



**Universitat Oberta
de Catalunya**

Interacción de síntomas neuropsiquiátricos,
Poligenic risk scores y *APOE* en la conversión del
deterioro cognitivo leve a demencia

Emilio Alarcón Martín

Bioinformática y bioestadística

Área 4 - Subárea 1: Estudios genéticos de enfermedades humanas

Dr Helena Brunel Montaner

Antonio Pérez Navarro

05/01/2020



Esta obra está sujeta a una licencia de
Reconocimiento-NoComercial-SinObraDerivada
[3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Interacción de síntomas neuropsiquiátricos, Poligenic risk scores y APOE en la conversión del deterioro cognitivo leve a demencia</i>
Nombre del autor:	<i>Emilio Alarcón Martín</i>
Nombre del consultor/a:	<i>Dr Helena Brunel Montaner</i>
Nombre del PRA:	<i>Antonio Pérez Navarro</i>
Fecha de entrega (mm/aaaa):	01/2020
Titulación:	<i>Bioinformática y bioestadística</i>
Área del Trabajo Final:	<i>Área 4 - Subárea 1: Estudios genéticos de enfermedades humanas</i>
Idioma del trabajo:	<i>Español</i>
Palabras clave	<i>Alzheimer disease, major depressive disorder, Polygenic Risk Score</i>
<p>Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.</i></p>	
<p>El descubrimiento de biomarcadores genéticos en forma de PRS supondría un avance en la evaluación del riesgo de la enfermedad de Alzheimer.</p> <p>Se muestran datos propios del proyecto GR@CE con 755 pacientes con deterioro cognitivo leve en los que se evalúa la conversión a demencia mediante datos de GWAS en forma de PRS (tanto de Alzheimer como de depresión mayor), incluyendo variables demográficas y síntomas neuropsiquiátricos.</p> <p>Mediante un modelo de Cox para evaluar el tiempo de supervivencia, se obtuvo un HR para ambas PRS_{AD} y PRS_{MDD}. de 1.3 y 1.2 respectivamente cuando empleamos los síntomas neuropsiquiátricos clave en la visita basal. La edad (HR = 1.05) y la apatía (HR = 1.6) son dos variables clave en la conversión.</p> <p>Con respecto a la predicción del síntoma neuropsiquiátrico depresivo, a través de un modelo de regresión logística, detectamos que el sexo está estrechamente relacionado (OR= 2.9), junto con los otros síntomas analizados (Apatía OR = 3.4 // Ansiedad OR = 6,5). Se percibe cierta tendencia con respecto a los PRS, pero la cohorte no tiene suficiente poder estadístico para discriminar el efecto de estas variables en comparación con los síntomas.</p>	

Abstract (in English, 250 words or less):

New findings of genetic biomarkers as PRS would represent a new approachment to Alzheimer's disease risk evaluation. Data used in this study has been collected from GR@CE Project. 755 patients with low cognitive impairment are shown, in which conversion to dementia is evaluated by GWAS data in PRS form (Alzheimer's disease as well as major depression), including demographic variables and neuropsychiatric symptoms. Using a Cox model to assess survival time, an HR was obtained both PRSAD and PRSMDD; 1.3 and 1.2 respectively when neuropsychiatric symptoms are key at the baseline visit. Age (HR= 1.05) and apathy (HR = 1.6) are used as key variables in conversion as well. Regarding the prediction of the depressive neuropsychiatric symptom, through a logistic regression model, we detected that sex is closely related (OR = 2.9), together with the other symptoms analyzed (Apathy OR = 3.4 // Anxiety OR = 6, 5). Certain trend has been found regarding to PRS, but the cohort does not have enough statistical power to discriminate the effect of those variables compared to the Discovery of genetic biomarkers in PRS form that would represent and advance in the Alzheimer's disease risk evaluation.

Índice

1. Introducción.....	1
1.1 Contexto y justificación del Trabajo	1
1.2 Objetivos del Trabajo	5
1.3 Enfoque y método seguido.....	5
1.4 Planificación del Trabajo. Tareas y calendario:.....	6
1.5 Breve resumen de productos obtenidos	7
1.6 Breve descripción de los otros capítulos de la memoria	7
2. Selección de pacientes y variables para el estudio	8
2.1. Variables trabajadas por la Unidad de Memoria.....	9
2.2. Variables genéticas.....	13
3. Validación del síntoma depresivo.....	16
3.1 Planteamiento teórico. Regresión logística, OR o RR	16
3.2 Modelos.....	17
4. Estudios de conversión MCI - Demencia.....	19
4.1 Planteamiento teórico.....	19
4.2. Modelos.....	21
5. Conclusiones.....	23
5.1. Ideas finales.....	23
5.2. Reflexión y análisis del seguimiento de planificación y metodología	23

5.3. Líneas de trabajo futuras	24
6. Glosario	25
7. Bibliografía.....	28
8. Anexos.....	33

Lista de figuras

Figura 1: Diagrama de Gantt para exponer el tiempo previsto para los distintos hitos del TFM

Figura 2: *Workflow* del número de pacientes según los criterios generales y específicos.

Figura 3: Gráficos para evaluar si la variable PRS_MDD sigue una distribución normal.

Figura 4: Gráfico de puntos comparando la distribución del PRS_{MDD} y PRS_{AD}.

Figura 5: Proporción de riesgos aplicada en el modelo de Cox para calcular HR

Figura 6: HR del modelo de Cox utilizando los síntomas neuropsiquiátricos basales

Figura 7: HR del modelo de Cox utilizando los síntomas neuropsiquiátricos totales

Lista de tablas

Tabla 1: Descriptivo general de variables demográficas

Tabla 2: Descriptivo general de variables clínicas

Tabla 3: Descriptivo general de variables genéticas

Tabla 4: Modelos de regresión logística para la predicción del síntoma depresivo.

Tabla 5: Modelos de Cox planteados para la conversión.

Agradecimientos

En estos tiempos de Covid, más que nunca, es importante dar las gracias a todos aquellos que, de una forma u otra, hacen posible que las cosas sigan adelante y que consiguen que las penas sean menos tristes y las alegrías más felices.

A mi tutora del máster Helena Brunel. Por aceptar este proyecto y tener la comprensión para entender mi situación particular. Sé que estas navidades no son las más comunes, pero que dediques parte de las mismas a revisar este proyecto es algo que agradezco inmensamente.

Como no podía ser de otra forma, dar las gracias a Fundació ACE, que con cercanía y calidez me han hecho sentir parte del equipo desde el minuto cero. A los compañeros del MS57, Sonia, Itziar y Pablo, por su predisposición a ayudar siempre que ha sido necesario y por hacerme ver que esto es un paso más de muchos. Al jefe de investigación, Agustín, que con sabio criterio siempre sabe ver aquello que solo la experiencia es capaz de detectar y ha sido capaz de aportar sentido común, que no siempre es el más común de los sentidos. En particular, a mi tutor y jefe de Fundació ACE, Sergi Valero. Con una disponibilidad absoluta y constante, siempre me ha hecho sentir respaldado, asesorado y acompañado en esta primera aventura combinando lo laboral y lo académico. Por valorar mi trabajo, tratarme como un igual y hacerme entender que no tenemos que venir sabiendo de todo, que nunca es mal momento para aprender algo nuevo y que cada paso que damos lo decidimos nosotros mismos.

A mis padres y mi hermana, que con todo su amor y cariño han estado ahí en todo momento, incluso cuando mi estado de humor hacía más agradable estar en la calle con mascarilla que dentro de casa. Que con paciencia y más paciencia, han escuchado cada uno de los soliloquios que he tenido durante este proyecto y que compartía con ellos para no sentir que estaba hablando más tiempo solo que acompañado.

A mis amigos y personas cercanas, vosotros que habéis sabido hacer de consejeros, de bufones y hasta de cuidadores... Entender que estar solo no implica sentirse solo es el primer paso para disfrutar de la vida en todas sus facetas. Vosotros hacéis que cada día sea más parecido a un viernes y menos a un lunes.

Por último, quiero dar las gracias a la única persona de los citados que sé que no podrá leer este trabajo. Abuela, este 2020 ha sido un año muy duro, pero más duro es saber que no estás aquí para festejar este logro de tu nieto. Ojalá que sigas siendo mi musa todos los días de mi vida como cuando estabas entre nosotros. Te echo de menos como nunca, te quiero como siempre.

1. Introducción

1.1 Contexto y justificación del Trabajo

1.1.2 Descripción general:

Un trastorno mental se define como la alteración de tipo emocional, cognitivo y/o conductual, en que quedan afectados procesos psicológicos básicos como son la emoción, la motivación, la cognición, el aprendizaje, el lenguaje, etc. Lo que dificulta a la persona su adaptación al entorno cultural y social en que vive y crea alguna forma de malestar subjetivo [1]. Algunos de los principales sistemas de clasificación de los trastornos mentales son:

- Capítulo V de la Clasificación Internacional de Enfermedades (CIE-10) creado por la Organización Mundial de la Salud (OMS) [2]
- Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-5) hechos por la Asociación Psiquiátrica Americana (APA) [3]

Siguiendo estas directrices, la demencia es un trastorno neurológico caracterizado por el deterioro de la función cognitiva más allá de lo que podría considerarse una consecuencia del envejecimiento normal. Afecta a la memoria, el pensamiento, la orientación, la comprensión, el cálculo, la capacidad de aprendizaje, el lenguaje y el juicio [4]. Es importante resaltar la diferencia entre el deterioro de la memoria asociado con la edad y el deterioro cognitivo leve (en inglés, MCI). En el MCI, el sujeto presenta una reducción de sus funciones cognitivas comparado con otras personas de su misma edad y además afecta a su desenvolvimiento diario. Hasta el 50% de los pacientes con MCI desarrolla demencia dentro de los 3 años [5]. Dado que se considera un diagnóstico intermedio entre la cognición normal y la demencia, se han dedicado muchos esfuerzos a la identificación de los individuos con MCI con una mayor vulnerabilidad de conversión a la demencia.

Es necesario especificar que Alzheimer y demencia no son sinónimos. La enfermedad de Alzheimer (en inglés, AD) es un tipo de demencia que aparece principalmente

durante la vejez. Es la más habitual, pues comprende entre un 60% y un 70% del total de casos de demencia [4]. Este trastorno neurodegenerativo se manifiesta con deterioro cognitivo y trastorno conductual que lleva a la pérdida de autonomía.

Distinguir entre los diversos trastornos mentales a menudo resulta un tema complejo, pues muchos presentan síntomas similares y un paciente puede presentar varios a la vez. En la actualidad, gracias a los avances en técnicas de genotipado, se plantea usar marcadores genéticos para discernir entre aquellas que puedan resultar similares, como el trastorno bipolar y la esquizofrenia [6].

Además de tener una cierta correlación genética, unas pueden ser comórbidas de otras. Este efecto está especialmente estudiado entre Alzheimer y depresión. Los pacientes con Alzheimer frecuentemente manifiestan síntomas neuropsiquiátricos como la depresión y la ansiedad, y las personas con trastorno depresivo mayor tienen un riesgo más elevado de desarrollar Alzheimer. Esto sugiere etiologías compartidas y vías de intersección entre ambas enfermedades [7].

El trastorno depresivo mayor (en inglés MDD, habitualmente denominado depresión) es una enfermedad mental que se caracteriza por un estado de ánimo invasivo y persistente acompañado de una baja autoestima y una pérdida de interés o de placer en actividades que normalmente se considerarían entretenidas. Es una enfermedad frecuente en todo el mundo, y se calcula que afecta a más de 300 millones de personas [8].

Además, coexisten síntomas depresivos durante los trastornos de ansiedad por dos vías: un 80% de pacientes con sintomatología depresiva no llegan a desarrollar MDD, viene producido por estrés crónico, un 20% comienzan manifestando aparentemente un trastorno de ansiedad, pero subyace una depresión. Este último caso conlleva mayor riesgo, debido a que la depresión puede agravarse hasta llevar al suicidio [9,10,11].

La apatía, síntoma neuropsiquiátrico caracterizado por una disminución de la motivación, presenta una elevada prevalencia en pacientes con demencia. No solo está asociada con una reducción de la actividad diaria y un aumento de estrés para el cuidador, sino también con un peor pronóstico de la enfermedad. Es uno de los primeros síntomas que aparece en estadios tempranos, sin embargo, son pocos los tratamientos

farmacológicos existentes y los estudios de intervenciones no farmacológicas aún no han sido replicados [12].

1.1.2. Justificación del TFM:

Hoy en día, el tiempo de exploración clínica es escaso, lo que dificulta el diagnóstico preciso de enfermedades complejas. Además, resulta laborioso trabajar de forma longitudinal, especialmente si se busca una frecuencia de visita elevada. Los sistemas de salud todavía no han dado una respuesta adecuada a la carga de trastornos mentales; en consecuencia, la divergencia entre la necesidad de tratamiento y su prestación es grande en todo el mundo. En los países de ingresos bajos y medios, entre un 76% y un 85% de las personas con trastornos mentales graves no recibe tratamiento; la cifra es alta también en los países de ingresos elevados: entre un 35% y un 50% [13].

Tras muchos años de investigación de la enfermedad, sigue sin existir terapia o tratamiento eficaz para curar la AD en estadios avanzados, por lo que la tendencia actual es la inclusión de sujetos en los ensayos de prevención de la AD pre-sintomática y MCI. El análisis del líquido cefalorraquídeo (en inglés, CSF), la evaluación neuropsicológica, y la tomografía de emisión de positrones amiloide (en inglés, PET) son métodos comunes usados para la detección temprana. El problema que presentan es que no son adecuados como cribado poblacional, ya sea por ser invasivos o costosos, lo que conlleva que no sean adecuados para entornos clínicos de atención primaria [14-17].

En el campo de la genética, gracias a la tecnología GWAS (estudio de asociación de genoma completo) se está pudiendo desarrollar herramientas útiles para el diagnóstico y prevención de enfermedades. El método implica el análisis de los genomas de múltiples personas y la búsqueda de marcadores genéticos que se pueden utilizar para predecir la presencia de una enfermedad. Una vez que dichos marcadores genéticos son identificados, se pueden utilizar para entender cómo los genes contribuyen a la enfermedad y desarrollar mejores estrategias de prevención y tratamiento [18]. Una de las herramientas más en alza es el cálculo de *polygenic risk scores* (PRS), el cual puede definirse como un valor que aúna el efecto estimado de múltiples variables genéticas para un fenotipo en un individuo concreto. El uso del PRS ha crecido recientemente por

su utilidad para detectar la etiología genética compartida entre los rasgos y puede actuar como un biomarcador para un fenotipo, entre otras aplicaciones. El problema que presentan este tipo de métodos es que requieren de una muestra amplia para el GWAS con un poder estadístico capaz de detectar regiones que, si bien tengan un efecto pequeño, sea consistente [19].

El consorcio de psiquiatría genómica (en inglés, PGC), busca la colaboración mundial para avanzar en el descubrimiento genético de conocimientos significativos desde el punto de vista biológico, clínico y terapéutico. Poseen, para varios trastornos psiquiátricos, estudios de GWAS que pueden ser usados para este tipo de abordajes. Entre estos, se encuentran los realizados sobre trastorno de depresión mayor, que podrían usarse para calcular el PRS de dicha enfermedad (PRS_{MDD}). El último artículo publicado señala un total de 101 loci de riesgo para dicha enfermedad [20,21].

El biobanco europeo de ADN sobre la enfermedad de Alzheimer (en inglés, EADB), en un proyecto multicéntrico que, en su último artículo, con un tamaño muestral de 39,106 AD y 401,577 controles, han reportado 75 loci de riesgo para la AD. Utilizaremos este GWAS para calcular el PRS (PRS_{AD}) [22,23].

Se ha descrito en la bibliografía el efecto de ambos PRS en la conversión de MCI a demencia. [24] Aquellos pacientes con un PRS_{AD} más elevado presentan una conversión más rápida. Este mismo comportamiento se observa al analizar PRS_{MDD} . Algunos autores sostienen que parte de las variantes alélicas de riesgo son comunes en entre ambas enfermedades [25] pero otros contraponen que estas enfermedades no presentan una heredabilidad conjunta [26].

Existen múltiples trabajos que relacionan síntomas neuropsiquiátricos como el depresivo, el ansioso y el apático con las demencias. También existen artículos sobre el uso de PRS_{AD} y PRS_{MDD} como biomarcadores para detectar el riesgo de desarrollar la enfermedad. El enfoque novedoso de este TFM es el uso de ambas herramientas de forma conjunta en el estudio de la conversión de pacientes MCI a demencia.

1.2 Objetivos del Trabajo

1.2.1. Objetivos generales:

1) Determinar el efecto predictivo de variables demográficas, neuropsiquiátricas, genómicas y su interacción en la conversión de un paciente con MCI a demencia

1.2.2. Objetivos específicos:

2) Familiarización del manejo de GWAS

3) Definir los PRS que se utilizarán en los distintos modelos de regresión (MDD, AD, sMDD y sAD)

4) Describir la asociación entre los PRS calculados y la sintomatología neuropsiquiátrica depresiva

5) Analizar el efecto de las variables de interés sobre el tiempo de supervivencia de los pacientes

1.3 Enfoque y método seguido

Dentro del centro en el que trabajo, Fundació ACE, poseemos una población de estudio proviene del proyecto GR@ACE de unos 10000 pacientes con GWAS [27]. En este conjunto de sujetos se encuentran diagnósticos de distintas demencias en diferentes estadios de progresión (quejas subjetivas, MCI, demencia) y de muchos de ellos se tienen seguimientos anuales que permiten estudios longitudinales. Otros datos de interés que permiten analizar la población con la que trabajamos son los datos demográficos (edad, género, educación, ...), varios test neuropsicológicos y el Inventario Neuropsiquiátrico (NPI) de Cummings [28]. El NPI es una entrevista estructurada cuyo objetivo es obtener información sobre la presencia de síntomas neuropsiquiátricos y conductuales en pacientes con alteraciones cerebrales y es ampliamente utilizado en el ámbito por ser válido, fiable y breve. En nuestro centro, el neurólogo/geriatra evalúa una versión reducida (NPI-Q) en español con la información suministrada por parte de un familiar o cuidador. [29]

En nuestro centro no tenemos información suficiente para poder determinar con certeza si un paciente presenta, además de una demencia, el MDD. Lo más similar que

poseemos es el ítem “Depresión” del NPI, que no es la enfermedad en sí misma, sino un síntoma neuropsiquiátrico, este es la estrategia más usual para la evaluación de esta sintomatología en demencia [30]. La diferencia radica en que, para hablar de MDD, dicho síntoma debe cronificarse y aparecer otros síntomas físicos, como dolores de cabeza o digestivos, y conductuales, como ansiedad o trastorno del sueño. Idealmente, esta enfermedad debe ser diagnosticada por un psicólogo tras un seguimiento exhaustivo y frecuente. En nuestro caso contamos con una frecuencia de seguimiento habitual de un año y la visita está realizada por neurólogos y centrada principalmente en la detección de demencia y su evolución. El NPI se realiza en cada visita, por lo tanto, poseemos datos longitudinales del mismo y podemos evaluar cómo se comporta en el tiempo [31].

Utilizaremos la regresión logística para construir los diferentes modelos para evaluar la utilidad de las covariables seleccionadas para discriminar qué participantes presentan sintomatología depresiva. Para el análisis de supervivencia, se empleará un modelo de riesgos proporcionales (regresión de Cox) el cual no impone como premisa una forma para la función de riesgo. Define un parámetro que es la razón entre ambas funciones de riesgo (*hazard ratio*) que estima el cociente entre las funciones de riesgo de dos curvas durante el período de seguimiento [32].

1.4 Planificación del Trabajo. Tareas y calendario:

1. Selección de pacientes y variables para el estudio [13-10-20 // 26-10-20]
 - a. Criterios de inclusión y exclusión de pacientes
 - b. Posible transformación de variables
2. Construcción de los PRS [27-10-20 // 16-11-20]
 - a. Completos para MDD (PRS_{MDD}) y AD (PRS_{AD})
 - b. Parciales, eliminando variantes genéticas comunes (PRS_{sMDD} y PRS_{sAD})
3. Analizar la asociación entre los modelos de PRS y las posibles evaluaciones longitudinales del síntoma neuropsiquiátrico “Depresión” y valoración de los mismos por curvas ROC [17-11-20 // 30-11-20]
 - a. Regresión logística para la discriminación de los individuos con sintomatología positiva

b. Curvas ROC de los modelos creados

4. Análisis de supervivencia con las variables estimadas [01-12-20 // 14-12-20]
5. Redactar memoria del TFM [15-12-20 // 05-01-21]
6. Preparación de la presentación de TFM [06-01-21 // 10-01-21]

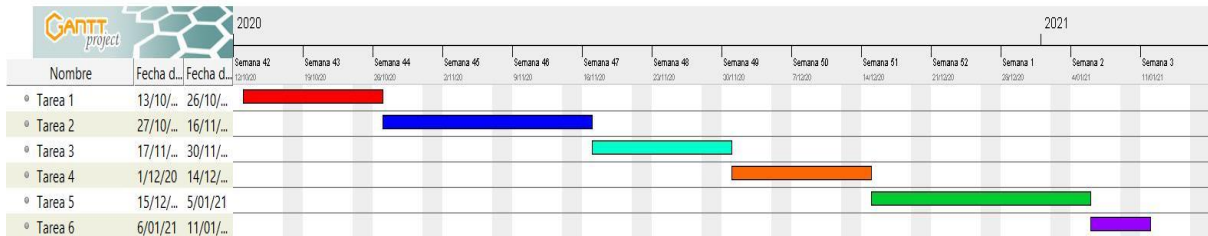


Figura 1: Diagrama de Gantt para exponer el tiempo previsto para los distintos hitos del TFM

1.5 Breve resumen de productos obtenidos

- Memoria del TFM
- Archivo Rmarkdown (.Rmd) y documento HTML con el preprocesamiento de los datos y la información desglosada de los diferentes modelos y gráficos planteados
- Presentación Powerpoint con vídeo

1.6 Breve descripción de los otros capítulos de la memoria

- **Capítulo 3:** Se presentan las variables demográficas, clínicas y genéticas con las que se trabajará, los criterios de selección empleados para definir la muestra y descriptivos básicos para tener una noción del tipo de población con la que estamos trabajando.
- **Capítulo 4:** Se analiza la relación entre el síntoma depresivo y el resto de las variables para entender cómo se comporta y si puede servir como *proxy* del MDD.
- **Capítulo 5:** Estudios de supervivencia para evaluar el tiempo de conversión de MCI a demencia para los pacientes.

2. Selección de pacientes y variables para el estudio

La cohorte proviene de Fundació ACE, donde los pacientes son evaluados en su Unidad de Trastornos de la Memoria, Instituto Català de Neurociències Aplicades, Barcelona, España [33]. El diagnóstico de MCI de cada paciente se realiza de forma consensuada por un equipo de neurólogos, neuropsicólogos y trabajadores sociales.

Los criterios de inclusión generales son:

- Tener un índice de demencia clínica (CDR) de 0,5 [34]
- A nivel escolar debe saber leer, escribir y saber hacer cálculo básico (no ser analfabeto funcional)
- No tener anomalías severas a nivel visual o auditivo
- Poseer un diagnóstico de MCI en la visita basal con la Unidad de Memoria
- Tener muestra de ADN disponible

Para este estudio en concreto, existen tres criterios más específicos:

- Disponibilidad de datos GWAS que hayan pasado el control de calidad
- Disponibilidad de genotipado de *APOE*
- Haberse realizado el cuestionario NPI en al menos dos visitas para poder realizar un estudio sobre su evolución

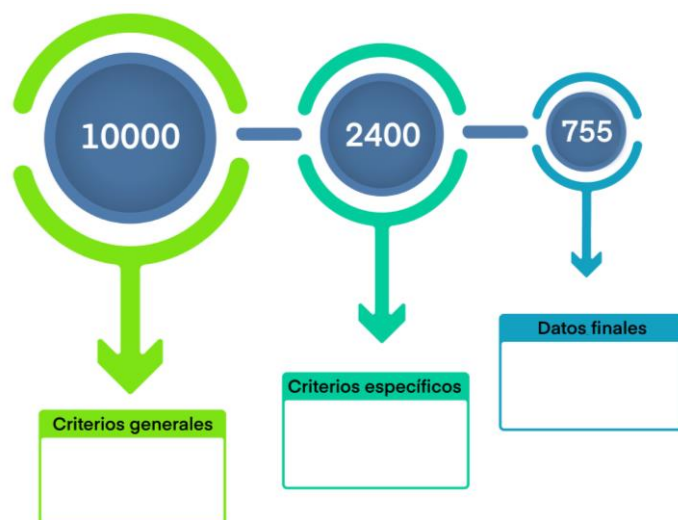


Figura 2: *Workflow* del número de pacientes según los criterios generales y específicos. La población final supone un 7.55% del total de los datos.

2.1. Variables trabajadas por la Unidad de Memoria

2.1.1. Variables demográficas

Age: Variable cuantitativa discreta. Edad del paciente en su primera visita

Sex: Variable cualitativa dicotómica. Sexo del paciente. Hombre // Mujer

Years_School: Variable cuantitativa discreta. Escolaridad expresada en años

Level_School_Dichotomic: Variable cualitativa dicotómica. Escolaridad clasificada en:

- Low: Llegueix i Escriu sense Escolaritat // Estudis Inferiors a 6 anys // Estudis Primaris Complerts de 6 anys // E. Primaris de 8 anys (Batx. Elemental o Primaria 6 + Comerç) // No consta
- High: Batxillerat Superior (Escola Tècnica, FP1) // Preu o COU o Mestre, ATS, Peritatge, Aparellador o FP2 // Diplomatures Universitaries // Llicenciatures //Mestratge o Doctorat

N= 755	
Sex, N (%):	
Hombre	267 (35.4%)
Mujer	488 (64.6%)
Age, Mediana [Q1;Q3]	
77.0 [71.0;82.0]	
Level_School_Dichotomic, N (%):	
Low	558 (73.9%)
High	197 (26.1%)
Years_School, Mediana [Q1;Q3]	
6.00[4.00;10.0]	

Tabla 1: Descriptivo general de variables demográficas

2.1.2. Variables clínicas

PROoPS: Variable cualitativa dicotómica. Probable // Posible. Se define como “MCI posible” cuando existen comorbilidades que puedan explicar o contribuir a los déficits cognitivos, y como “probable” cuando no hay ninguna.

DPI: Variable cuantitativa discreta. Muestra el número de dominios afectados en el paciente, utilizando los criterios de diagnóstico de Petersen [35,36]. Suele dicotomizarse como unidominio/multidominio.

Amnestic_non_amnestic_NBACE: Variable cualitativa dicotómica. Amnésico // No amnésico. Esta clasificación, utilizando los criterios de diagnóstico de Petersen [35,36], hace relación a cuál es la capacidad que se ve más afectada en el deterioro, en caso de ser la memoria, se considera que es un MCI amnésico (más frecuente y mayor probabilidad de evolucionar a AD). Para aquellos pacientes en los que el deterioro predomina en las otras zonas cognitivas, como puedan ser atención y función ejecutiva,

se consideran MCI no amnésicos (tienden a desarrollar otro tipo de demencias según el área más afectada) [37].

MMSE_DN: Variable cuantitativa discreta. Prueba escrita de cribado para evaluar MCI o demencia. Su puntuación máxima es 30, generalmente se considera como cognición alterada valores por debajo de 24 puntos [38].

Conversion_MCI_dementia: Variable dicotómica. Determina si, en algún punto del seguimiento del paciente, ha cambiado su diagnóstico de MCI a demencia.

Months_followup: Variable numérica. Determina el tiempo en meses hasta la conversión (en caso de que el paciente haya convertido a demencia) o el último diagnóstico (en caso de que el paciente no haya convertido en ninguna visita).

AP_depresio_A: Variable dicotómica. Presencia de MDD previo a la visita basal con la Unidad de Memoria.

Depressio_Disforia_G_SN.X // Apatia_indiferencia_G_SN.X // Ansietat_G_SN.X: Variable dicotómica. Ítem del test NPI que determina si presenta dicho síntoma neuropsiquiátrico en su visita número X, sin tener en cuenta la gravedad del mismo.

Depressio_Disforia_G_SN.total // Apatia_indiferencia_G_SN.total // Ansietat_G_SN.total: Variable dicotómica. Ítem del test NPI que determina si presenta dicho síntoma neuropsiquiátrico en alguna visita del seguimiento.

N= 755	
AP_depresio_A, N (%)	
No	527 (69.8%)
Yes	228 (30.2%)
Amnestic_non_amnestic_NBACE, N (%)	
Amnestic	605 (80.1%)
Non-amnestic	150 (19.9%)
PROoPS, N(%)	
Probable	320 (42.4%)
Possible	435 (57.6%)
Conversion_MCI_dementia, N (%)	
No	257 (34.0%)
Yes	498 (66.0%)
Months_followup, Mediana [Q1;Q3]	
18.0 [11.0;28.0]	
MMSE_DN, Mediana [Q1;Q3]	
26.0 [24.0;27.0]	
Depressio_Disforia_G_SN.1, N (%)	
No	305 (40.4%)
Yes	450 (59.6%)
Apatia_indiferencia_G_SN.1, N (%)	
No	395 (52.3%)
Yes	360 (47.7%)
Ansietat_G_SN.1, N (%)	
No	298 (39.5%)
Yes	457 (60.5%)

Tabla 2: Descriptivo general de variables clínicas

2.2. Variables genéticas

2.2.1. Genotipación, control de calidad e imputación

La extracción de ADN se realizó a partir de sangre periférica. Los participantes fueron genotipados usando el Axiom 815K del Biobanco Español (Thermo Fisher) en el Centro Nacional de Genotipado (CeGEN, Santiago de Compostela, España).

Los controles de calidad a nivel de muestra son:

- Tasa de genotipado $> 97\%$
- Correlación entre el sexo genotipado y el reportado
- Eliminación de pacientes atípicos a nivel genético *European cluster of 1000 Genomes* con un análisis de componentes principales (en inglés, PCA)
- Eliminación de pacientes genéticamente emparentados con otras muestras de la cohorte

Los controles de calidad a nivel de variante genética son:

- Tasa de genotipado $> 95\%$
- Frecuencia alélica menor (MAF) $> 0,01$
- En equilibrio Hardy-Weinberg ($p < 1 \times 10^{-4}$ en los controles)
- Sin faltas diferenciales entre los casos y los controles ($p < 1 \times 10^{-4}$)

La imputación se realizó utilizando el consorcio de referencia de haplotipos [39] (HRC, panel completo) y el panel de referencia de 1000 Genomas [40] (sólo para las inserciones/deleciones, comúnmente abreviadas “indels”) en el servidor de Imputación de Michigan [41]. Posteriormente, se descartan las variantes raras (MAF $< 0,001$) y con baja calidad de imputación ($R^2 < 0,30$) para no entorpecer el análisis posterior.

Los modelos de regresión logística, creados usando PLINK (v2.00a), fueron ajustados para los cuatro primeros componentes principales de la ascendencia. El número de variantes genéticas obtenido para el análisis es 14.542.816.

Toda esta información sobre el procesamiento genético de los datos del centro se encuentra ampliada en los artículos ya publicados del grupo de investigación al que pertenezco. [42,43]

2.2.2. *APOE*

El gen *APOE*, localizado en el cromosoma 19 y habitualmente asociado a la AD, tiene tres formas comunes [44]:

- *APOE* $\epsilon 2$ disminuye el riesgo de padecer AD.
- *APOE* $\epsilon 4$ aumenta el riesgo de padecer AD.
- *APOE* $\epsilon 3$ (la más común), no parece incidir en el riesgo

Con estos alelos, los genotipos posibles son $\epsilon 2\epsilon 2$, $\epsilon 2\epsilon 3$, $\epsilon 2\epsilon 4$, $\epsilon 3\epsilon 3$, $\epsilon 3\epsilon 4$, $\epsilon 4\epsilon 4$. Se plantea trabajar con *APOE* utilizando dos transformaciones diferentes:

APOE_dicot: Variable cualitativa dicotómica. Genotipo *APOE* dicotomizado como ausencia/presencia del alelo $\epsilon 4$.

APOE_dosage: Variable cuantitativa discreta. *APOE* categorizada según la dosis del alelo $\epsilon 4$. Para ello consideramos como $\epsilon 2$ como protector (-1), $\epsilon 3$ como neutro (0) y $\epsilon 4$ como riesgo (+1).

2.2.3. Construcción de PRS

El punto de partida de los PRS son los resultados de un GWAS. Para MDD será el último artículo publicado por el PGC y para AD, el último artículo publicado por EADB.

En líneas generales, se han seleccionado aquellas variantes genéticas estadísticamente significativas e independientes entre sí. Para cada una de las variantes genéticas, la dosis del alelo fue multiplicada por el logaritmo de la OR y se realizó el sumatorio para obtener el PRS. Para mejor comprensión a nivel descriptivo, se aplica una transformación adicional para que el PRS represente el número de alelos de riesgo "promedio" [45].

Como el efecto de *APOE* en AD es tan consistente, en el PRS_{AD} no se incluye *APOE*, sino que se plantea como una variable independiente [45].

Se crean dos variables nuevas para reflejar el contexto genético de los pacientes:

PRS_AD: Variable cuantitativa continua. Estima el efecto de múltiples variantes genéticas sobre Alzheimer.

PRS_MDD: Variable cuantitativa continua. Estima el efecto de múltiples variantes genéticas sobre el trastorno de depresión mayor.

N= 755	
APOE_dichotomic, N (%)	
Without e4	477 (63.2%)
With e4	278 (36.8%)
APOE_genotype, N (%)	
e2e2	4 (0.53%)
e2e3	46 (6.09%)
e2e4	10 (1.32%)
e3e3	417 (55.2%)
e3e4	247 (32.7%)
e4e4	31 (4.11%)
APOE_dosage, N(%)	
-2	4 (0.53%)
-1	46 (6.09%)
0	427 (56.6%)
1	247 (32.7%)
2	31 (4.11%)
PRS_AD_GRACE, Media (SD)	
66.3 (5.21)	
PRS_MDD_GRACE, Media (SD)	
97.9 (6.55)	

Tabla 3: Descriptivo general de variables genéticas


```
> plot(t3["PRS_MDD_GRACE"])
```

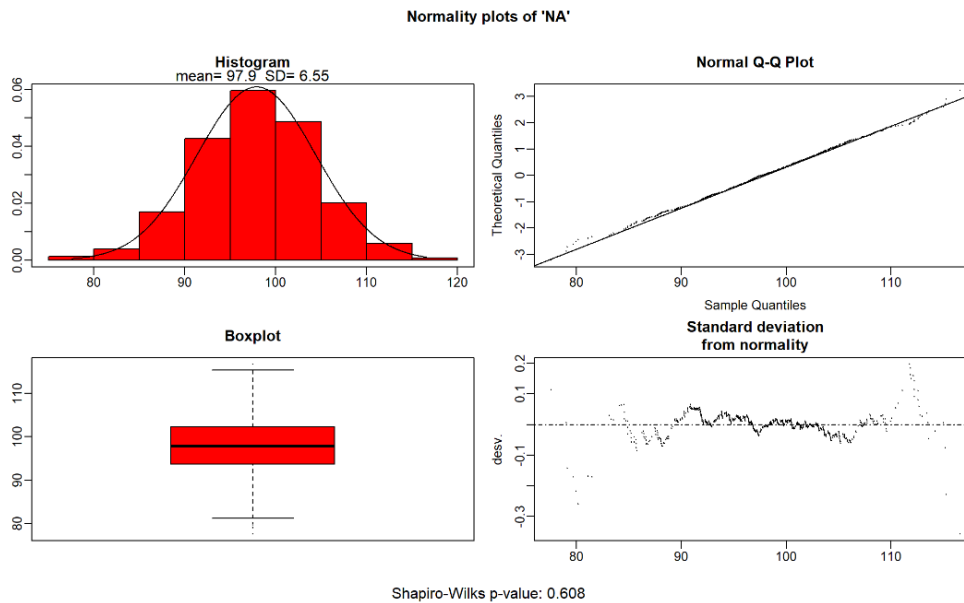


Figura 3: Gráficos para evaluar si la variable PRS_MDD sigue una distribución normal. Puede apreciarse que se asemeja bastante a dicha distribución. La variable PRS_AD tiene un comportamiento similar al aquí mostrado.

3. Validación del síntoma depresivo

3.1 Planteamiento teórico. Regresión logística, OR o RR

Para modelar la probabilidad de un evento dicotómico como presencia/ausencia, en estadística suele ser común el uso de un modelo de regresión logística. Para ello, calcula el logaritmo de las probabilidades como una combinación lineal de las variables independientes añadidas (que pueden ser continuas o categóricas) [46].

La magnitud del efecto de las variables en el modelo puede plantearse en forma de *odds ratio* (OR) o *risk ratio* (RR). La diferencia entre ambas es que OR es la razón entre la probabilidad de que un evento ocurra y la probabilidad de que no ocurra mientras que RR es la razón del riesgo de ocurrencia entre los dos grupos (el riesgo de ocurrencia es la división del número de casos de interés entre el total de casos). Ambas opciones son viables, pero el OR presentan ciertas ventajas como [47]:

- El OR determina un estimado para las relaciones entre variables dicotómicas, lo que aporta una interpretación más adecuada en términos de efectividad.
- El OR permite evaluar el efecto de otras variables sobre las relaciones usando la regresión logística.

En R, la función genérica `glm()` permite el cálculo del modelo logístico [48]. *family* sirve para indicar el componente aleatorio del modelo y la función de enlace que se utilizará. En nuestro caso, `family = binomial(link=logit)`.

3.2 Modelos

3.2.1 Correlación entre ambos PRS

Con el fin de evaluar la relevancia de los PRS generados junto con las variables clínicas y demográficas en la presencia/ausencia del síntoma depresivo, se calculan un total de 16 modelos de regresión logística combinando distintas hipótesis:

- El sexo se emplea únicamente como variable de ajuste o existe interacción con alguno de los PRS
- El síntoma neuropsiquiátrico depresivo se evalúa como presencia/ausencia en la visita basal o se plantea como presencia/ausencia en alguna visita
- Otros síntomas neuropsiquiátricos relevantes (ansiedad y apatía) influyen o no en el modelo
- *APOE* se plantea como variable dicotómica o como dosis alélica

Para confirmar que ambos PRS son independientes entre sí, vamos a enfrentar ambas variables entre sí en un gráfico de dispersión [Figura 4] y calcular una correlación de Pearson. No se observa un patrón en la nube de polvo y el p-valor de la prueba es 0.7, por lo que mantenemos la hipótesis nula de que no existe correlación entre ambas variables.

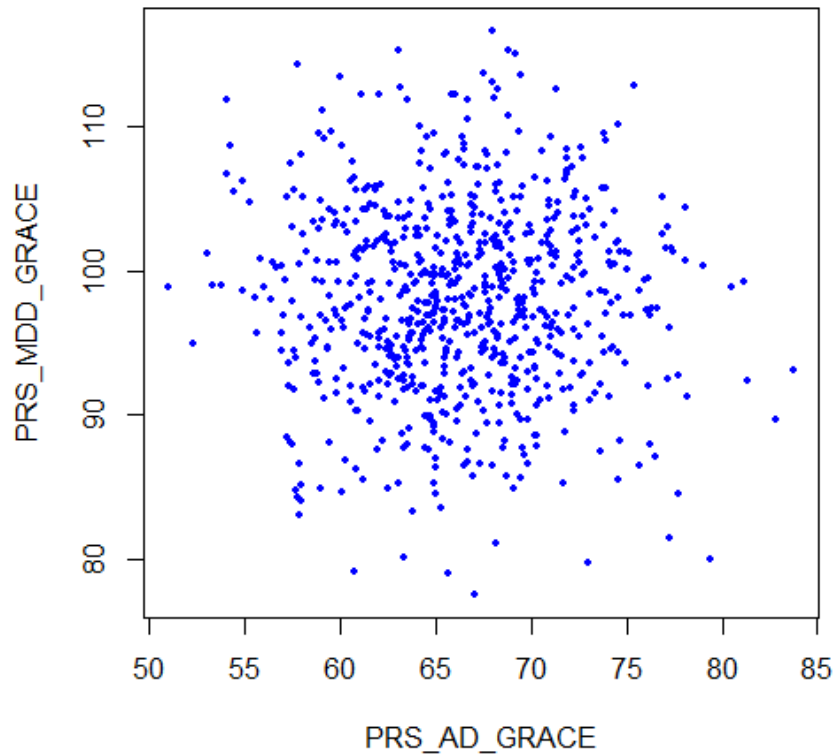


Figura 4: Gráfico de puntos comparando la distribución del PRS_{MDD} y PRS_{AD}.

3.2.2. Regresión logística

Variables	Modelos	
	1	2
PRS_AD_GRACE	1.44 (0.90 - 2.31)	1.70 (0.97 - 2.95) .
PRS_MDD_GRACE	1.29 (0.94 - 1.77)	1.42 (0.97 - 2.07) .
PRS_AD_GRACE *PRS_MDD_GRACE	0.99 (0.99 - 1.00)	0.99 (0.99 - 1.00) .
APOE_dichotomic	1.25 (1.05 - 1.50) **	0.99 (0.80 - 1.23)
Sex	1.96 (1.37 - 2.83) ***	2.87 (1.88 - 4.39) ***
Age	1.04 (1.02 - 1.07) ***	1.04 (1.02 - 1.07) ***
Years_School	1.05 (1.01 - 1.09) **	1.04 (0.99 - 1.09)
Apatia_indiferencia_G_SN.1 // total	3.26 (2.31 - 4.65) ***	3.39 (2.16 - 5.34) ***
Ansietat_G_SN.1 // total	5.63 (3.99 - 8.02) ***	6.52 (4.18 - 10.26) ***

Tabla 4: Modelos de regresión logística para la predicción del síntoma depresivo. En el modelo 1 se trabaja con los síntomas basales y en el modelo 2 se trabaja con los totales. P-value: <0.001 ‘***’ // ‘**’ <0.01 // ‘*’ <0.05 // ‘.’ <0.1

Aunque parece que existe cierta tendencia por parte de ambos PRS y de la interacción entre ellos, queda diluida frente a los otros dos síntomas neuropsiquiátricos, los cuales

en ambos modelos presentan una asociación bastante fuerte. Esto es coherente ya que en muchas ocasiones suelen presentarse de forma conjunto o enmascarse unos con otros. Variables demográficas como sexo o edad, también resultan significativas, siendo más común en mujeres que en hombres.

4. Estudios de conversión MCI - Demencia

4.1 Planteamiento teórico

Este tipo de abordajes entran dentro de los análisis de supervivencia. Un análisis de supervivencia se define como un conjunto de técnicas estadísticas planteadas para estudiar el tiempo que tarda en transcurrir un evento. En nuestro centro, el punto de partida es ese primer diagnóstico de MCI en la Unidad de Memoria y el evento a estudiar es el cambio de diagnóstico a demencia.

Dentro de las diferentes herramientas existentes, se realizará el estudio construyendo modelos de Cox (también llamado de riesgos proporcionales). Se basa en calcular la ratio entre dos funciones de riesgos para obtener el *hazard ratio* (HR). Este método está clasificado como “semiparamétrico” porque, aunque no asume ninguna forma con respecto a la función de riesgo basal, $h_0(t)$, las covariables entran a través de un modelo lineal con sus correspondientes parámetros [Figura 5]. Esto permite focalizar en qué características permiten alargar ese tiempo de vida más que en concretar qué tiempo concreto vivirá (entendiendo ese tiempo de vida como periodo sin demencia). La asunción que sí realiza el modelo es que el riesgo de los grupos de pacientes se mantiene constante en el tiempo [32].

$$HR = \frac{h_0(t) \exp\left(\sum_{j=1}^p \beta_j X_j^*\right)}{h_0(t) \exp\left(\sum_{j=1}^p \beta_j X_j\right)} = \frac{\exp\left(\sum_{j=1}^p \beta_j X_j^*\right)}{\exp\left(\sum_{j=1}^p \beta_j X_j\right)} = \exp\left(\sum_{j=1}^p \beta_j (X_j^* - X_j)\right)$$

Figura 5: Proporción de riesgos aplicada en el modelo de Cox para calcular HR

Una de las principales ventajas es que permite considerar al mismo tiempo distintas variables y evaluar si existen relación entre ellas y si su efecto persiste incluso teniendo en cuenta otras. Además, la interpretación de estas variables es sencilla, ya que a cada una de ellas se les define una HR concreta [49].

Una de las limitaciones que tenemos con respecto a los datos es que no nos es posible aplicar el concepto de “censura” al estudio. Este término hace referencia a aquellos pacientes de los que no se tiene información completa, ya sea por razones tan variadas como fallecimiento o abandono del estudio. Para ninguno de estos dos supuestos disponemos de información al respecto [49].

En Rstudio, la función `coxph()` del paquete *survival* permite, a partir de un objeto `Surv` creado con la función `survfit()` y de las covariables de interés, calcular un modelo con las siguientes variables básicas [50]:

- **coef:** Coeficiente estimado de la beta.
- **exp(coef):** Exponencial elevado al coeficiente beta estimado. HR
- **se(coef):** La desviación estándar de la estimación.
- **z:** Valor del estadístico para prueba Wald (testar la hipótesis $\beta_i = 0$, en muestras grandes se asemeja a una distribución normal).
- **p:** P-valor de la prueba Wald.

Además, la función `ggforest()` permite representar de forma rápida y cómoda los HR de las variables incluidas en el modelo. Además, aporta información como:

- **AIC:** *Criterio de información de Akaike*. Medida que aporta una solución de compromiso entre la bondad de ajuste y la complejidad de un modelo, pudiendo ser usada como aproximación a la calidad del modelo [51].
- **Concordance index:** Medida orientativa de la concordancia. $0 \leq x \leq 1$ [52].

A la hora de la interpretación, una variable que presente un HR mayor a 1 indica que aumenta el riesgo cuando esa característica aumenta. Lo opuesto ocurre cuando el HR es inferior a 1.

Al igual que con las regresiones logísticas, no se plantean en este documento todos los modelos diseñados, para mantener la claridad en el trabajo. El resto de las tablas realizadas se localizan en el Anexo 1.

4.2. Modelos

En primer lugar, planteamos el modelo empleando únicamente las dos variables de estudio, el PRS_{MDD} y el PRS_{AD}. A continuación, probamos comprobaremos un modelo en el que se incluyan todas las variables de interés, planteando los síntomas neuropsiquiátricos solo de la visita basal. Finalmente, repetiremos el modelo anterior con los síntomas longitudinales, para ver qué aproximación se ajusta mejor con respecto a la conversión. Los datos se muestran tanto de forma individual [Figuras 6 y 7] como comparativas [Tabla 5].

Puede observarse que ambos PRS aumentan el riesgo de forma estadísticamente significativa cuando planteamos los síntomas basales. Sin embargo, al realizar un enfoque más longitudinal, comprobamos que la apatía presenta un gran impacto en el riesgo. Teniendo en cuenta que la AD aparece con mayor frecuencia con la edad, es coherente que sea una variable que resulte muy relevante en ambos modelos.

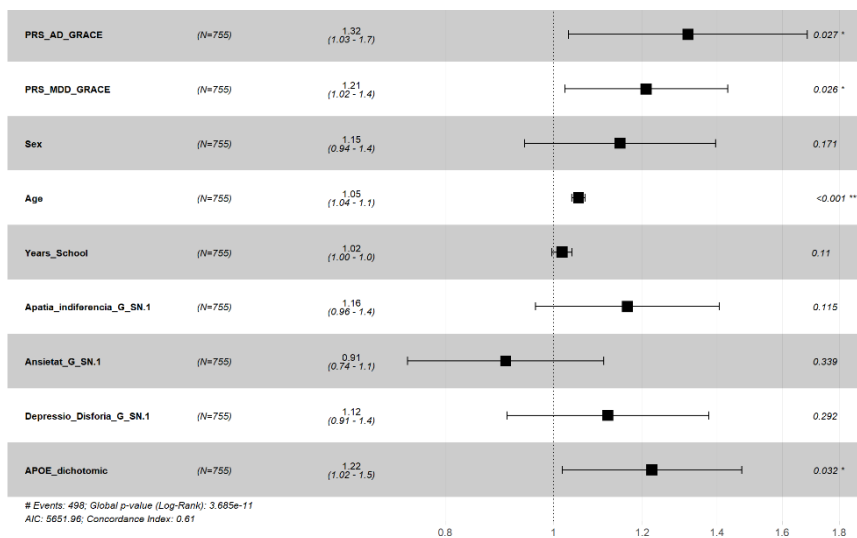


Figura 6: HR del modelo de Cox utilizando los síntomas neuropsiquiátricos basales

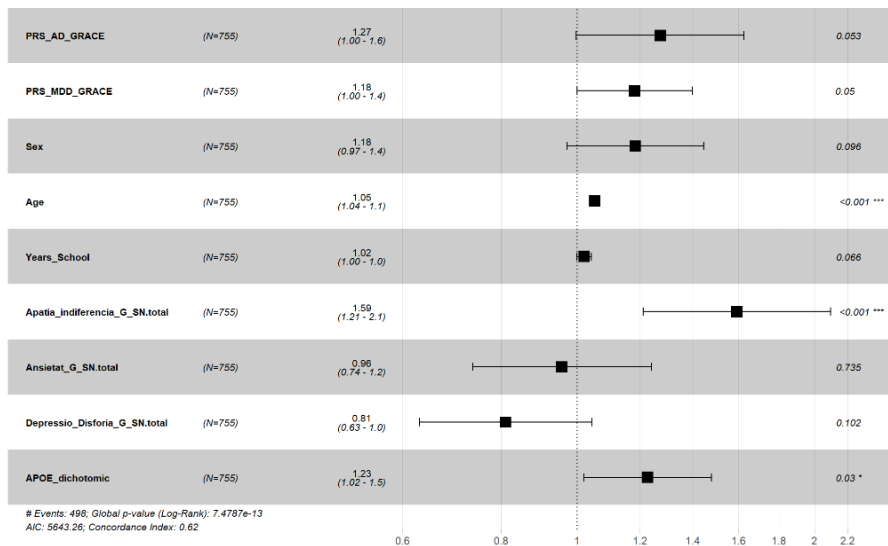


Figura 7: HR del modelo de Cox utilizando los síntomas neuropsiquiátricos totales

Variables	Modelos		
	0	1	2
PRS_AD_GRACE	1.003 (0.9861 - 1.020)	1.3188 (1.0316 - 1.6860) *	1.2738 (0.9969 - 1.627) .
PRS_MDD_GRACE	1.006 (0.9925 - 1.021)	1.2010 (1.0231 - 1.4307) *	1.1821 (0.9998 - 1.398) .
PRS_AD_GRACE * PRS_MDD_GRACE		0.9972 (0.9947 - 0.9997) *	0.9976 (0.9951 - 1.000) .
APOE_dichotomic		1.2244 (1.0176 - 1.4732) *	1.2279 (1.0197 - 1.479) *
Sex		1.1467 (0.9426 - 1.3951)	1.1845 (0.9701 - 1.446) .
Age		1.0525 (1.0379 - 1.0673) ***	1.0526 (1.0380 - 1.067) ***
Years_School		1.0171 (0.9962 - 1.0386)	1.0198 (0.9987 - 1.041) .
Depressio_Disforia_G_SN.1		1.1183 (0.9082 - 1.3768)	0.8109 (0.6306 - 1.043)
Apatia_indiferencia_G_SN.1		1.1640 (0.9637 - 1.4060)	1.5932 (1.2112 - 2.096) ***
Ansietat_G_SN.1		0.9061 (0.7402 - 1.1092)	0.9560 (0.7365 - 1.241)

Tabla 5: Modelos de Cox planteados para la conversión. HR (95% IC). P-value: <0.001
‘***’ // ‘**’ <0.01 // ‘*’ <0.05 // ‘.’ <0.1

5. Conclusiones

5.1. Ideas finales

- Existe un aumento del riesgo de conversión de MCI a demencia en pacientes apáticos con elevado PRS de Alzheimer y depresión. Ambos PRS presentan una interacción con tendencia a la significación. La edad es una variable clave en la degeneración cognitiva del paciente.
- En nuestra cohorte no tenemos un tamaño muestral suficiente para poder evaluar la relación entre los PRS y la sintomatología depresiva. Sin embargo, podemos afirmar que está altamente relacionada con otros síntomas neuropsiquiátricos como la ansiedad o la apatía y que el sexo es una variable relevante; las mujeres tienen una tendencia más elevada a presentarla.
- Plantear enfoques longitudinales de síntomas neuropsiquiátricos, aumenta el peso de estos en la conversión en detrimento de los PRS.

5.2. Reflexión y análisis del seguimiento de planificación y metodología

Tras este trabajo, he podido entender lo complejo que resulta llevar a cabo un análisis estadístico completo. Por un lado, por la heterogeneidad de la población, que en casos como este puede haber una ventana de más de 20 años entre la visita del primer paciente y la del último. Por otro, por la necesidad encontrar un equilibrio entre el número de variables a analizar y la exhaustividad en la elaboración del modelo predictivo, pues en ocasiones plantear y evaluar todas las posibilidades puede suponer un proceso lento.

El objetivo principal se ha cumplido, y gran parte de los secundarios también. Se ha decidido dejar para líneas futuras el cálculo de los PRS específicos de ambas enfermedades eliminando aquellas variables que fuesen comunes (PRS_{sMMD} y PRS_{sAD}), tanto por falta de tiempo como por contradicciones en la bibliografía. Con respecto a la planificación temporal del proyecto, no ha sido posible cumplirla debido a varios motivos:

- La temporada de Covid ha alterado las tareas realizadas por los miembros del centro, lo que supuso un retraso de los datos de partida superior a una semana.

Esto muestra que, aparte del tiempo real de dedicación a un proyecto, siempre es necesario dejar un margen para posibles contratiempos, ya sean laborales o personales.

- Subestimación del tiempo de aprendizaje de softwares nuevos. Herramientas bioinformáticas como PLINK pueden resultar áridas en un primer abordaje si no se han utilizado o estudiado con anterioridad. Es conveniente informarse antes de forma general para tener una aproximación sobre la complejidad de las operaciones a realizar con dicho software.
- No plantear la redacción completa de los capítulos tras la realización de los mismos, dejando para el final toda la labor de redacción y revisión. Aunque esto permite ser flexible en los primeros capítulos al tener ya conocimiento del contenido de los últimos, puede suponer un riesgo en caso de que se retrase el proyecto, haciendo que las revisiones del mismo no puedan ser todo lo exhaustivas que debiese.
- Discrepancias entre las regresiones logísticas calculadas según Rstudio y según STATA. De aquí podemos plantear que, en ocasiones, es interesante realizar alguno de los análisis en otra plataforma para verificar que se obtienen resultados similares y no existen problemas de formato.

5.3. Líneas de trabajo futuras

- Aumentar la población del estudio. Este tipo de abordajes suele requerir un tamaño muestral elevado para poder apreciar efectos sutiles de las variables. Entre las posibilidades para ello se encuentra genotipar a más pacientes del centro (fue uno de los criterios que más redujo la población de estudio) y diseñar sistemas nuevos de inclusión de pacientes (como aquellos que se diagnostican de forma basal como “queja subjetiva de memoria” y en visitas posteriores son diagnosticados como MCI).
- Replicar este estudio con otro set de pacientes genotipados, para verificar que el efecto se mantiene en diferentes muestras y no es producto del azar.

- Exploración de distintos bloques de PRS. Se ha cuestionado si ambas enfermedades presentan variantes alélicas comunes. Sería un buen planteamiento replicar esos estudios y trabajar de forma conjunta con PRS parciales y completos para evitar aplicar por duplicados aquellas variantes de riesgo para ambas enfermedades.
- Ampliar los análisis de supervivencia, ya sea aplicando diferentes transformaciones a las variables actuales (trabajar los PRS por terciles para agrupar a los pacientes por riesgo bajo, medio y alto) o incluyendo nuevas variables clínicas, como puedan ser el DPI, PROoPS, etc.
- Ampliar la colección del centro de CSF para poder añadir biomarcadores al análisis. Actualmente es una información que se usa en nuestro centro como parte del diagnóstico, pero aún requiere un número de muestras más elevado para poder emplearse en estudios de estas características.
- Explorar la conversión MCI→AD en concreto en lugar de la conversión a demencia. También podría plantearse la reversión de estos pacientes AD→MCI.

6. Glosario

AD: *Alzheimer's disease*. Tipo de demencia más frecuente, que aparece principalmente durante la vejez [4].

AIC: *Criterio de información de Akaike*. Medida que aporta una solución de compromiso entre la bondad de ajuste y la complejidad de un modelo, pudiendo ser usada como aproximación a la calidad del modelo [51].

DNA: *Deoxyribonucleic acid*. Ácido nucleico con la información genética para el funcionamiento y desarrollo de los seres vivos.

APOE: *Apolipoprotein E*. Proteína que participa en el transporte de lípidos. Su alelo ε4 tiene una elevada prevalencia en pacientes con AD [44].

CDR: *Clinical dementia rating*, es una escala global de cinco puntos. Sus contenidos se dividen en seis ejes (memoria, orientación, juicio y solución de problemas, trabajo en la comunidad, rendimientos en casa y en aficiones y cuidado personal) [34].

CIE-10: Clasificación internacional de enfermedades, 10.^a edición. Clasificación y codificación de las enfermedades y una amplia variedad de signos, síntomas, hallazgos anormales, denuncias, circunstancias sociales y causas externas de daños y/o enfermedad [2].

CSF: *Cerebrospinal fluid*. Líquido que rodea el cerebro y la médula espinal. En el diagnóstico de AD se extrae para evaluar la concentración de biomarcadores como β -amiloide, la proteína tau-total y la proteína tau-fosforilada (p-tau) [53].

SD: *Standard deviation*. Medida de dispersión más común en estadística.

DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Sistema de clasificación de los trastornos mentales su diagnóstico, estudio y tratamiento por parte de clínicos e investigadores [3].

EADB: Banco de *DNA* europeo para el descifrado de la heredabilidad de la AD [22].

GWAS: *Genome-wide association study*. Análisis de una variación genética a lo largo de todo el genoma humano con el objetivo de identificar su asociación a un rasgo observable [22].

HR: *Hazard ratio*. Riesgo relativo de un evento comparando dos grupos de un ensayo [32].

MAF: *Minor allele frequency*. Frecuencia del alelo menos común en un locus concreto para una población determinada.

MCI: *Mild Cognitive Impairment*. Situación de cambios cognitivos que no se justifican por la edad, sino que son consecuencia de alguna condición médica, ya sea la AD u otra [5].

MDD: *Major depressive disorder*. Trastorno emocional que cursa, entre otros síntomas, con un sentimiento de tristeza constante y una pérdida de interés en realizar diferentes actividades [8].

MMSE: *Mini-Mental State Examination*. Prueba escrita de cribado para evaluar MCI o demencia. Su puntuación máxima es 30, generalmente se considera como cognición alterada valores por debajo de 24 puntos [38].

NPI-Q: *Neuropsychiatric Inventory Questionnaire*. Entrevista estructurada cuyo objetivo es obtener información sobre la presencia de síntomas neuropsiquiátricos y conductuales en pacientes con alteraciones cerebrales [28].

OMS: Organización Mundial de la Salud. Organismo para alcanzar para todos los pueblos el máximo grado de salud (física, mental y social).

OR: *Odds ratio*. Medida estadística definida como la razón entre la probabilidad de que un evento ocurra y la probabilidad de que no ocurra [47].

PCA: *Principal component analysis*. Técnica para describir un conjunto de datos como nuevas variables no correlacionadas entre sí, buscando reducir la dimensionalidad de estos [54].

PGC: *Psychiatric Genomics Consortium*. El consorcio de psiquiatría genómica busca la colaboración mundial para avanzar en el descubrimiento genético de conocimientos significativos desde el punto de vista biológico, clínico y terapéutico [20].

PRS: *Polygenic risk scores*. Valor que aúna el efecto estimado de múltiples variables genéticas para un fenotipo en un individuo concreto, normalmente calculado como una suma ponderada de alelos [19].

RR: *Risk ratio*. Medida estadística definida como la razón del riesgo de ocurrencia entre los dos grupos [47].

7. Bibliografía

- [1] World Health Organization. (2018, 19 abril). Mental disorders. https://www.who.int/mental_health/management/en/
- [2] Uribe, M. O. (1996). Clasificación Internacional de Enfermedades, Organización Mundial de la Salud. Décima Versión CIE-10. *Salud Mental*, 19(Supl 2), 11-18.
- [3] Reus, V. I., Fochtmann, L. J., Bukstein, O., Eyler, A. E., Hilty, D. M., Horvitz-Lennon, M., ... & McIntyre, J. (2018). The American Psychiatric Association practice guideline for the pharmacological treatment of patients with alcohol use disorder. *American Journal of Psychiatry*, 175(1), 86-90.
- [4] World Health Organization: WHO. (2020b, septiembre 21). Demencia. [who. https://www.who.int/es/news-room/fact-sheets/detail/dementia](https://www.who.int/es/news-room/fact-sheets/detail/dementia)
- [5] Manuales MSD. (2020). Demencia. Manual MSD versión para profesionales. <https://www.msdmanuals.com/es-es/professional/trastornos-neuro%C3%B3gicos/delirio-y-demencia/demencia>
- [6] Ruderfer, D. M., Fanous, A. H., Ripke, S., McQuillin, A., Amdur, R. L., Gejman, P. V., ... & Kelsoe, J. R. (2014). Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Molecular psychiatry*, 19(9), 1017-1024.
- [7] Lutz, M. W., Sprague, D., Barrera, J., & Chiba-Falek, O. (2020). Shared genetic etiology underlying Alzheimer's disease and major depressive disorder. *Translational psychiatry*, 10(1), 1-14.
- [8] Depresión. (n.d.). Retrieved September 28, 2020, from <https://www.who.int/es/news-room/fact-sheets/detail/depression>
- [9] Robins LN, Regier DA. Psychiatric disorders in America: the Epidemiologic Catchment Area Study. New York, The Free Press, 1991.
- [10] Consorcio Internacional en Epidemiología Psiquiátrica. Comparación transnacional de la prevalencia de los trastornos mentales y los factores con ellos correlacionados, 2004.
- [11] Vallejo Ruiloba, J., & Gastó Ferrer, C. (2000). Trastornos afectivos: ansiedad y depresión.
- [12] Sherman, C., Liu, C. S., Herrmann, N., & Lanctôt, K. L. (2018). Prevalence, neurobiology, and treatments for apathy in prodromal dementia. *International psychogeriatrics*, 30(2), 177-184.

[13] Trastornos mentales. (n.d.). Retrieved September 28, 2020, from <https://www.who.int/es/news-room/fact-sheets/detail/mental-disorders>

[14] Soininen, H. S., & Scbelstens, P. (1998). Early diagnostic indices for the prevention of Alzheimer's disease. *Annals of medicine*, 30(6), 553-559.

[15] Monge-Argilés, J. A., Sanchez-Paya, J., Munoz-Ruiz, C., Pampliega-Perez, A., Montoya-Gutierrez, J., & Leiva-Santana, C. (2010). Biomarkers in the cerebrospinal fluid of patients with mild cognitive impairment: a meta-analysis of their predictive capacity for the diagnosis of Alzheimer's disease.

[16] Toledo, J. B., Shaw, L. M., & Trojanowski, J. Q. (2013). Plasma amyloid beta measurements—a desired but elusive Alzheimer's disease biomarker. *Alzheimer's research & therapy*, 5(2), 8.

[17] Aizenstein, H. J., Nebes, R. D., Saxton, J. A., Price, J. C., Mathis, C. A., Tsopelas, N. D., ... & Bi, W. (2008). Frequent amyloid deposition without significant cognitive impairment among the elderly. *Archives of neurology*, 65(11), 1509-1517.

[18] Estudio de asociación de genoma completo (GWAS) | NHGRI. (2020). *Genome.gov*. <https://www.genome.gov/es/genetics-glossary/Estudio-de-asociacion-de-genoma-completo>

[19] Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genet*, 9(3), e1003348.

[20] Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., ... & Bacanu, S. A. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics*, 50(5), 668-681.

[21] Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shirali, M., ... & Alloza, C. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience*, 22(3), 343-352.

[22] <https://www.era-learn.eu/network-information/networks/jpco-fund/jpcfund-a-call-for-european-research-projects-on-neurodegenerative-diseases-risk-and-protective-factors-longitudinal-cohort-approaches-and-advanced-experimental-models/european-dna-bank-to-decipher-the-alzheimer-disease-missing-heritability>

[23] Bellenguez, C., Küçükali, F., Jansen, I., Andrade, V., Moreno-Grau, S., Amin, N., ... & Boland, A. (2020). New insights on the genetic etiology of Alzheimer's and related dementia. *medRxiv*.

- [24] Xu, J., Li, Q., Qin, W., Jun Li, M., Zhuo, C., Liu, H., ... & Yu, C. (2018). Neurobiological substrates underlying the effect of genomic risk for depression on the conversion of amnesic mild cognitive impairment. *Brain*, 141(12), 3457-3471.
- [25] Lutz, M. W., Sprague, D., Barrera, J., & Chiba-Falek, O. (2020). Shared genetic etiology underlying Alzheimer's disease and major depressive disorder. *Translational psychiatry*, 10(1), 1-14.
- [26] Brainstorm Consortium. (2018). Analysis of shared heritability in common disorders of the brain. *Science (New York, NY)*, 360(6395).
- [27] Web, A. (n.d.). GR@ACE Genomic Research at Fundació ACE: Making Alzheimer's history. Retrieved September 28, 2020, from <https://www.fundacioace.com/es/grace.html>
- [28] Cummings, J. L. (1997). The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. *Neurology*, 48(5 Suppl 6), 10S-16S.
- [29] Sergi, V., Marta, M., De Itziar, R., Ana, E., Sonia, M. G., Adelina, O., ... & Montse, A. (2020). Interaction of neuropsychiatric symptoms with APOE ϵ 4 and conversion to dementia in MCI patients in a Memory Clinic. *Scientific Reports*, 10(1), 1-10.
- [30] Manuales MSD. (2020). Demencia. Manual MSD versión para profesionales. <https://www.msmanuals.com/es-es/professional/trastornos-neurologicos/delirio-y-demencia/demencia>
- [31] Lai, C. K. (2014). The merits and problems of Neuropsychiatric Inventory as an assessment tool in people with dementia and other neurological disorders. *Clinical interventions in aging*, 9, 1051.
- [32] Boj del Val, E. (2014). El modelo de regresión de Cox.
- [33] Boada, M., Tárraga, L., Hernández, I., Valero, S., Alegret, M., Ruiz, A., ... & Clinic, M. (2014). Design of a comprehensive Alzheimer's disease clinic and research center in Spain to meet critical patient and family needs. *Alzheimer's & Dementia*, 10(3), 409-415.
- [34] Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., & Berg, L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. *Archives of neurology*, 58(3), 397-405.

- [35] Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of neurology*, 56(3), 303-308.
- [36] Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, 256(3), 183-194.
- [37] Pose, M., & Manes, F. (2010). Deterioro cognitivo leve. *Acta Neurol Colomb*, 26(3 suppl 3), S7-12.
- [38] Shulman, K. I., Herrmann, N., Brodaty, H., Chiu, H., Lawlor, B., Ritchie, K., & Scanlan, J. M. (2006). IPA survey of brief cognitive screening instruments. *International psychogeriatrics*, 18(2), 281.
- [39] McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., ... & Luo, Y. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics*, 48(10), 1279-1283.
- [40] 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68-74.
- [41] <https://imputationserver.sph.umich.edu>
- [42] Moreno-Grau, S., de Rojas, I., Hernández, I., Quintela, I., Montreal, L., Alegret, M., ... & Rosende-Roca, M. (2019). Genome-wide association analysis of dementia and its clinical endophenotypes reveal novel loci associated with Alzheimer's disease and three causality networks: The GR@ ACE project. *Alzheimer's & Dementia*, 15(10), 1333-1347.
- [43] de Rojas, I., Moreno-Grau, S., Tesi, N., & Grenier-Boley, B. (2019). Common variants in Alzheimer's disease: Novel association of six genetic variants with AD and risk stratification by polygenic risk scores. *medRxiv*, (19012021).
- [44] *Stolerman, I. (Ed.). (2010). Encyclopedia of psychopharmacology. Springer Science & Business Media. (Online ed.). Berlin: Springer. ISBN 978-3540686989.*
- [45] *Chouraki, V., Reitz, C., Maury, F., Bis, J. C., Bellenguez, C., Yu, L., ... & Larson, E. B. (2016). Evaluation of a genetic risk score to improve risk prediction for Alzheimer's disease. Journal of Alzheimer's Disease, 53(3), 921-932.*
- [46] Aedo, S., Pavlov, S., & Clavero, F. (2010). Riesgo relativo y Odds ratio¿ Qué son y cómo se interpretan. *Rev Obstet Ginecol*, 5(1), 51-4.
- [47] *Tolles, J., & Meurer, W. J. (2016). Logistic regression: relating patient characteristics to outcomes. Jama, 316(5), 533-534.*

- [48] Reche, J. L. C. (2013). Regresión logística. Tratamiento computacional con R.
- [49] Gómez, G., & Langohr, K. (2004). Análisis de supervivencia. Barcelona, España.
- [50]http://rstudio-pubs-static.s3.amazonaws.com/316989_83cbe556125645b698c9ff6cf88c4c1a.html#4_modelo_de_riesgo_proporcionales_de_cox
- [51] Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- [52] Therneau, T. M., & Lumley, T. (2015). Package 'survival'. *R Top Doc*, 128, 112.
- [53] Wright, B. L., Lai, J. T., & Sinclair, A. J. (2012). Cerebrospinal fluid and lumbar puncture: a practical review. *Journal of neurology*, 259(8), 1530-1545.
- [54] Shlens, J. (2014). A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*.

8. Anexos

Debido a la extensión de los modelos planteados, se adjunta el documento “Anexo 1”. Es un archivo Rmarkdown que recoge de forma detallada la creación de la tabla de partida, los modelos logísticos y los modelos de Cox.