entre áreas	patrón	# pendientes	Σ pendientes
Parahipp_L_WM RT/Sup_Front_Gyr_R_GM RT	2	42	2,76
Mid_Front_R_CSF RM/Thal_L_CSF RM	4	40	2,17
Parahipp_L_WM RT/Precuneus_R_GM RT	3	39	2,65
Parahipp_L_WM RT/Mid_Front_L_GM RT	2	39	2,63
Parahipp_L_WM RT/Precuneus_L_GM RT	2	39	2,61
Parahipp_L_WM RT/Sup_Front_Gyr_L_GM RT	2	38	2,61
Hipp_L_GM RM/Sup_Front_Gyr_L_GM RM	2	38	2,14
Sup_Temp_Post_R_CSF RM/Mid_Inf_Temp_R_CSF RM	3	37	2,56
Parahipp_L_WM RT/Insul_L_GM RT	4	37	2,52
Parahipp_L_WM RT/Caudate_L_GM RT	4	37	2,40
Ant_temp_lat_L_WM RT/Sup_Front_Gyr_L_GM RT	1	37	2,38
Parahipp_L_WM RT/Lat_Occip_L_GM RT	2	37	2,38
Parahipp_L_WM RT/Caudate_R_GM RT	4	36	2,49
Parahipp_L_WM RM/Brainstem_R_WM RM	2	36	2,42
Mid_Inf_Temp_L_WM RT/Sup_Front_Gyr_R_GM RT	3	36	2,41
Inf_Lat_Pariet_L_WM RT/Precuneus_R_GM RT	2	36	2,41
Parahipp_L_WM RT/Ant_Orbit_Gyr_R_GM RT	2	36	2,39
Mid_Inf_Temp_R_CSF RM/Post_Temp_R_CSF RM	4	35	2,44
Mid_Inf_Temp_L_WM RT/Sup_Front_Gyr_L_GM RT	3	35	2,41
Parahipp_L_WM RT/Insul_R_GM RT	4	35	2,38
Mid_Inf_Temp_R_WM RT/Sup_Front_Gyr_L_GM RT	3	35	2,37
Ant_temp_lat_R_WM RT/Sup_Front_Gyr_L_GM RT	3	35	2,36
Sup_Temp_Post_R_CSF RM/Ant_Orbit_Gyr_L_CSF RM	3	34	2,56
Sup_Temp_Post_R_CSF RM/Straight_Gyr_R_CSF RM	3	34	2,42
Thal_L_CSF RM/Ant_Orbit_Gyr_R_CSF RM	3	34	2,39
Parahipp_L_WM RT/Post_Centr_Gyr_L_GM RT	2	34	2,35
Inf_Lat_Pariet_L_WM RT/Precuneus_L_GM RT	3	34	2,33
Caudate_R_CSF RM/Ant_Orbit_Gyr_R_CSF RM	1	33	2,53
Caudate_R_CSF RM/Ant_Orbit_Gyr_L_CSF RM	1	33	2,48
Thal_R_CSF RM/Ant_Orbit_Gyr_R_CSF RM	3	33	2,45
Caudate_R_CSF RM/Straight_Gyr_R_CSF RM	1	33	2,44
Sup_Temp_Post_R_CSF RM/Ant_Orbit_Gyr_R_CSF RM	3	33	2,43

Tabla 4.12: Patrón que siguen las variables con pendientes moderadas

Pendientes fuertes

Área o ratio entre áreas	patrón	# pendientes	Σ pendientes
Caudate_R_WM RM/Sup_Front_Gyr_R_WM RM	1	18	2,18
Hipp_R_CSF RT	1	17	2,29
Hipp_L_CSF RT	1	16	2,31
Caudate_R_WM RM	1	16	1,88
Hipp_L_GM RT/Hipp_L_WM RT	4	15	2,08
Caudate_L_WM RM/Accumb_L_GM RM	1	15	2,00
Post_Temp_R_CSF RT	1	15	1,98
Caudate_R_WM RM/Sup_Front_Gyr_L_WM RM	1	15	1,96
Insul_L_WM RM/Caudate_L_WM RM	1	15	1,95
Parahipp_R_CSF RT	1	14	2,40
Parahipp_L_CSF RT	1	14	2,11
Post_Temp_L_CSF RT	1	14	2,02
Lingual_Gyr_L_CSF RT	3	14	1,91
Sup_Temp_Post_R_CSF RT	1	14	1,73

Tabla 4.13: Patrón que siguen las variables con pendientes fuertes

El patrón 1, que es el que soporta la hipótesis de CITA se repite en varias ocasiones en las variables con pendientes fuertes.

4.4. Análisis comparativo entre el total de sujetos y sujetos en base al test

Los neurólogos de la fundación CITA-Alzheimer indicaron que sería interesante analizar cómo pueden variar los resultados si se seleccionan las mismas curvas pero esta vez en base a la curva roja (sujetos positivos) y en base a la curva verde (sujetos negativos).

Este proceso va a consistir en analizar aquellas variables que tengan pendientes fuertes, ya que parecen ser más interesantes como se ha visto anteriormente (muchas siguen el mismo patrón). Al igual que se ha hecho con la curva azul, se seleccionan las variables que tengan 33 o más pendientes moderadas y las variables que tengan 14 o más pendientes fuertes.

En primer lugar se realiza el proceso que se ha mencionado para los sujetos positivos. Se muestra la lista con las curvas seleccionadas (ver tabla 4.14).

Áreas o ratio entre áreas	# pendientes	Σ pendientes
Caudate_L_WM_VOL.35_RM/Ant_Orbit_Gyr_R_WM_VOL.56_RM	20	2,86
Caudate_R_WM_VOL.36_RM/Ant_Orbit_Gyr_R_WM_VOL.56_RM	19	3,07
Insul_L_WM_VOL.21_RM/Caudate_L_WM_VOL.35_RM	19	2,10
Pallidum_L_GM_VOL.43_RT/Ant_Orbit_Gyr_R_WM_VOL.56_RT	18	3,38
Pallidum_L_GM_VOL.43_RT/Precuneus_L_WM_VOL.63_RT	18	3,04
C_Callosum_R_GM_VOL.46_RT/Ant_Orbit_Gyr_R_WM_VOL.56_RT	18	2,58
Pallidum_L_GM_VOL.43_RT/Precuneus_R_WM_VOL.64_RT	17	3,11
Pallidum_R_GM_VOL.44_RT/Precuneus_R_WM_VOL.64_RT	17	3,06
Pallidum_L_GM_VOL.43_RT/Lat_Temp_Ventr_R_WM_VOL.48_RT	17	2,76
Pallidum_L_GM_VOL.43_RT/Cuneus_R_WM_VOL.68_RT	17	2,52
GM_VOL_hipp_R_RM/Pre_Centr_Gyr_L_WM_VOL.51_RM	17	2,17
Pallidum_R_GM_VOL.44_RT/Ant_Orbit_Gyr_R_WM_VOL.56_RT	16	3,08
Pallidum_L_GM_VOL.43_RM/Ant_Orbit_Gyr_R_WM_VOL.56_RM	16	3,05
Pallidum_L_GM_VOL.43_RM/Cuneus_R_GM_VOL.68_RM	16	2,88
Pallidum_L_GM_VOL.43_RM/Cuneus_L_GM_VOL.67_RM	16	2,86
Pallidum_R_GM_VOL.44_RT/Precuneus_L_WM_VOL.63_RT	16	2,78
Pallidum_L_GM_VOL.43_RT/Third_Ventr_R_WM_VOL.50_RT	15	3,23
Pallidum_L_GM_VOL.43_RM/Precuneus_R_WM_VOL.64_RM	15	2,85
Pallidum_R_GM_VOL.44_RM/Ant_Orbit_Gyr_R_WM_VOL.56_RM	15	2,84
Pallidum_L_GM_VOL.43_RM/Precuneus_L_WM_VOL.63_RM	15	2,65
Ant_Orbit_Gyr_R_GM_VOL.56_RM/Ant_Orbit_Gyr_R_WM_VOL.56_RM	15	2,63
Pallidum_R_GM_VOL.44_RT/Third_Ventr_R_WM_VOL.50_RT	14	2,89
Pallidum_L_GM_VOL.43_RM/Third_Ventr_R_WM_VOL.50_RM	14	2,88
Pallidum_R_GM_VOL.44_RM/Precuneus_R_WM_VOL.64_RM	14	2,70
Pallidum_R_GM_VOL.44_RM/Cuneus_L_GM_VOL.67_RM	14	2,67
Pre_Centr_Gyr_L_WM_VOL.51_RM/Ant_Orbit_Gyr_R_WM_VOL.56_RM	14	2,60

Tabla 4.14: Variables con más pendientes fuertes con la curva roja

A continuación, se realiza el proceso para los sujetos negativos. Se muestra la lista con las curvas seleccionadas (ver tabla 4.15).

Áreas o ratio entre áreas	# pendientes	∑ pendientes
Ant_temp_med_R_WM_VOL.6_RM/Mid_Front_R_WM_VOL.30_RM	16	2,63
Ant_temp_med_L_WM_VOL.5_RM/Post_Temp_R_WM_VOL.32_RM	16	2,41
Ant_temp_med_L_WM_VOL.5_RM/Post_Temp_R_GM_VOL.32_RM	16	2,24
Ant_temp_med_L_WM_VOL.5_RM/Mid_Front_R_WM_VOL.30_RM	15	2,59
Ant_temp_med_R_WM_VOL.6_RM/Post_Temp_R_WM_VOL.32_RM	15	2,29
Ant_temp_med_R_WM_VOL.6_RM/Pre_Centr_Gyr_R_WM_VOL.52_RM	15	2,28
Ant_temp_med_R_WM_VOL.6_RM/Pre_Centr_Gyr_L_WM_VOL.51_RM	15	2,22
Ant_temp_med_R_WM_VOL.6_RM/Post_Temp_R_GM_VOL.32_RM	15	2,16
Mid_Front_R_GM_VOL.30_RT/Post_Temp_L_WM_VOL.31_RT	14	2,54
Ant_temp_med_R_WM_VOL.6_RM/Inf_Lat_Pariet_L_WM_VOL.33_RM	14	2,21
Straight_Gyr_R_GM_VOL.54_RT/Precuneus_R_WM_VOL.64_RT	14	2,19
Ant_temp_med_L_WM_VOL.5_RM/Inf_Lat_Pariet_L_GM_VOL.33_RM	14	2,13
Ant_temp_med_L_WM_VOL.5_RM/Pre_Centr_Gyr_L_WM_VOL.51_RM	14	2,09

Tabla 4.15: Variables con más pendientes fuertes con la curva verde

Teniendo estas variables y haciendo una comparativa con las que se han obtenido en base a la curva azul (todos los sujetos), casi no hay variables que coincidan con esta última. Las variables que se repiten en la curva azul y en la curva roja son:

- “Sup_Temp_Post_R_CSF_VOL.12_RM/Ant_Orbit_Gyr_R_CSF_VOL.56_RM”
- “Sup_Temp_Post_R_CSF_VOL.12_RM/Ant_Orbit_Gyr_L_CSF_VOL.55_RM”
- “Insul_L_WM_VOL.21_RM/Caudate_L_WM_VOL.35_RM”

Y las que se repiten en la curva azul y en la curva verde es:

- “Mid_Inf_Temp_L_WM_VOL.13_RT/Sup_Front_Gyr_R_GM_VOL.60_RT”

Para ver las variables de manera más visual, se construye el mismo scatterplot para la curva roja y verde al igual que se hizo para la curva azul (ver figuras 4.22 y 4.23).

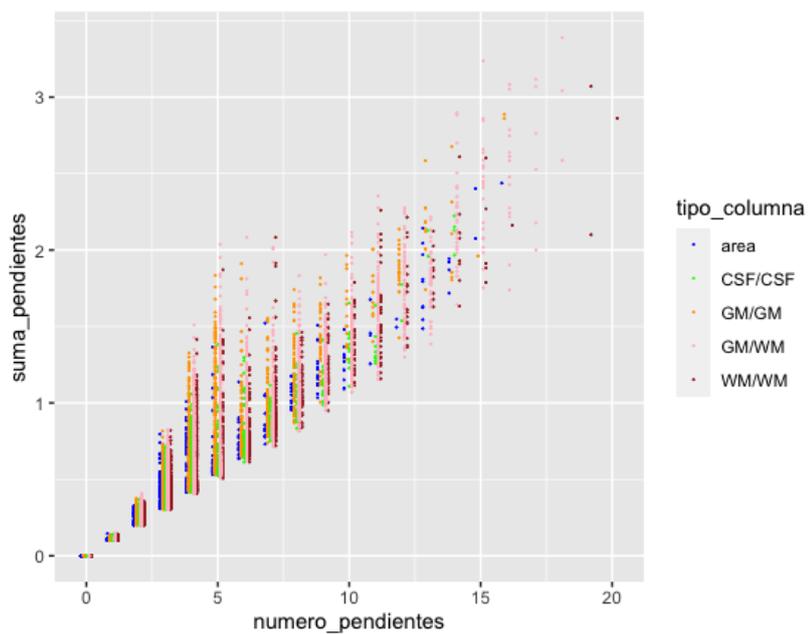


Figura 4.22: Pendientes fuertes curva roja

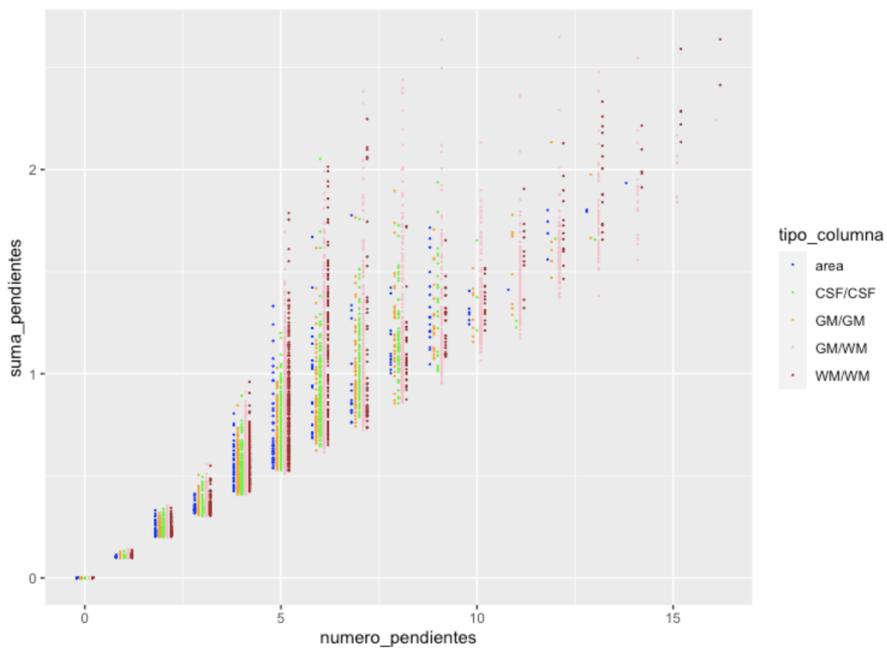


Figura 4.23: Pendientes fuertes curva verde

Por lo que se puede observar, las gráficas parecen diferir notablemente en comparación con el mismo scatterplot para la curva azul (ver figura 4.17).

Estos resultados, que muestran que hay diferencias considerables entre los casos seleccionados según un criterio u otro, hacen intuir que el número de pendientes fuertes de un área o ratio varía notablemente entre la curva roja y la verde. Para analizar visualmente este efecto hemos dibujado los siguientes gráficos (ver figuras 4.24 y 4.25). Consisten en dibujar los puntos de la curva azul y los de la curva roja y verde en un mismo scatterplot, dibujando así una línea que indica dónde está el punto (variable) en la curva azul y dónde está en las curvas roja y verde. Se construyen 10 gráficas: 5 por grupo (áreas, ratios GM/GM, WM/WM, CSF/CSF y GM/WM) y por test (roja y verde). Los puntos o variables que se muestran pertenecen a variables con pendientes fuertes.

Únicamente se incluyen las figuras que hacen referencia al grupo de las áreas para que se visualice un ejemplo de los scatterplots mencionados. El resto de figuras se incluyen en el apéndice.

Estas gráficas son similares entre sí. Cabe destacar que las líneas cortas indican cambios leves y líneas largas indican cambios grandes. Se han quitado las líneas que van al punto (0,0) debido a que hay excesivas líneas que van al punto de origen y no ayuda a la visualización. También se han quitado las líneas que van a los puntos con una única pendiente fuerte, para que así el resto se vea de manera más clara (ver en el apéndice 4 y 5).

Se confirma que el número (y suma) de las pendientes fuertes varía mucho dependiendo de si se seleccionan todos los sujetos o solo aquellos que tienen un resultado positivo (o negativo) en la punción. La gran cantidad de líneas y su longitud así lo indican.

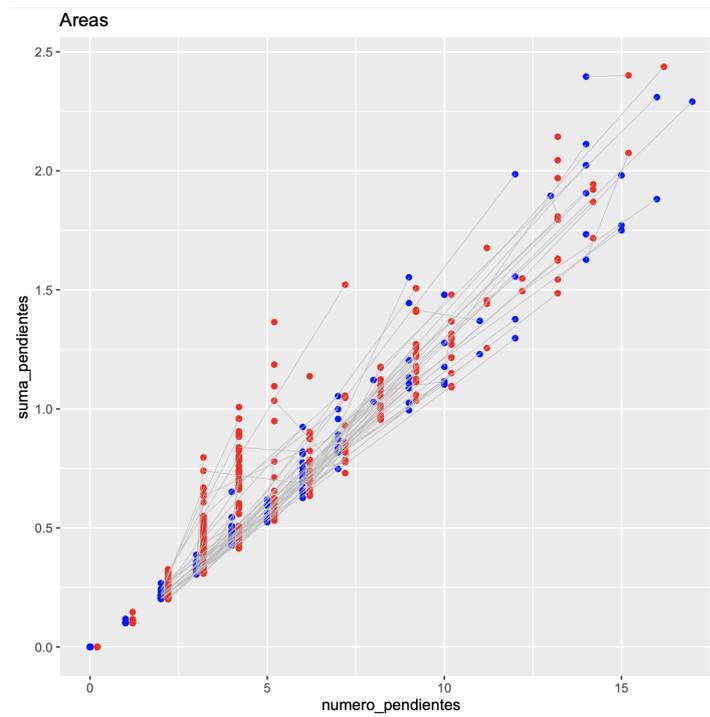


Figura 4.24: Diferencias entre puntos en base a la curva roja

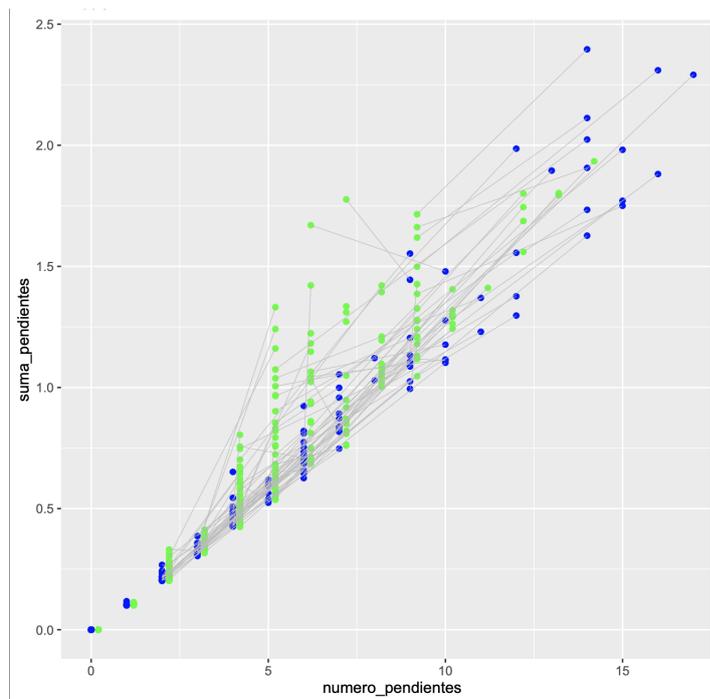


Figura 4.25: Diferencias entre puntos en base a la curva verde

5. CAPÍTULO

Conclusiones y trabajo futuro

En este trabajo, se ha analizado la evolución de las áreas cerebrales de 537 sujetos, tanto sanos como en estados iniciales de la enfermedad de Alzheimer.

En el análisis realizado, teniendo en cuenta todos los datos (sujetos con resultado negativo, positivo y no realizado) y por otro lado los negativos, únicamente se encuentra en común una única variable que hace referencia a un ratio entre 2 áreas cerebrales.

Sin embargo, en el análisis teniendo en cuenta todos los datos y también los positivos se han encontrado 3 coincidencias en variables que también son ratios.

Se puede mencionar que si se seleccionan las áreas (o ratios entre áreas) que más varían con la edad, la variación entre los sujetos sanos y los sujetos en estado preclínico del Alzheimer no es igual en muchos casos.

Las áreas o ratios con mayor variación con la edad, difieren mucho si se seleccionan en base a todos los sujetos o solo en base a sujetos negativos o positivos.

Las zonas que más interesan para que en un futuro se sigan analizando, son las curvas que se han definido como pendientes fuertes porque indican cambios bruscos en las curvas, por lo que algo podría estar sucediendo en esas áreas o ratios.

Se podría analizar las curvas para detectar las áreas donde la evolución de las curvas rojas y verdes difiera más. Así, se detectarían las áreas más afectadas en los sujetos en estado preclínico del Alzheimer.

También sería interesante en un futuro analizar los puntos pares cercanos (ver figuras [4.24](#)

y 4.25), ya que estos son cambios leves entre curvas seleccionadas (ya sea azul con roja o azul con verde).

Como un posible trabajo futuro, además de completar el análisis de este proyecto sería interesante usar técnicas de Machine Learning para que con la información obtenida durante este trabajo, se pueda ser capaz de hacer una estimación a los pacientes preclínicos para determinar si son susceptibles de padecer la enfermedad de Alzheimer sin necesidad de tener que realizarse el test de punción lumbar.

A día de hoy existen medicamentos que son capaces de retrasar los síntomas, una vez que el individuo ya padece la enfermedad. Pero, si en un futuro se da con el/las área/s exactas que son influyentes para el desarrollo de dicha enfermedad, sería un gran avance para la medicina y para la ciencia en general, ya que todavía no se sabe cómo funciona de manera exacta el cerebro, pero se averiguaría las zonas cerebrales que se deberían analizar y “atacar”, para que de alguna forma (quién sabe) se pudiera curar (o evitar) la enfermedad.

Anexos

Apéndice

Los siguientes puntos que se muestran a continuación indican los documentos en formato PDF que se deben consultar en caso de querer acceder a la información que se referencia a lo largo del capítulo 4. Incluyen todas aquellas gráficas que se han obtenido durante el proyecto y no se han podido mostrar, debido a que de haber incluido todas estas gráficas, la memoria hubiera sido muy extensa.

1. **Boxplots**
2. **Boxplots por franjas de edad**
3. **Scatterplots con líneas.**
4. **Azules con rojos**
5. **Azules con verdes**

Bibliografía

- [1] Deterioro cognitivo leve. <https://medlineplus.gov/spanish/mildcognitiveimpairment.html>.
- [2] Punción espinal (punción lumbar). <https://www.mayoclinic.org/es-es/tests-procedures/lumbar-puncture/about/pac-20394631>.
- [3] Introducción a los trastornos. <http://www.alzfae.org/fundacion/504/introduccion-a-los-trastornos>, 2011.
- [4] Enfermedad de alzheimer. <https://medlineplus.gov/spanish/alzheimersdisease.html>, 2019.
- [5] Chernecky CC and Berger BJ. Magnetic resonance imaging (mri) - diagnostic. in: Chernecky cc, berger bj, eds. laboratory tests and diagnostic procedures. 6th ed. pages 754–757, 2013.
- [6] Barras CD and Bhattacharya JJ. Current status of imaging of the brain and anatomical features. in: Adam a, dixon ak, gillard jh, schaefer-prokop cm, eds. grainger allison's diagnostic radiology. 2021.
- [7] M Corbetta and G. L. Shulman. Control of goal-directed and stimulus-driven attention in the brain. nat rev neurosci, 3. pages 201–215, 2002.
- [8] María Sagrario Manzano Palomo. Qué es el envejecimiento cerebral. <https://knowalzheimer.com/que-es-el-envejecimiento-cerebral/>, 2014.