

**MÁSTER UNIVERSITARIO EN I. A. AVANZADA:
FUNDAMENTOS, MÉTODOS Y APLICACIONES**

TRABAJO DE FIN DE MÁSTER

*Aplicación de algoritmos de minería de datos sobre
estudios de RS-fMRI*

Presentado por:

D^a CARMEN DE LA TORRE PALOMO

Dirigido por:

Dr. MARIANO RINCÓN

Curso Académico 2013 – 2014

D. Mariano Rincón

CERTIFICA:

Que el trabajo titulado: "*Aplicación de algoritmos de minería de datos sobre estudios de RS-fMRI*", ha sido realizado bajo mi dirección dentro del Máster Universitario en I. A. Avanzada: Fundamentos, Métodos y Aplicaciones por la alumna D^a Carmen de la Torre Palomo, y cuenta con mi aprobación para ser defendido ante la comisión evaluadora de los Trabajos de Fin de Máster nombrada al efecto.

Fecha 07 de septiembre de 2014.

Firmado:.....

ÍNDICE

1.- INTRODUCCIÓN.....	9
1.1.- Identificación del problema.....	9
1.2.- Motivación.....	9
1.3.- Objetivos.....	10
1.4.- Metodología.....	11
1.5.- Principales aportaciones.....	13
1.6.- Organización del documento.....	14
2.- ANTECEDENTES.....	16
2.1.- Resting State.....	16
2.2.- La señal BOLD	17
2.3.- Conectividad funcional en reposo.....	18
2.4.- Métodos de análisis de imágenes fMRI.....	20
2.5.- Relación entre la conectividad funcional y estructural.....	21
2.6.- Estudios previos.....	23
3.- OBJETIVOS.....	24
4.- MATERIAL Y MÉTODOS.....	26
4.1.- Descripción de la muestra	26
4.2.- Selección del método de obtención de los mapas de conectividad funcional.....	28
4.2.1.- Selección de parámetros	30
4.2.1.1.- Selección de las semillas.....	31
4.2.1.2.- Radio del ROI.....	33
4.3.- Selección de parámetros en REST.....	33
4.3.1.- Tamaño mínimo del cluster.....	34
4.3.2.- Límite.....	35

4.4.- Selección de las variables a considerar.....	39
4.5.- Aplicación de técnicas de minería de datos.....	46
4.5.1.- Preprocesado del conjunto de datos.....	47
4.5.2.- Reducción de la dimensionalidad.....	47
4.5.3.- Algoritmo de agrupación.....	49
5.- RESULTADOS.....	50
6.- DISCUSIÓN.....	54
7.- CONCLUSIÓN.....	58
8.- ANEXOS.....	59
8.1.- Métodos de análisis de imágenes fMRI.....	59
8.2.- Comparativa métodos de análisis de imágenes fMRI.....	60
9.- BIBLIOGRAFÍA.....	61
9.1.- Fuentes electrónicas.....	66

GLOSARIO DE ABREVIATURAS

ACC	Anterior Cingulate Cortex
ALFF	Amplitude of Low Frequency Fluctuation
BOLD	Blood Oxygenation Level Dependent
DCL	Deterioro Cognitivo Leve
DICOM	Digital Imaging and Communications in Medicine
DMN	Default Mode Network
DPARSF	Data Processing Assistant for Resting-State fMRI
DTI	Diffusion Tensor Imaging
EA	Enfermedad de Alzheimer
fMRI	Functional Magnetic Resonance Imaging
HC	Hippocampus
ICA	Independent Component Analysis
IPL	Inferior Parietal Lobule
MNI	Montreal Neurological Institute
mPFC	Medial PreFrontal Cortex
NIFTI	Neuroimaging Informatics Technology Initiative
PCA	Análisis de Componentes Principales
PCC	Posterior Cingulate Cortex
QSM	Quejas Subjetivas de Memoria
ReHo	Regional Homogeneity
REST	Resting-State fMRI Data Analysis Toolkit
R-fMRI	Resting state Functional Magnetic Resonance Imaging
RM	Resonancia Magnética
ROI	Region Of Interest
RSN	Resting State Network
SPM	Statistical Parametric Mapping
TNN	Task-Negative Network

ÍNDICE DE TABLAS

Tabla 1	Ventajas e inconvenientes de la técnica fMRI.
Tabla 2	Regiones consideradas como parte de la DMN.
Tabla 3	Parámetros de adquisición de la serie funcional.
Tabla 4	Pacientes incluidos en la muestra.
Tabla 5	Semillas de la región PCC utilizadas en DPARSF para obtener los mapas z de conectividad.
Tabla 6	Semillas de la región precuneus utilizadas en DPARSF para obtener los mapas z de conectividad.
Tabla 7	Semillas de la región mPFC utilizadas en DPARSF para obtener los mapas z de conectividad.
Tabla 8	Semillas de la región IPL utilizadas en DPARSF para obtener los mapas z de conectividad.
Tabla 9	Semillas de la región HC utilizadas en DPARSF para obtener los mapas z de conectividad.
Tabla 10	Semillas de la región ACC utilizadas en DPARSF para obtener los mapas z de conectividad.
Tabla 11	VARIABLES extraídas por cada paciente para la selección de la mejor semilla.
Tabla 12	Mejores semillas de cada una de las regiones de la DMN.
Tabla 13	Representación del histograma del mapa z de conectividad funcional generado por la mejor semilla de cada región y su aproximación por una función gaussiana (paciente 17).
Tabla 14	Datos del escenario 2 del paciente 17 cuando se calcula el valor z de los voxeles que se superponen con la media.
Tabla 15	Datos del escenario 2 del paciente 17 cuando se calcula el valor z de los voxeles que se superponen con la mediana.

Tabla 16	Selección de atributos empleando diferentes métodos de búsqueda en el escenario 1.
Tabla 17	Selección de atributos empleando diferentes métodos de búsqueda en el escenario 2.
Tabla 18	Enumeración y descripción de los atributos seleccionados en cada uno de los escenarios.
Tabla 19	Pacientes igualmente agrupados utilizando los mapas de difusión, independientemente del número de variables del mapa y del escenario.
Tabla 20	Comparación de los resultados obtenidos utilizando los mapas de difusión y la selección de variables.
Tabla 21	Análisis para la identificación de los grupos.

ÍNDICE DE FIGURAS

- Figura 1 Prevalencia de los distintos tipos de demencia.
- Figura 2 Esquema de las distintas fases del estudio.
- Figura 3 Flujo sanguíneo en los estados de reposo y activación.
- Figura 4 Conexiones estructurales existentes entre algunas de las regiones de la DMN.
- Figura 5 Visualización de un mapa z de conectividad funcional en REST.
- Figura 6 Corrección del límite realizada por AlphaSim en función del núcleo gaussiano empleado para el suavizado y el criterio de conectividad considerado.
- Figura 7 Mediana del número de clusters por semilla del PCC en función del límite.
- Figura 8 Mediana del número de clusters por semilla del precuneus en función del límite.
- Figura 9 Mediana del número de clusters por semilla del mPFC en función del límite.
- Figura 10 Mediana del número de clusters por semilla del IPL en función del límite.
- Figura 11 Mediana del número de clusters por semilla del HC en función del límite.
- Figura 12 Mediana del número de clusters por semilla del ACC en función del límite.
- Figuras 13 y 14 Media y mediana de las regiones conectadas con las semillas del PCC y precuneus respectivamente.
- Figuras 15 y 16 Media y mediana de las regiones conectadas con las semillas del mPFC e IPL respectivamente.
- Figuras 17 y 18 Media y mediana de las regiones conectadas con las semillas del HC y ACC respectivamente.
- Figura 19 Representación parcial del fichero de datos introducido en R en el primer escenario.
- Figura 20 Representación parcial del fichero de datos introducido en R en el segundo escenario.
- Figura 21 Agrupación de pacientes aplicando la selección de atributos.
- Figura 22 Agrupación de pacientes aplicando la transformación de atributos con mapas de difusión (2 a 6 variables).

1.- INTRODUCCIÓN

Con este apartado, se intenta que el lector adquiriera una visión global del estudio realizado. En él, se utiliza una técnica denominada resonancia magnética funcional (fMRI) que permite explorar la redes funcionales del cerebro humano con una resolución relativamente alta. Esta técnica se puede emplear en reposo, en cuyo caso hablamos de resonancia magnética funcional en reposo (R-fMRI) o mientras el paciente realiza un conjunto de tareas, previamente planificadas (paradigmas), durante el proceso de adquisición de las imágenes.

1.1.- Identificación del problema

El diagnóstico precoz de desórdenes neuronales relacionados con la demencia es fundamental para el tratamiento y la calidad de vida de los pacientes que desarrollen la Enfermedad de Alzheimer (EA en adelante). Hasta ahora, los estudios se han centrado en la detección de pacientes con Deterioro Cognitivo Leve (DCL en adelante). Sin embargo, en el presente trabajo se intenta adelantar aún más el proceso hasta el punto de poder reconocer pacientes con Quejas Subjetivas de Memoria (QSM en adelante).

1.2.- Motivación

El envejecimiento de la población, entendido como el aumento de la población con más de 60 años, acompañado por una reducción en el número de niños menores de 15 años, ha hecho que la comunidad científico médica dedique más recursos y esfuerzos al estudio de las enfermedades relacionadas con la edad, como es la demencia. Según la Fundación del cerebro, esta enfermedad supone el deterioro adquirido de las funciones intelectuales (memoria, capacidad para expresarse y comprender, sentido de la orientación, capacidad para razonar, etc) que impide llevar a cabo de manera autónoma y adecuada las actividades de la vida diaria. Según las deficiones de demancia y Enfermedad de Alzheimer extraídas de Wikipedia (2014), esta dolencia es un tipo de demencia que se manifiesta por un deterioro cognitivo acompañado de trastornos conductuales. También se caracteriza por una pérdida de la memoria inmediata y de otras capacidades mentales, a medida que las células nerviosas mueren y diferentes zonas del cerebro se atrofian. La EA es la forma más

común de demencia, incurable y terminal.

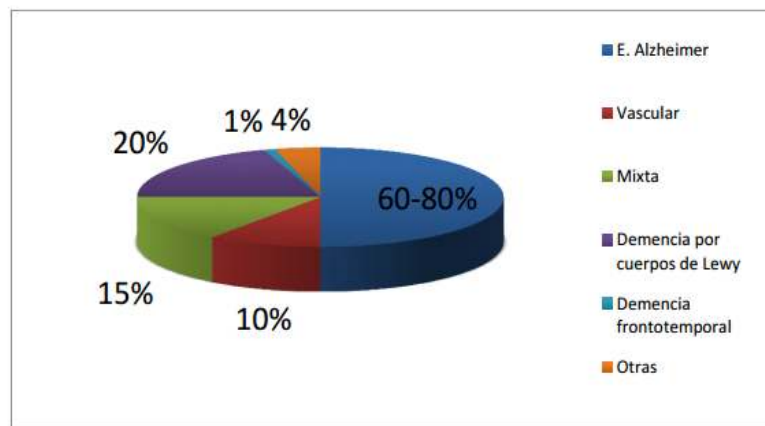


Figura 1: Prevalencia de los distintos tipos de demencia.

Fuente: Prieto et al., (2011). Prevalencia de los distintos tipos de demencia.

Un estadio intermedio entre el envejecimiento normal y la demencia, es lo que se denomina DCL. Los pacientes aquejados de la misma, presentan pérdidas de memoria y alteración en la realización de sus actividades habituales. En el presente estudio, se intenta adelantar aún más la detección precoz de éstos desórdenes, incluyendo pacientes que han sido diagnósticos con QSM. También se han considerado pacientes sanos (controles) y pacientes con DCL.

La demencia ocupa el tercer puesto de las enfermedades más costosas, por detrás de las enfermedades cardiovasculares y el cáncer, según Prieto et al., 2011. Aunque no existe en la actualidad curación para la EA, existen tratamientos que mejoran los síntomas y pueden estabilizar la progresión de la enfermedad. Por tanto, un diagnóstico temprano contribuye a la mejora de la calidad de vida del paciente y de sus cuidadores, reduciendo además los costos asistenciales asociados.

1.3.- Objetivos

- ▶ Establecer una herramienta de ayuda al diagnóstico para detectar precozmente pacientes con QSM y DCL en función del análisis estadístico de sus estudios de R-fMRI.
- ▶ Análisis de los métodos de estudio de fMRI.
- ▶ Selección de los parámetros más óptimos para el análisis de la conectividad .

- ▶ Determinación de las variables a medir.
- ▶ Planteamiento de posibles formas de abordar el problema.
- ▶ Análisis y selección de los algoritmos de minería de datos adecuados al problema.
- ▶ Aplicación de herramientas de minería de datos al conjunto de variables consideradas en los escenarios planteados.

1.4.- Metodología

Cuando el cerebro no está realizando ninguna tarea específica, se sigue detectando actividad cerebral. Un gran número de estudios previos mencionan la existencia de regiones del cerebro conectadas funcionalmente en reposo, dando lugar a las denominadas redes en reposo (Beckmann et al., 2005; Biswal et al., 1995; Damoiseaux et al., 2006; De Luca et al., 2006; Salvador et al., 2005a; Van Den Heuvel et al., 2010). La más importante de ellas, y que se analizará en el presente trabajo, es la denominada DMN (Default Mode Network). Según Hafkemeijer et al., 2012, esta red está constituida por la corteza cingulada posterior (PCC), el precuneus, la corteza media prefrontal (mPFC), los lóbulos parietales inferiores (IPL), la corteza cingulada anterior (ACC) y el hipocampo (HC). Los pacientes con DCL presentan alteraciones en la conectividad de las regiones que componen la DMN (Gili et al., 2011 y Hafkemeijer et al., 2012).

En el presente estudio, partimos de una muestra de 30 pacientes constituida por personas con DCL, QSM y personas mayores sin ninguna alteración detectada, que constituyen el grupo de control. No se dispone del diagnóstico de cada uno de estos pacientes, por lo que se debe recurrir a alguna técnica de aprendizaje no supervisado para establecer agrupaciones de pacientes que compartan alguna característica o diagnóstico clínico. A cada uno de los pacientes, se le realiza una R-fMRI. Con la toolbox de Matlab llamada DPARSF v3.0 (Data Processing Assistant for Resting-State fMRI) e introduciendo las semillas y los parámetros adecuados, se obtienen los mapas z de conectividad funcional de cada uno de los pacientes. Estos mapas se procesan con otra toolbox también de Matlab denominada REST v1.8 (Resting-State fMRI Data Analysis Toolkit). Esta herramienta, en su forma original, no nos proporciona las variables que se desean explotar. Con el conocimiento del dominio que disponemos, y en base a la bibliografía consultada, centraremos

en valorar la conectividad funcional de cada región de la DMN con todas las demás regiones que componen esta red con el objetivo de establecer la agrupación de pacientes deseada. Se ha generado una toolbox, en base al visor de REST, que nos permite extraer esta información de la forma más eficiente posible. En este punto, se plantean dos formas distintas de abordar el problema:

- Escenario 1: en base a los resultados obtenidos con todas las semillas de una región se valora, empleando la mediana, qué semilla muestra una conectividad mayor entre las regiones que componen la DMN en el conjunto de pacientes. Esta selección, de la que denominamos como "mejor semilla", se repite para todas las regiones consideradas (PCC, Precuneus, mPFC, IPL, HC y ACC).

- Escenario 2: se realiza un análisis acumulado, es decir, para valorar la conectividad entre las regiones de la DMN, se tienen en cuenta los resultados obtenidos con todas las semillas de una misma región.

Una vez realizadas las modificaciones en el código de la toolbox REST en función de los dos escenarios expuestos, se generan dos ficheros con las variables que se van a medir en cada uno de los escenarios en el formato adecuado para su utilización en el software R. Con esta aplicación y con el software WEKA v3.6.10 realizamos un preprocesado de los datos, una reducción de la dimensionalidad del problema, y después aplicamos una técnica de agrupamiento o clustering, que nos permita agrupar a los pacientes en subgrupos. A priori, no se dispone del diagnóstico de los pacientes (clase), por lo que se debe recurrir a una técnica de aprendizaje no supervisado.

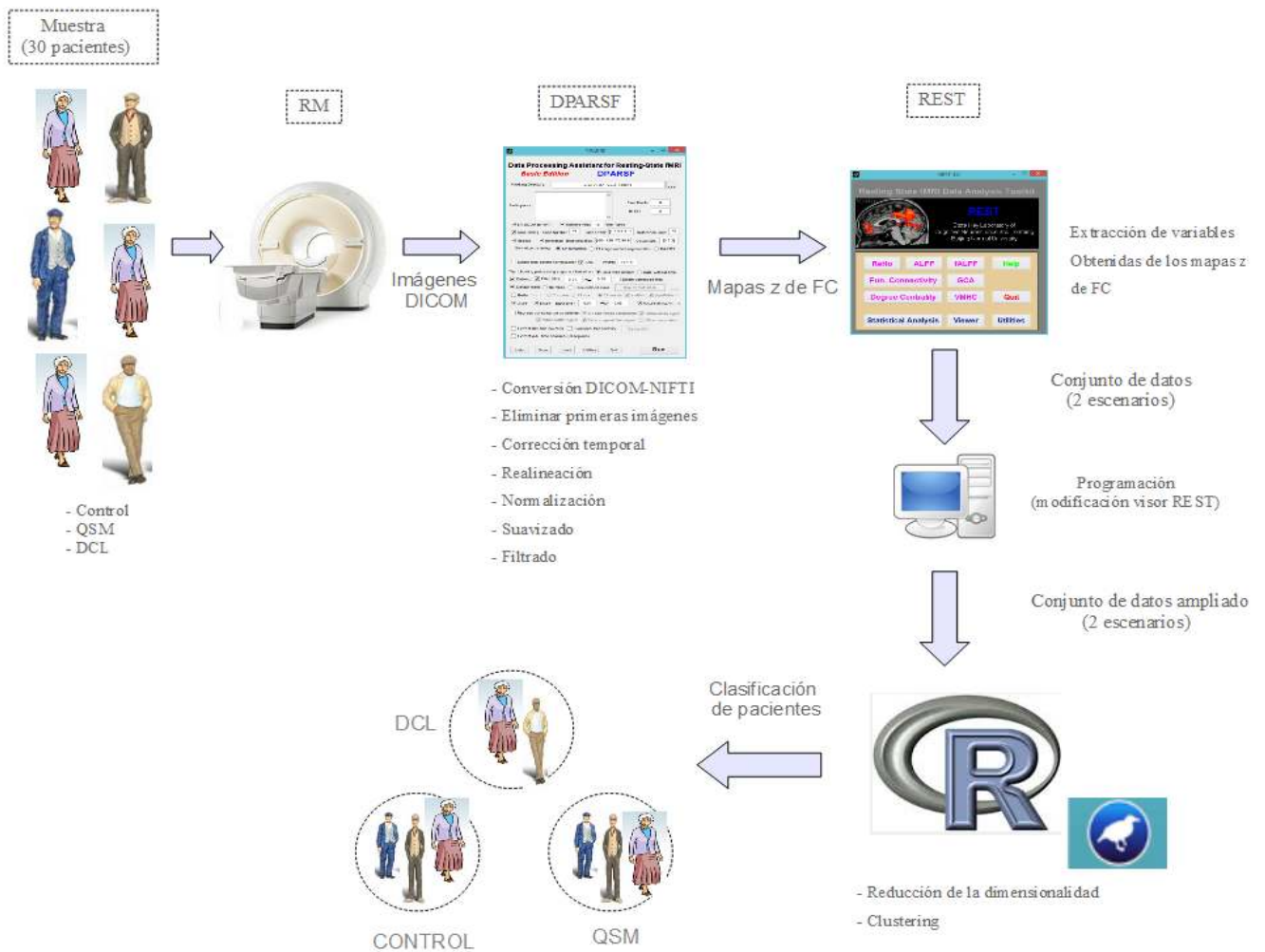


Figura 2: Esquema de las distintas fases del estudio.

1.5.- Principales aportaciones

Este estudio abre un campo de investigación poco desarrollado hasta el momento, la aplicación de algoritmos de minería de datos a estudios de R-fMRI. Además, la mayor parte de los estudios previos de R-fMRI realizan análisis estadísticos de grupos, mientras que en el presente trabajo se intenta dar un paso más allá, generando una herramienta de ayuda al diagnóstico precoz de desórdenes neuronales relacionados con la edad. Para ello, ha sido imprescindible determinar qué variables nos pueden ayudar a establecer agrupaciones de pacientes. Consideramos que el estudio de la conectividad entre las regiones de la DMN es la clave que nos permitirá identificar grupos de pacientes potencialmente ligados a los distintos estados de una enfermedad

neurodegenerativa. En este sentido, también se ha intentado introducir un punto innovador al incluir en nuestra muestra, no solo pacientes con DCL, sino también con QSM.

A pesar de no disponer del diagnóstico de los pacientes, se han analizado los resultados obtenidos encontrando evidencias de que el agrupamiento realizado podría corresponderse, sino totalmente, por lo menos parcialmente con la realidad. Se ha obtenido un grupo más numeroso, que es el que presenta una mayor conectividad funcional entre las regiones de la DMN, que podría ser considerado el grupo de control. Por otro lado, el grupo más reducido, que además presenta una conectividad menor entre las regiones mencionadas, se podría suponer que incluye a los pacientes con DCL. Los pacientes restantes, serían los aquejados de QSM.

1.6.- Organización del documento

A continuación, se describe brevemente el contenido de los próximos apartados:

- **Antecedentes:** en este apartado se explica el concepto de "Resting-State", el fundamento de la fMRI, la señal BOLD (Blood Oxygenation Level Dependent) y se comenta la existencia de las denominadas redes en reposo, focalizándonos en la DMN. Se enumeran los cambios que sufre, según la bibliografía, la DMN de los pacientes que sufren DCL. Brevemente, se exponen las dos grandes categorías de métodos de análisis de fMRI, los métodos basados en modelos y los basados en datos. Los anexos I y II completan la información sobre este punto. Finalmente, se describe la relación existente entre la conectividad funcional y estructural de las regiones cerebrales.
- **Objetivos:** se describe el objetivo del presente trabajo y los distintos hitos intermedios para su consecución.
- **Material y métodos:** apartado bastante amplio en el que se describe la muestra empleada así como el funcionamiento de las herramientas utilizadas en cada una de las fases del estudio y los parámetros de cada una de ellas debidamente justificados. Las herramientas consideradas han sido:
 - DPARSF para la obtención de los mapas z de conectividad funcional a partir de los estudios de R-fMRI.

- REST modificado, para la extracción de las variables a considerar en cada uno de los dos escenarios. Estas variables serán obtenidas a partir de los mapas z de conectividad funcional generados con DPARSF.
 - R, software estadístico que permite aplicar múltiples algoritmos de minería de datos.
 - WEKA, que dispone de múltiples algoritmos de aprendizaje automático aplicables en minería de datos.
-
- Resultados: En este apartado se incluirán las simulaciones realizadas, comparando los resultados obtenidos aplicando distintas técnicas de reducción de la dimensionalidad y comparando también los resultados de los dos escenarios planteados.
 - Discusión: se justificarán los resultados obtenidos en el apartado anterior.
 - Conclusiones: una vez expuestos y justificados los resultados del presente estudio, en este apartado se incluirán algunas recomendación para estudios futuros y se plantearán posibles líneas de investigación.
 - Anexos: En este apartado se incluye un esquema que muestra la clasificación de los métodos de estudio de fMRI así como una tabla comparativa entre varios de los métodos.

2.- ANTECEDENTES

La resonancia magnética funcional (fMRI), es una técnica que estudia la función del cerebro de forma no invasiva y supone una potente herramienta utilizada desde principios de los noventa, tanto en el entorno clínico como en el de la investigación. La técnica más popular utiliza el contraste entre los distintos niveles de oxigenación de la sangre, BOLD (Blood Oxygenation Level Dependent). En la siguiente tabla, se muestran las ventajas e inconvenientes de la fMRI:

VENTAJAS	INCONVENIENTES
Permite obtener una alta resolución espacial y temporal	Sensible a los movimientos de la cabeza
Sin radioactividad	Desfase de entre 3 y 6 segundos desde que la región del cerebro se activa y el flujo de sangre se incrementa
No invasiva	Ruido del imán
Reproducible	Claustrofobia en RM cerradas

Tabla 1: Ventajas e inconvenientes de la técnica fMRI.

2.1.- Resting State

Cuando el cerebro no está realizando ninguna tarea específica, se sigue detectando actividad cerebral debido a las reminiscencias autobiográficas, imágenes mentales, planes de futuro, etc, que ocupan nuestros pensamientos. A este estado mental se le denomina "**resting-state**" en inglés, estado de reposo consciente, pensamiento independiente del estímulo o función cerebral en modo por defecto en español (Greicuis et al., 2002).

En el contexto de la experimentación, "rest" es una definición operacional referida a una condición constante en la que no se impone ningún estímulo o evento conductista notable. Los ojos pueden estar abiertos o cerrados, con o sin fijación. El objetivo de los experimentos en resting state es capturar las propiedades estadísticas de la actividad neuronal generada. Por el contrario, el objetivo de los estudios relacionados con eventos es medir respuestas invocadas o inducidas (Snyder et al., 2012).

Inicialmente, hubo una discusión sobre cómo se debía llamar a este fenómeno. La mayoría de los autores describieron el estado "OFF" como "rest", "línea base" o "condición de control". También aparecieron los términos "intrínseco" y "espontáneo", pero finalmente prevaleció el

término "resting-state". Debido a que el Dr. Karl Friston había empleado el término "conectividad funcional" para describir la correlación temporal entre regiones del cerebro físicamente distantes, finalmente el fenómeno tomó el nombre de "resting-state functional connectivity" en inglés o "conectividad funcional en reposo" en español (Biswal, 2012).

2.2.- La señal BOLD

Normalmente, los núcleos de los átomos están orientados aleatoriamente, pero bajo la influencia de un campo magnético los núcleos se alinean en la dirección del mismo. De tal forma que, cuando mayor es el campo magnético aplicado, mayor es el alineamiento producido. Cuando los núcleos están apuntando en la misma dirección, las señales magnéticas de los núcleos individuales se suman, dando lugar a una señal que es lo suficientemente significativa como para ser medida. En fMRI lo que se detecta es la señal magnética procedente de los núcleos de hidrógeno del agua (H_2O). La magnitud de la señal medida proporciona información sobre el tipo de tejido que estamos analizando, y nos permite distinguir entre materia gris, blanca o líquido espinal cerebral en imágenes de RM (Devlin, 2007).

El oxígeno se proporciona a las neuronas a través de la hemoglobina. Cuando la actividad neuronal se incrementa, crece la demanda de oxígeno y por tanto se incrementa el flujo sanguíneo en la zona de activación neuronal (Devlin, 2007).

La hemoglobina es diamagnética cuando está oxigenada y paramagnética cuando está desoxigenada. Esta diferencia en las propiedades magnéticas genera la señal BOLD que permite detectar las zonas con diferente oxigenación en las imágenes de fMRI (Devlin, 2007).

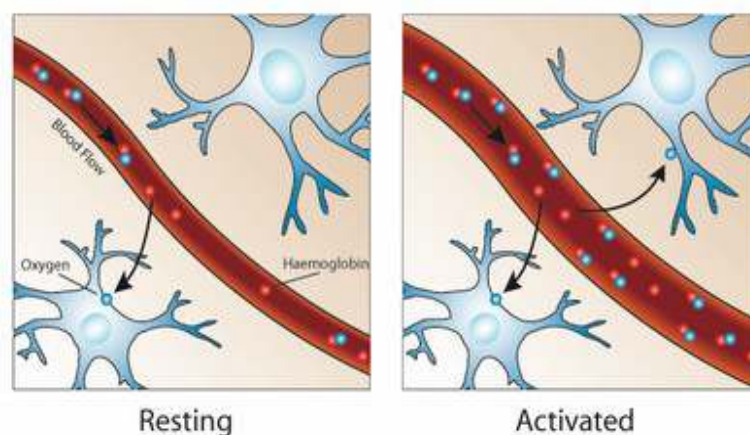


Figura 3: Flujo sanguíneo en los estados de reposo y activación.

Fuente: Devlin, (2007).

Al activarse las neuronas, se incrementa el flujo de sangre y la oxigenación, lo que provoca un cambio (decaimiento) del parámetro T2* de la resonancia magnética. Estos cambios en el flujo de sangre y en la oxigenación (vascular o hemodinámica) se retrasan temporalmente debido al encendido neuronal, factor conocido como retraso hemodinámico (Calhoun et al., 2003). Seguidamente, el flujo de sangre se incrementa por encima del nivel de oxígeno demandado. La oxigenación de la sangre se incrementa en función de la activación neuronal (Devlin, 2007).

La señal BOLD muestra fluctuaciones lentas espontáneas que, inicialmente, se consideró como ruido. No se contempló que estas fluctuaciones eran de origen neuronal hasta que Biswal y sus colegas demostraron que las señales BOLD en resting state están correladas temporalmente con el sistema somatomotor (Snyder et al., 2012). Aplicando esta técnica, se han podido identificar varias RSNs o áreas separadas espacialmente que presentan, en reposo, fluctuaciones sincrónicas de la señal BOLD (Lee et al., 2012). El significado funcional de estas fluctuaciones fue presentado por primera vez por Biswal en 1995 (Lee et al., 2012).

La señal BOLD tiene una frecuencia baja ($<0.1\text{Hz}$) con una contribución mínima de oscilaciones cardíacas y respiratorias a más altas frecuencias ($> 0.3\text{Hz}$). Por tanto, el ancho de banda de interés es de 0.01-0.1 Hz (Van Den Heuvel et al., 2010).

2.3.- Conectividad funcional en reposo

Un gran número de estudios previos mencionan la existencia de regiones del cerebro conectadas funcionalmente en reposo, dando lugar a las denominadas redes en reposo (Beckmann et al., 2005; Biswal et al., 1995; Damoiseaux et al., 2006; De Luca et al., 2006; Salvador et al., 2005a; Van Den Heuvel et al., 2010). La más importante de ellas, y que se analizará en el presente trabajo, es la denominada red del modo por defecto. Esta red, denominada también Default Mode Network (DMN) o Task-Negative Network (TNN) es la más conocida. En contraste con otras RSNs, las regiones de la DMN muestran un elevado nivel de actividad neuronal cuando el sujeto no está realizando ninguna tarea específica, lo que sugiere que la actividad de esta red refleja un estado por defecto de actividad neuronal en el cerebro humano (Calhoun et al., 2003). Cuando la DMN está activa, el sujeto está recordando, pensando sobre su futuro, sobre el mundo que le rodea, etc. Aunque la red del modo por defecto está más activa durante el reposo, también lo puede estar durante la realización de tareas con limitada demanda cognitiva, no así en tareas que requieran una demanda mayor.

Respecto a las regiones que componen la DMN, autores como Beckmann et al., 2005 y Damoiseaux et al., 2006 consideran que está constituida por la corteza cingulada posterior (posterior cingulate cortex, en inglés), la corteza media prefrontal (medial prefrontal cortex, en inglés) y los lóbulos parietales inferiores (inferior parietal lobules, en inglés). Sin embargo, otros estudios como el de Hafkemeijer et al., 2012 consideran que el Precuneus, la corteza cingulada anterior (anterior cingulate cortex, en inglés) y el hipocampo (Hippocampus, en inglés) también forman parte de esta red. En el presente estudio se ha considerado la última clasificación mencionada.

REGIONES CONSIDERADAS COMO PARTE DE LA DMN
Corteza cingulada posterior
Precuneus
Corteza media prefrontal
Lóbulos parietales inferiores
Corteza cingulada anterior
Hipocampo

Tabla 2: Regiones consideradas como parte de la DMN.

A partir de ahora, la mención a las regiones incluidas en la tabla 2 se realizará en base a sus acrónimos en inglés, incluidos en el correspondiente glosario.

Con el envejecimiento normal, se reducen tanto la conectividad funcional entre las distintas regiones que componen la DMN como la desactivación de la DMN ante tareas con demanda cognitiva. Este decremento es mayor en pacientes con DCL y más severo en pacientes con EA (Hafkemeijer et al., 2012).

Según Wang et al., 2012, el decremento de la conectividad funcional entre las regiones que componen la DMN en pacientes con EA está ampliamente confirmado en multitud de estudios previos como son los de Wang et al. 2006, Zhang et al. 2009, Greicius et al 2004, Agosta et al 2012 y Binnewijzend et al. 2012. Sin embargo, en el caso de los pacientes con DCL hay cierta controversia entre los resultados obtenidos en los estudios previos. Según este mismo autor (Wang et al., 2012) mientras que en algunos de ellos (Bai et al., 2009; Sorg et al., 2007; Qi et al., 2010) se observan cambios en la conectividad entre las regiones de la DMN, en otros no (Rombouts et al., 2009).

El diagnóstico de pacientes con DCL es sumamente importante ya que, aproximadamente la mitad de estos pacientes, padecerán la EA en un plazo de 3 a 5 años (Hafkemeijer et al., 2012).

A los pacientes implicados en el presente estudio, como se ha mencionado anteriormente, no se le aplica ningún paradigma, por lo que no se puede observar el nivel de desactivación de la DMN cuando los pacientes están realizando alguna tarea que requiera suficiente atención. Respecto a los cambios en la conectividad funcional entre las regiones de la DMN, Gili et al., 2011 y Hafkemeijer et al., 2012, al igual que Wang et al., 2012, afirman que los pacientes con DCL presentan alteraciones en la conectividad de DMN. Los cambios comentados en la bibliografía son los siguientes:

- La conectividad funcional entre el HC y el PCC, presente en los pacientes sanos, no se observa en pacientes con DCL (Sorg et al., 2007).
- Se reduce la conectividad entre las siguientes regiones de la DMN (Gili et al., 2011):
 - PCC y ACC bilateralmente.
 - mPFC y PCC unilateralmente.
 - mPFC y ACC bilateralmente.

La reducción de la conectividad es bilateral cuando ese efecto se produce en los dos sentidos, es decir, en el caso de la disminución de la conectividad bilateral entre el PCC y el ACC, disminuye tanto la conexión del PCC con el ACC, como la del ACC con el PCC. Si solo se produce en un sentido, se considera unilateral. Para poder contrastar estas afirmaciones al final de nuestro estudio, será necesario aplicar una técnica de estudio de fMRI basada en semilla, y seleccionar la semilla del area que se desea saber si está conectada con otra o no.

2.4.- Métodos de análisis de imágenes fMRI

Se han propuesto muchos métodos para el estudio de la fMRI. Estos métodos se pueden clasificar, según Li et al., 2009, en dos categorías:

- Métodos basados en modelos.
- Métodos basados en datos.

Estos últimos tienen una gran variedad de subcategorías que a su vez se dividen en:

- Métodos basados en descomposición
- Métodos de análisis por agrupamiento (clustering)

A continuación, se expone una breve descripción de cada una de las categorías y en el anexo 1 y 2 se puede encontrar un esquema de los mismos, así como una tabla comparativa con las ventajas e inconvenientes de algunos de estos métodos más relevantes.

a) Métodos basados en modelos

La forma más directa de examinar las conexiones funcionales de una region particular del cerebro es correlar la serie temporal en reposo de la zona de interés con la serie temporal del resto de regiones dando lugar a un mapa de conectividad (fcMap) que define las conexiones funcionales de la región cerebral predefinida. A esta región de interés se le denomina semilla (Van Den Heuvel et al., 2010).

b) Métodos basados en datos

Este tipo de métodos intenta superar las limitaciones del anterior, por ejemplo, estos métodos son independientes de la información conocida a priori o de los modelos que se hayan desarrollado. Existen dos grandes ramas de métodos, los que intentan expresar el conjunto de datos fMRI como una combinación lineal de vectores básicos (PCA/SVD) o como componentes independientes estadísticamente (ICA) y los métodos por agrupamiento tales como el agrupamiento difuso (FCA) o el jerárquico (HCA) (Li et al., 2009).

2.5.- Relación entre la conectividad funcional y estructural

Los caminos de sustancia blanca son conexiones estructurales de nuestro cerebro y se pueden visualizar empleando la tractografía o Diffusion Tensor Imaging (DTI) (Van Den Heuvel et al., 2009). Ésta es una técnica de resonancia que permite reconstruir el tracto de materia blanca en el cerebro. Se mide el perfil de difusión de moléculas libres de agua en el tejido cerebral, las cual son

conocidas por expandirse a lo largo de una dirección determinada en el tracto de materia blanca debido a la composición compacta de las colas de los axones (Van Den Heuvel et al., 2010).

Recientemente, un gran número de estudios han sugerido la existencia de una relación directa entre la conectividad funcional y estructural en el cerebro humano combinando estudios de fMRI en reposo y de tractografía (Van Den Heuvel et al., 2010). Concretamente, en el estudio realizado por Heuvel et al. 2009, se muestra la existencia de interconexiones de tracto de materia blanca en ocho de las nueve redes funcionales en reposo consideradas en el estudio, en una proporción considerable de los pacientes incluidos en la muestra. A continuación, se incluye un esquema con la conexiones estructurales presentes entre algunas regiones de la DMN (Van Den Heuvel et al., 2009):

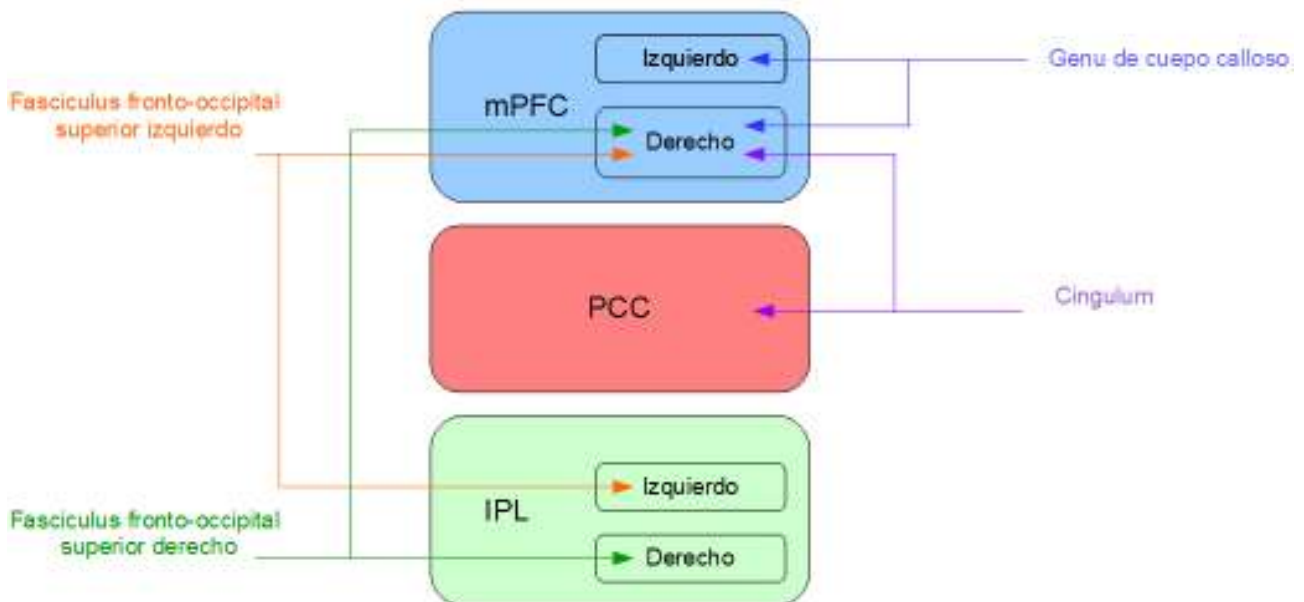


Figura 4: Conexiones estructurales existentes entre algunas de las regiones de la DMN.

Aunque la correlación se emplea como una medida de la conectividad funcional entre dos regiones, no está claro si esta conectividad funcional refleja una comunicación directa entre estas regiones o si éstas están relacionadas indirectamente a través de una tercera región. Estudios realizados demuestran que una gran cantidad de las regiones relacionadas funcionalmente están directamente interconectadas estructuralmente (Van Den Heuvel et al., 2009).

Sin embargo, la conectividad funcional muestra propiedades dinámicas que no se encuentran en la conectividad estructural. Por ejemplo, la conectividad funcional está modulada por el estado emocional y cognitivo, por la excitación sexual y el sueño, mientras que la conectividad estructural no se ve afectada por estos factores (Biswal et al., 2010).

Los estudios realizados sugieren que solamente ciertas partes de la región considerada están estructuralmente interconectadas. Esto podría indicar que RSNs relacionadas funcionalmente están estructuralmente interconectadas de una manera jerárquica, lo que significaría que solo una subparte de la región de la RSN está directa y estructuralmente conectada a otra región de la red y que las otras subpartes de la región están conectadas a este hub de conectividad estructural. Cabe mencionar que los hallazgos obtenidos hasta el momento no aportan una evidencia directa sobre ello (Van Den Heuvel et al., 2009).

2.6.- Estudios previos

Los estudios previos consultados relacionados con los cambios en la DMN en pacientes con DCL realizan estudios de grupos. Estos estudios muestran, por ejemplo, las diferencias en la conectividad funcional y estructural entre dos subtipos de pacientes con DCL (Li et al., 2014), o comparan la conectividad funcional de la DMN en pacientes con DCL y pacientes sanos (Yan et al., 2013).

En nuestro caso, no se ha realizado un estudio de grupos, sino que se ha definido una serie de variables a estudiar por paciente en función de las cuales, se intenta establecer grupos de pacientes utilizando un algoritmo de clustering del campo de la minería de datos. Este término no se debe confundir con los métodos de clustering disponibles para el tratamiento de estudios fMRI.

Esta herramienta de ayuda al diagnóstico que se pretende generar, sería un gran apoyo para los clínicos a la hora de diagnosticar si un paciente sufre DCL o no, sin más coste que la fMRI realizada y sin más inconvenientes para el paciente que la realización de una serie adicional (la serie funcional) en su estudio de RMI. Conviene recordar que la fMRI es un método no invasivo, reproducible, que no supone ninguna radiación para el paciente y que proporciona una resolución espacial y temporal alta.

3.- OBJETIVOS

► Establecer una herramienta de ayuda al diagnóstico para detectar precozmente pacientes con QSM y DCL en función del análisis estadístico de sus estudios de R-fMRI.

► Análisis de los métodos de estudio de R-fMRI. Uno de los pasos previos para conseguir el objetivo final, fue decidir qué método emplear para el análisis de los estudios de R-fMRI de los pacientes. Se barajaron dos opciones, un método ICA implementado con la toolbox de Matlab llamada GIFT (Calhoun, 2004) y un método basado en semilla, implementado también con una toolbox de Matlab y denominado DPARSF. Finalmente, se utilizó DPARSF ya que se quiere medir la conectividad funcional entre cada área de la DMN con las demás áreas de esa misma red. Además, en el presente estudio sólo se va a considerar la DMN. Si empleáramos algún método ICA, tendríamos el mapa z de conectividad funcional de todas las redes existentes en reposo y, además, sería necesario identificar visualmente las componentes obtenidas, clasificándolas como parte de cada una de las redes en reposo o considerándolas ruido. El inconveniente principal de los métodos basados en semilla es tener que seleccionarlas, pero este impedimento se superó consultando la bibliografía existente al respecto. Como se puede observar en las tablas del apartado 4.2.1.1., todas las semillas empleadas en este estudio lo han sido antes en otros.

Se seleccionó REST como herramienta para la extracción de las variables de cada uno de los escenarios porque permitía extraer gran parte de las variables que se necesitaban. Al modificar el código, se consiguió obtenerlas todas.

► Selección de los parámetros más óptimos para el análisis de la conectividad. Para determinar qué parámetros utilizar, como son las semillas, el radio de las mismas y el tamaño mínimo del cluster se consultó la bibliografía relacionada.

► Determinación de las variables a medir. Esta selección, se realizó teniendo en cuenta que multitud de estudios avalan la existencia de cambios en la conectividad de la DMN en pacientes con DCL (Gili et al., 2011; Hafkemeijer et al., 2012; Wang et al., 2012). Por tanto, nos centramos en extraer variables que miden la existencia de conectividad entre las distintas regiones que la componen. Sin embargo, se han

introducido otras variables como el tamaño y el valor z máximo del mapa z de conectividad (en ambos escenarios), la media y la desviación típica del histograma de mapa z de conectividad (escenario 1), etc.

► Planteamiento de posibles formas de abordar el problema. Conscientes de que el mapa z obtenido con cada una de las semillas era dependiente de las mismas, se consideró conveniente realizar también un análisis acumulado, que tuviera en cuenta los resultados obtenidos con todas las semillas de cada una de las regiones y no solo una de ellas.

► Análisis, selección y aplicación de los algoritmos de minería de datos adecuados para la selección de las variables más significativas del problema. La muestra, 30 pacientes, se puede considerar pequeña respecto al número de variables que manejamos en cada uno de los escenarios, 79 en el escenario 1 y 55 en el escenario 2, por lo que debemos recurrir a técnicas de reducción de la dimensionalidad.

► Análisis, selección y aplicación de herramientas de minería de datos al conjunto de variables consideradas para obtener el agrupamiento de pacientes deseado. Al no disponer de la clase o diagnóstico de los pacientes se debe recurrir a una técnica de aprendizaje no supervisado, por lo que seleccionamos un algoritmo de clustering, en concreto K-means.

4.- MATERIAL Y MÉTODOS

Este trabajo forma parte de un proyecto de investigación iniciado por varios Hospitales de la Comunidad de Madrid sobre desórdenes neuronales relacionados con la edad. En nuestro caso, el grupo de pacientes está compuesto por personas mayores sin ningún desorden detectado (grupo de control), con quejas subjetivas de memoria y con deterioro cognitivo leve, a los que se les ha realizado una R-fMRI. Las imágenes de estos pacientes fueron anonimizadas convenientemente antes de realizar el análisis de las mismas. La única información que se conoce a priori de los pacientes es su edad y sexo.

4.1.- Descripción de la muestra

Se ha trabajado con una muestra de 30 pacientes con edades comprendidas entre los 65 y 81 años, siendo la edad media de 71,73 años. El 63,33% de los pacientes son mujeres y el 36,67% varones y fueron instruidos para no realizar ninguna tarea específica durante el tiempo de adquisición de las imágenes de R-fMRI. Para poder realizar el análisis de las mismas, se ha adquirido, por cada paciente, una serie funcional de 6.160 imágenes y los parámetros incluidos en la tabla 3 con una RM GE Signa HDxt. La plantilla de grupo empleada para representar la serie funcional se ha generado a partir de la serie T1 coronal de los pacientes de la muestra. Dichas series constan de 252 imágenes cada una.

ETIQUETA DICOM	DESCRIPCIÓN	VALOR
(0018,0080)	Repetition Time (TR)	3.500 ms (3.5s)
(0020,0105)	NumberOfTemporalPositions (Time Points)	140
(0020,1002)	ImagesInAdquisition	44

Tabla 3: Parámetros de adquisición de la serie funcional.

IDENT. IMÁGENES	IDENT. ANÁLISIS	EDAD	SEXO	FECHA ADQUISICIÓN
MNEMOS_UCD005_81V	Sub_001	81	VARON	30/11/13
MNEMOS_UCD011_76F	Sub_002	76	MUJER	23/11/13
MNEMOS_UCD015_77V	Sub_003	77	VARON	23/11/13
MNEMOS_UCD016_69F	Sub_004	69	MUJER	30/11/13
MNEMOS_UCD018_80F	Sub_005	80	MUJER	30/11/13
MNEMOS_UCD020_69V	Sub_006	69	VARON	30/11/13
MNEMOS_UCD024_72V	Sub_007	72	VARON	30/11/13
MNEMOS_UCD026_70F	Sub_008	70	MUJER	30/11/13
MNEMOS_UCD032_65F	Sub_009	65	MUJER	30/11/13
MNEMOS_UMA0016^71F	Sub_010	71	MUJER	24/11/13
MNEMOS_UMA001^72V	Sub_011	72	VARON	24/11/13
MNEMOS_UMA002^70F	Sub_012	70	MUJER	24/11/13
MNEMOS_UMA004^80V	Sub_013	80	VARON	24/11/13
MNEMOS_UMA010^69V	Sub_014	69	VARON	24/11/13
MNEMOS_UMA014^65F	Sub_015	65	MUJER	24/11/13
MNEMOS_UCD023_69V	Sub_016	69	VARON	23/11/13
MNEMOS_UCD003_69V	Sub_017	69	VARON	30/11/13
MNEMOS_UCD010_72F	Sub_018	72	MUJER	01/12/13
MNEMOS_UCD021_66F	Sub_019	66	MUJER	23/11/13
MNEMOS_UCD002_74F	Sub_020	74	MUJER	23/11/13
MNEMOS_UCD019_73F	Sub_021	73	MUJER	23/11/13
MNEMOS_UCD027_65F	Sub_022	65	MUJER	23/11/13
MNEMOS_UCD033_66F	Sub_023	66	MUJER	30/11/13
MNEMOS_UCD035_69F	Sub_024	69	MUJER	30/11/13
MNEMOS_UMA005_69V_REG	Sub_025	69	VARON	24/11/13
MNEMOS_UCD031_75F	Sub_026	75	MUJER	01/12/13
MNEMOS_UCD034_78F	Sub_027	78	MUJER	01/12/13
MNEMOS_UCD030_68F	Sub_028	68	MUJER	01/12/13
MNEMOS_UCD008_75V	Sub_029	75	VARON	23/11/13
MNEMOS_UMA009_78F	Sub_030	78	MUJER	24/11/13

Tabla 4: Pacientes incluidos en la muestra.

No se ha descartado ningún paciente por movimiento excesivo de la cabeza durante el tiempo de adquisición de las imágenes.

Cabe mencionar que, aunque la muestra inicial era mayor, los estudios de algunos pacientes tuvieron que descartarse al no haberse adquirido la serie funcional con los parámetros incluidos en la tabla 3.

4.2.- Selección del método de obtención de los mapas de conectividad funcional

Entre todos los métodos existentes para el procesado de los datos de fMRI (consultar anexo 1), los más ampliamente utilizados por la comunidad científica, han sido los métodos basados en semilla, los de análisis de componente principal (PCA), los de análisis de componentes independiente y los de agrupamiento o clustering (Van Den Heuvel et al., 2010).

En el presente estudio se ha empleado GIFT (Group ICA), que es una toolbox que se instala sobre MATLAB (Calhoun, 2004) y que implementa un método de análisis por componentes independientes. Aunque este método permite obtener patrones de conectividad del cerebro completo sin necesidad de definir una semilla, no nos permitiría medir los valores z de conectividad entre un región determinada de la DMN y otra. Precisamente, estas medidas son las que pretendemos obtener. Además, los métodos ICA presentan la dificultad de tener que clasificar visualmente cada componente obtenida entre las distintas redes funcionales presentes durante el reposo o bien descartarlas como ruido. Finalmente, se optó por utilizar un método basado en semilla. Además, se debe añadir que, de las siete redes funcionales en reposo existentes (Van den Heuvel et al., 2009), en el presente estudio solo nos interesa una de ellas, la DMN. Utilizar GIFT nos obligaría a analizar gran cantidad de componentes que pertenecen a redes excluidas del presente estudio.

DPARSF (Data Processing Assistant for Resting-State fMRI), es una toolbox de MATLAB para realizar el análisis de datos de fMRI definiendo una o varias semillas a priori y que utiliza funciones de SMP (Statistical Parametric Mapping) y REST (Resting-State fMRI Data Analysis). Las simulaciones con DPARSF tienen las siguientes fases:

- **Preparación de los datos:** Consiste en estructurar las carpetas de imágenes de una determinada forma. Se crean dos carpetas en el directorio de trabajo, una denominada "*FunRaw*" y otra "*T1Raw*" en las que se deben incluir las carpetas de la serie funcional y estructural respectivamente de los pacientes de la muestra. Una vez hecho esto, se selecciona en DPARSF el directorio de trabajo en el que se encuentran dichas carpetas.

- **Preprocesado de los datos:**

- En esta fase se realiza la conversión de las imágenes generadas por los equipos de RM de DICOM (Digital Imaging and Communications in Medicine) a NIFTI (Neuroimaging

Informatics Technology Initiative) realizando una llamada a dcm2nii del software MRICroN. Este formato de imágenes es el que se emplea para el tratamiento de imágenes ya que, entre otras ventajas, mejora la localización espacial de las imágenes (Chao-Gan et al., 2010)

- Elimina los primeros volúmenes de imágenes funcionales (generalmente 10) para que la señal se equilibre y haya dado tiempo a que el paciente se adapte al ruido del equipo de RM (Chao-Gan et al., 2010).

- Permite corregir las diferencias en el tiempo de adquisición de las imágenes de los distintos cortes. Más necesario aún, cuanto mayor sea el tiempo de repetición (TR). Se debe especificar el número de cortes, el orden de los mismos y el corte de referencia. DPARSF invoca a SPM para realizar este procesado (Chao-Gan et al., 2010). En el caso que nos ocupa, para obtener la serie funcional se han realizado 44 cortes de 140 imágenes cada uno, lo que proporciona un total de 6.160 imágenes.

- Corrige el movimiento de la cabeza de los paciente e incluso se genera un fichero con los parámetros de realineación generados por SPM. Todos los pacientes de la muestra presentan un movimiento inferior a 3mm en cualquiera de los ejes x, y y z a excepción del sujeto 18 y 24. Sin embargo, estos dos pacientes no se han descartado de la muestra ya que se han encontrado estudios previos en los que se han considerado pacientes con este mismo nivel de movimiento (Duncan et al., 2009 y Zhang et al., 2009).

- Se realiza una normalización espacial ya que la forma, tamaño, orientación y anatomía del cerebro de cada individuo puede ser diferente. En SMP esta normalización se puede realizar de dos formas, utilizando el espacio MNI (Montreal Neurological Institute) o empleando la segmentación unificada de la imagen T1 de los pacientes (Chao-Gan et al., 2010). En el presente estudio se opta por la primera opción, que remuestrea las imágenes con un tamaño de voxel de 3 x 3 x 3 mm³. Las coordenadas de las semillas son coordenadas MNI también. Tras la simulación con DPARSF, si se ha seleccionado esta opción, se genera una figura por paciente que nos permite comprobar el efecto de la normalización.

- El suavizado reduce los efectos negativos que pueda haber producido la normalización. Para ello, SPM emplea un núcleo gaussiano de 4 x 4 x 4 mm³ que tiene la forma de una distribución normal (Chao-Gan et al., 2010).

- Finalmente, se realiza un filtrado (0,01 – 0,08Hz aproximadamente) para considerar solo la señal de baja frecuencia (Chao-Gan et al., 2010). La señal BOLD tiene una frecuencia baja (<0.1Hz).

■ **Obtención de resultados:** DPARSF permite explotar más información, no solo la conectividad funcional, como se explica a continuación.

- Conectividad funcional: DPARSF puede calcular la conectividad de dos formas, "*voxel-wise*" o "*ROI-wise*". Por defecto, y en el presente estudio, se emplea el método "*voxel-wise*" que mide la correlación entre el punto seleccionado como semilla y el resto de voxeles del cerebro para generar el mapa de conectividad funcional. El mapa de coeficientes de correlación se convierte en mapa z aplicando la transformación de Fisher de r a z para mejorar la normalidad (Chao-Gan et al., 2010). En este estudio, es necesario que la conectividad funcional se mida "*voxel-wise*", ya que se quiere observar la conectividad funcional entre las semillas seleccionadas y el resto de voxeles del cerebro para determinar que regiones están conectadas con las semillas.

- ReHo (Regional Homogeneity): es una medida de la actividad cerebral que se realiza por voxel y que evalúa la similitud o sincronización entre las series temporales de ese voxel con sus vecinos más próximos (Zhang et al., 2004). Para realizar el cálculo de esta parámetro, las imágenes no deben estar suavizadas.

- ALFF (Amplitude of Low Frequency Fluctuation): Esta medida representa la amplitud de las fluctuaciones de baja frecuencia de la actividad cerebral en el estado de reposo (Zang et al., 2007). Según Zuo et al., 2010, el cerebro es capaz de generar multitud de ondas oscilatorias durante la función cerebral. En la materia gris, las oscilaciones de amplitudes mayores aparecen en las regiones asociadas con la DMN.

4.2.1.- Selección de parámetros

En este apartado se describe los parámetros seleccionados en la aplicación empleada (DPARSF) para la obtención de los mapas z de conectividad funcional de los pacientes de la muestra.

4.2.1.1.- Selección de las semillas

Los mapas z de conectividad obtenidos con DPARSF requieren la selección, a priori, de una semilla. Para abordar este problema de la forma más rigurosa posible, se han considerado varias semillas por cada una de las regiones consideradas como parte de la DMN. Todas ellas, han sido utilizadas como semillas en estudios anteriores. A continuación, se incluye una tabla con las coordenadas MNI de las semillas consideradas en cada región junto con su referencia.

PCC	COORDENADAS MNI	REGIÓN	ETIQUETA	REFERENCIA
101	[-2 -51 41]	Cingulate Gyrus	52	Lynch et al., 2013
102	[-5 -49 40]	Cingulate Gyrus	52	Fox et al., 2005
103	[-8 -56 26]	Cingulate Gyrus	52	Brewer et al., 2011
104	[0 -52 25]	Cingulate Gyrus	52	Crone et al., 2011
105	[-4 -40 43]	Cingulate Gyrus	52	Pizoli et al., 2011
106	[0 -27 30]	Cingulate Gyrus	52	Harrison, et al., 2008
107	[-6 -39 27]	Cingulate Gyrus	52	Harrison, et al., 2008
108	[-2 -36 35]	Cingulate Gyrus	52	Lynch et al., 2013
109	[-4 -36 44]	Cingulate Gyrus	52	Salgado-Pineda et al., 2011

Tabla 5: Semillas de la región PCC utilizadas en DPARSF para obtener los mapas z de conectividad.

PRECUNEUS	COORDENADAS MNI	REGIÓN	ETIQUETA	REFERENCIA
201	[-3 -69 42]	Precuneus	48	Wang et al., 2012
202	[-3 -39 48]	Precuneus	48	Wang et al., 2012
203	[-4 -74 40]	Precuneus	48	Anchisi et al., 2005
204	[0 -56 50]	Precuneus	48	Crone et al., 2011
205	[3 -72 42]	Precuneus	48	Wang et al., 2012
206	[-9 -63 27]	Precuneus	48	Harrison, et al., 2008
207	[-5 -53 41]	Precuneus	48	Fox et al., 2005
208	[4 -62 34]	Precuneus	48	Bluhm et al., 2009

Tabla 6: Semillas de la región precuneus utilizadas en DPARSF para obtener los mapas z de conectividad.

mPFC	COORDENADAS MNI	REGIÓN	ETIQUETA	REFERENCIA
301	[-6 52 -2]	Medial Frontal Gyrus	31	Brewer et al., 2011
302	[-6 51 21]	Medial Frontal Gyrus	31	Harrison, et al., 2008
303	[12 50 8]	Medial Frontal Gyrus	31	Bryant et al., 2008
304	[0 48 24]	Medial Frontal Gyrus	31	Coricelli et al., 2009
305	[12 42 21]	Medial Frontal Gyrus	31	Blakemore, 2008
306	[-2 60 10]	Medial Frontal Gyrus	31	Bluhm et al., 2009
307	[-6 42 -12]	Medial Frontal Gyrus	31	Lemogne et al., 2011
308	[10 60 10]	Medial Frontal Gyrus	31	Hipp et al., 2012
309	[10 60 18]	Medial Frontal Gyrus	31	Bluhm et al., 2009

Tabla 7: Semillas de la región mPFC utilizadas en DPARSF para obtener los mapas z de conectividad.

IPL	COORDENADAS MNI	REGIÓN	ETIQUETA	REFERENCIA
401	[54 -45 54]	Inferior Parietal Lobule	51	Harrison, et al., 2008
402	[52 -51 43]	Inferior Parietal Lobule	51	Watanabe et al., 2013
403	[-45 -39 42]	Inferior Parietal Lobule	51	Metereau et al., 2013
404	[54 -48 48]	Inferior Parietal Lobule	51	Happé et al., 2010
405	[-48 -50 50]	Inferior Parietal Lobule	51	Happé et al., 2010
406	[-36 -36 42]	Inferior Parietal Lobule	51	Happé et al., 2010
407	[-42 -54 42]	Inferior Parietal Lobule	51	Happé et al., 2010
408	[39 -42 42]	Inferior Parietal Lobule	51	Happé et al., 2010

Tabla 8: Semillas de la región IPL utilizadas en DPARSF para obtener los mapas z de conectividad.

HC	COORDENADAS MNI	REGIÓN	ETIQUETA	REFERENCIA
501	[-27 -24 -9]	Sub-Gyral	11	Duncan et al., 2009
502	[-20 -30 -8]	Sub-Gyral	11	Groen et al., 2011
503	[-27 -24 -8]	Sub-Gyral	11	Duncan et al., 2009
504	[-32 -31 -8]	Sub-Gyral	11	Duncan et al., 2009
505	[29 -20 -11]	Sub-Gyral	11	Duncan et al., 2009
506	[-22 -36 -4]	Parahippocampa Gyrus	20	
507	[24 -28 -10]	Parahippocampa Gyrus	20	Groen et al., 2011
508	[26 -36 -10]	Parahippocampa Gyrus	20	Groen et al., 2011
509	[-29 -10 -21]	Parahippocampa Gyrus	20	Duncan et al., 2009

Tabla 9: Semillas de la región HC utilizadas en DPARSF para obtener los mapas z de conectividad.

ACC	COORDENADAS MNI	REGIÓN	ETIQUETA	REFERENCIA
601	[-12 42 6]	Anterior Cingulate	39	Wang et al., 2012
602	[-1 47 -4]	Anterior Cingulate	39	Fox et al., 2005
603	[-10 42 16]	Anterior Cingulate	39	Bryant et al., 2008
604	[-8 36 12]	Anterior Cingulate	39	Lemogne et al., 2012
605	[12 30 28]	Anterior Cingulate	39	Lemogne et al., 2012
606	[9 33 18]	Anterior Cingulate	39	Cieslik et al., 2012
607	[-12 22 28]	Anterior Cingulate	39	Otsuka et al., 2006

Tabla 10: Semillas de la región ACC utilizadas en DPARSF para obtener los mapas z de conectividad.

Para identificar a qué parte del cerebro pertenecen tanto las semillas consideradas como los clusters obtenidos en los mapas z de conectividad funcional, se va a emplear la función "cuixuFindStructure" que se ejecuta sobre Matlab y que requiere la instalación de un fichero con información anatómica "Tddatabase.mat" descargable gratuitamente en <http://www.alivelearn.net/xjview8>.

4.2.1.2.- Radio del ROI

En DPARSF se debe introducir tanto las coordenadas de la semilla con la que se desea realizar el análisis como el radio de la misma. En toda la bibliografía consultada, las semillas tienen forma esférica, aunque el radio sí varía según la fuente consultada. En el presente estudio, se han considerado semillas con 6 mm de radio al igual que en Fox et al., 2005 y Marrelec et al., 2011. Según este último autor, dado el tamaño de las regiones a analizar (las que componen la DMN), un radio de 6 mm alcanza un compromiso entre la especificidad anatómica y la sensibilidad de la señal.

En DPARSF se pueden introducir tanto coordenadas de Talairach como MNI. En el presente estudio, siempre que se haga referencia a coordenadas, éstas serán coordenadas MNI.

Otros posibles tamaños del radio puede ser 5 mm (Lemogne et al., 2011), 10 mm (Bluhm et al., 2009), etc.

4.3.- Selección de parámetros en REST

Los mapas z de conectividad están constituidos por regiones que están correladas, en mayor o menor medida, con la semilla que se haya introducido como parámetro. En la siguiente imagen, se muestra el mapa z de conectividad funcional de la semilla 1 del PCC del paciente 1:

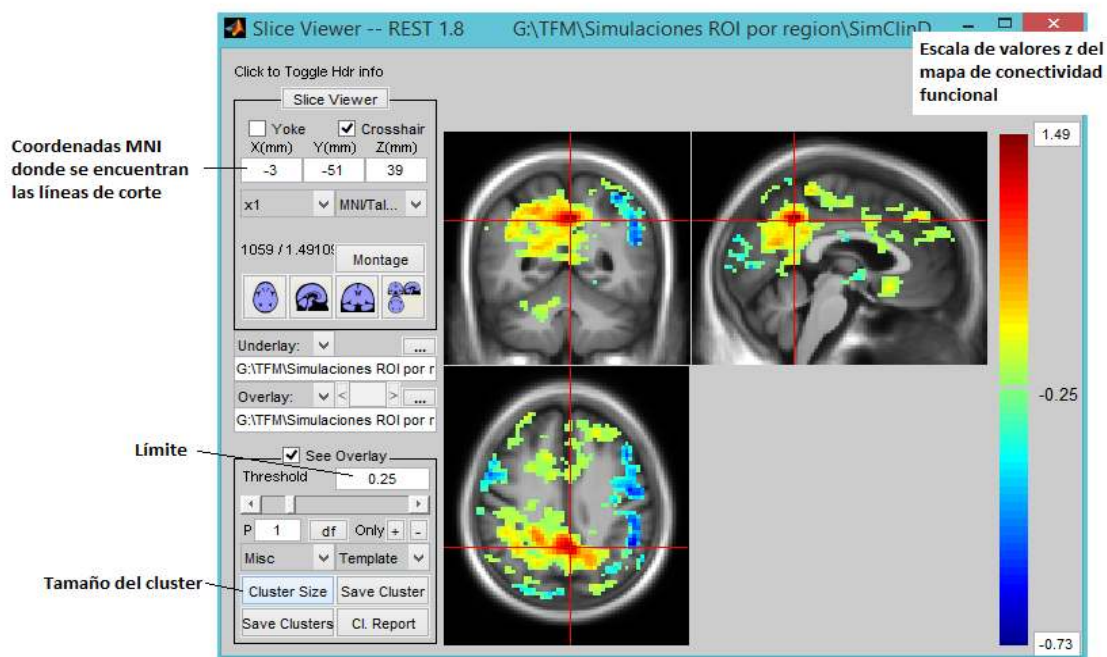


Figura 5: Visualización de un mapa z de conectividad funcional en REST.

Los mapas z de conectividad funcional generados en REST contienen tanto regiones positivamente correladas con la semilla, como negativamente correladas. En nuestro estudio, centrado en medir la conectividad funcional entre las distintas regiones de la DMN, solo se han considerado las regiones correladas positivamente con la semilla. Además, según la bibliografía consultada, algunos autores como Murphy et al., 2009 piensan que el origen de estas anticorrelaciones se deben a los diversos procesados realizados sobre la señal fMRI, aunque hay otros que afirman que el origen de las mismas si es biológico (Chai et al., 2012).

Un paso previo a la selección de variables, es definir el tamaño mínimo del clusters, es decir, agrupaciones de voxeles de tamaño inferior a este parámetro no serán considerados en el presente estudio, y el límite.

4.3.1.- Tamaño mínimo del cluster

En la aplicación REST, para obtener la información sobre los clusters existentes en los mapas z de conectividad funcional generados con DPARSF, es necesario fijar determinados parámetros que determinarán los resultados obtenidos. Uno de ellos, es el criterio de conectividad, que indica qué características deben tener dos voxeles para que se considere que pertenecen al mismo cluster. En este caso, sus caras deben tocarse, es decir, se ha empleado el criterio de conectividad $rmm = 4$. Durante la etapa de preprocesado de los datos, tras la normalización, se aplica un núcleo gaussiano de $4 \times 4 \times 4 \text{ mm}^3$ para suavizar los datos. Una vez conocidos estos parámetros, consultando el apartado "*Correction thresholds by AlphaSim*", disponible en la opción "*Misc*" de la interfaz de usuario del visor de REST, se observa que el tamaño del cluster que se debe emplear es 54 voxeles. Esta corrección se realiza para minimizar los errores de tipo I, es decir, reducir los falsos positivos que se puedan obtener por voxel.

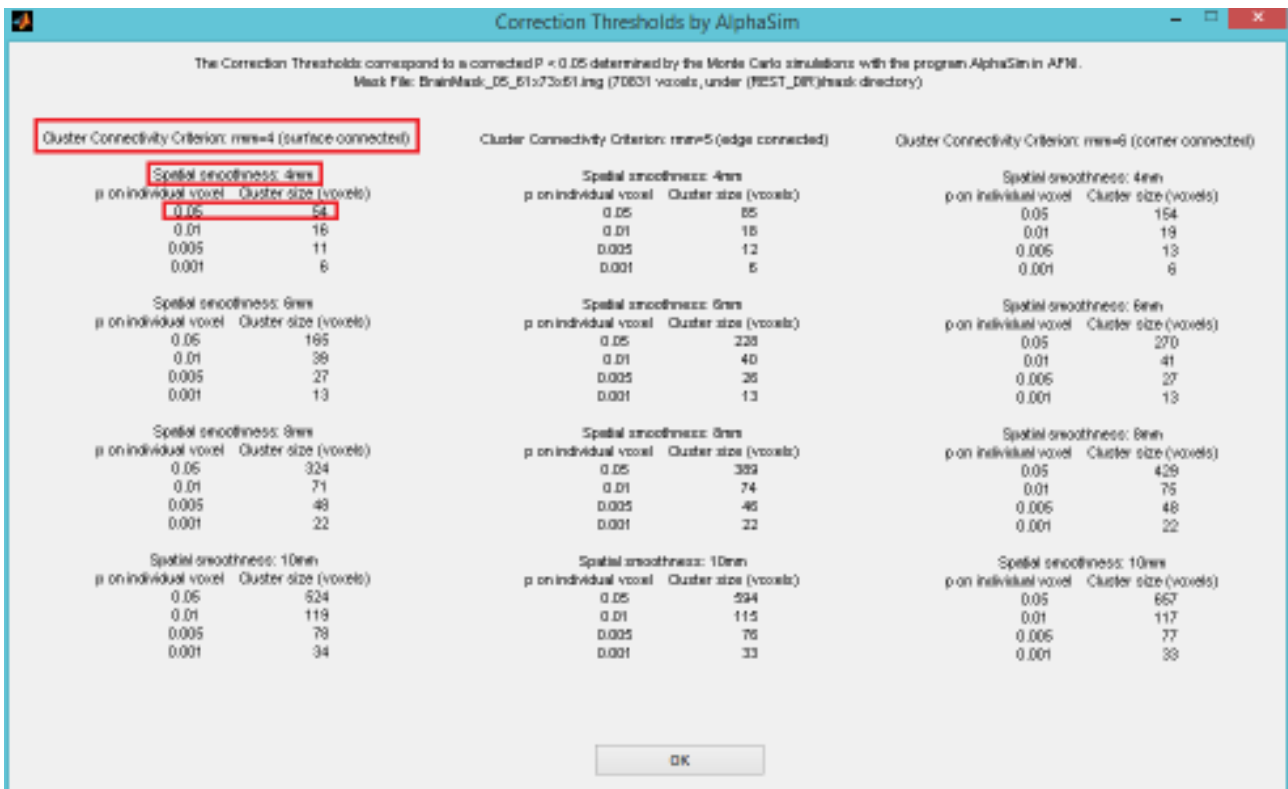


Figura 6: Corrección del límite realizada por AlphaSim en función del núcleo gaussiano empleado para el suavizado y el criterio de conectividad considerado.

4.3.2.- Límite

Como último paso en REST, para obtener la información de los clusters de los mapa z de conectividad, es imprescindible determinar el límite o threshold (en inglés) que se va a utilizar. Para su selección, se ha realizado un estudio, por cada una de las regiones de la DMN, del número de clusters reportados por REST en función del límite seleccionado. Las simulaciones se han realizado con nueve valores del límite, 0.15, 0.2, 0.25, 0.30, 0.35, 0.4, 0.45, 0.5 y 0.55. A continuación, se incluyen las gráficas de la mediana de las distintas semillas del número de clusters obtenidos con cada límite y con cada uno de los pacientes de la muestra.

■ **PCC:**

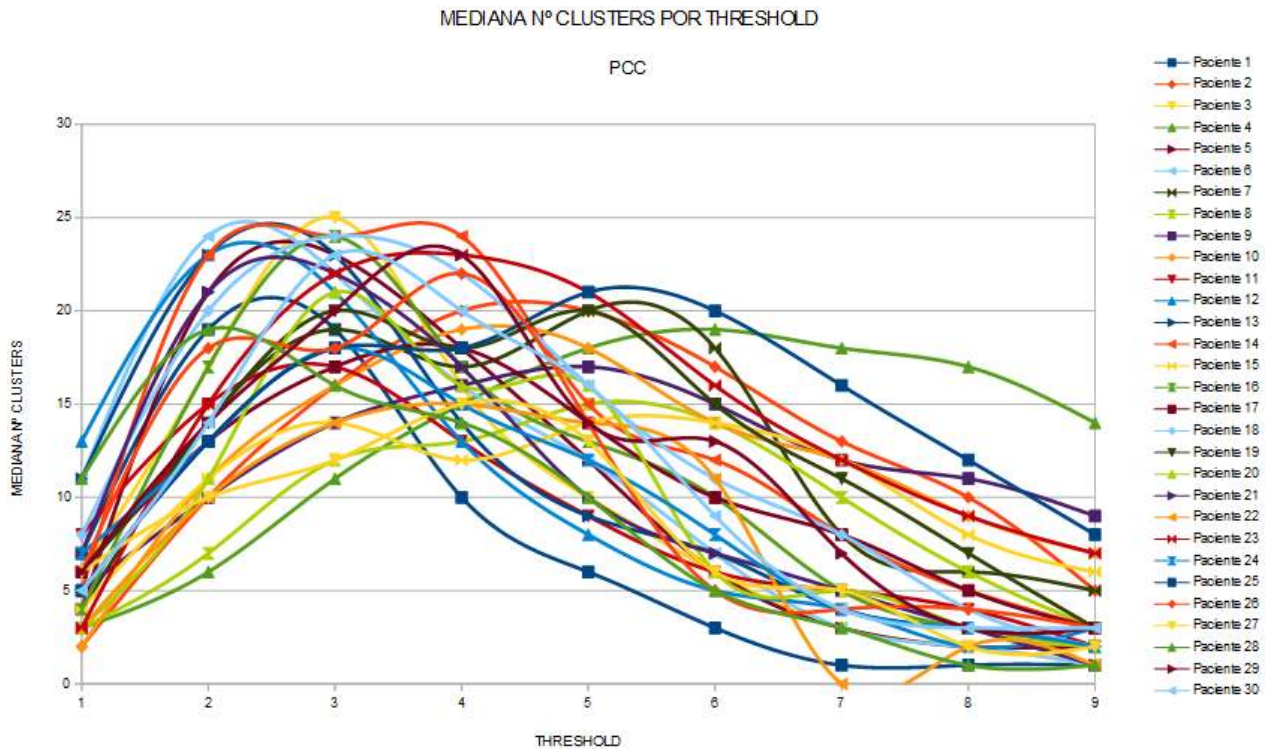


Figura 7: Mediana del número de clusters por semilla del PCC en función del límite.

■ **Precuneus:**

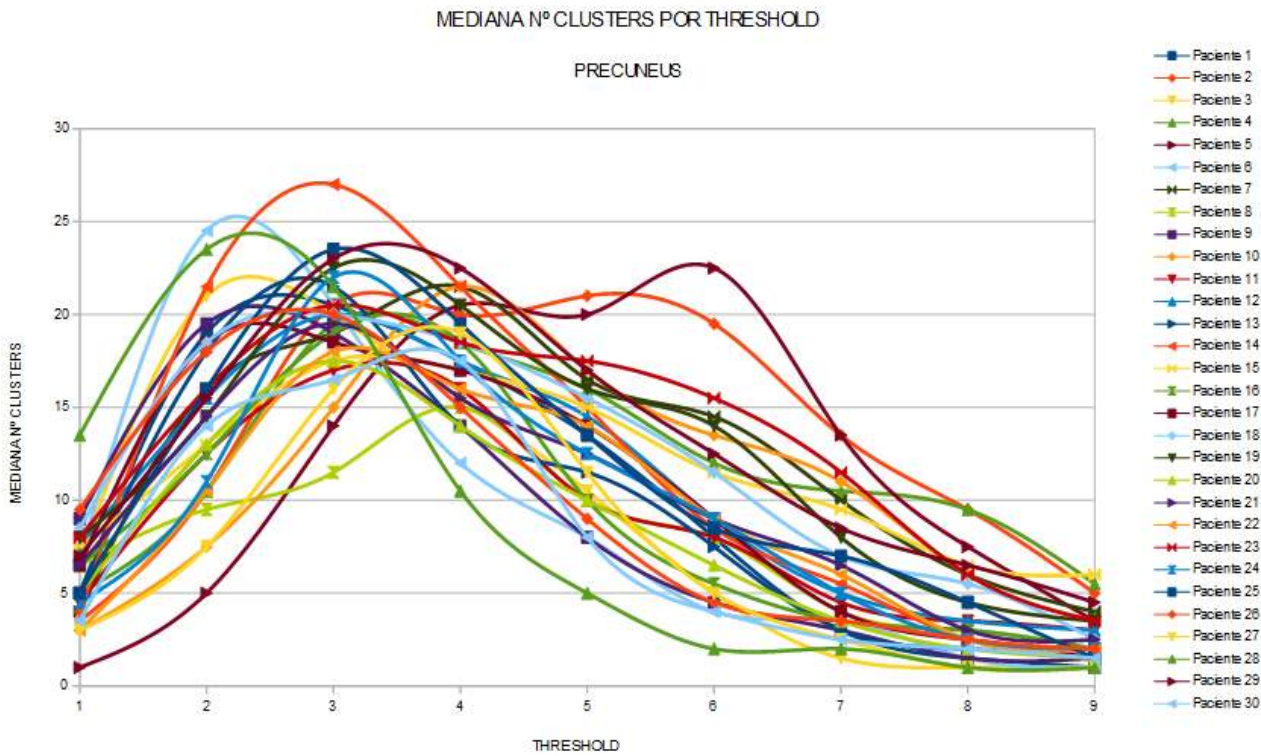


Figura 8: Mediana del número de clusters por semilla del precuneus en función del límite.

■ **mPFC:**

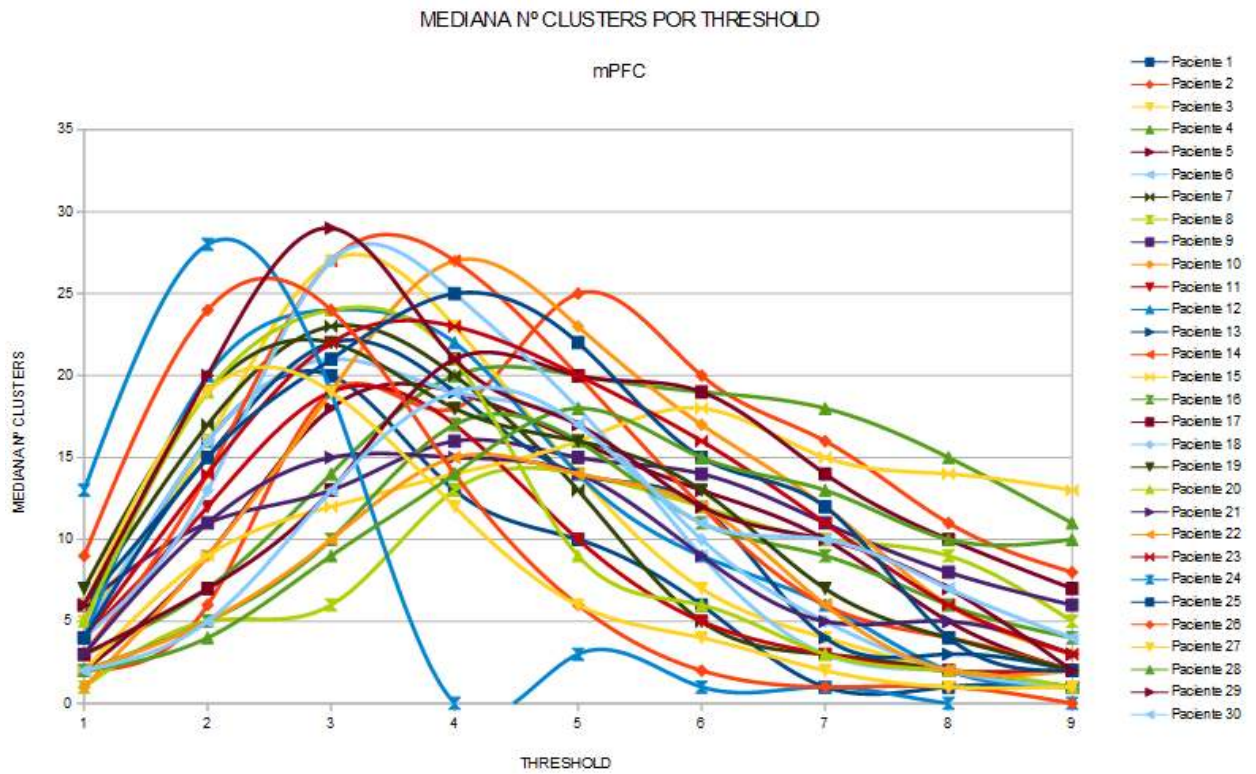


Figura 9: Mediana del número de clusters por semilla del mPFC en función del límite.

■ **IPL:**

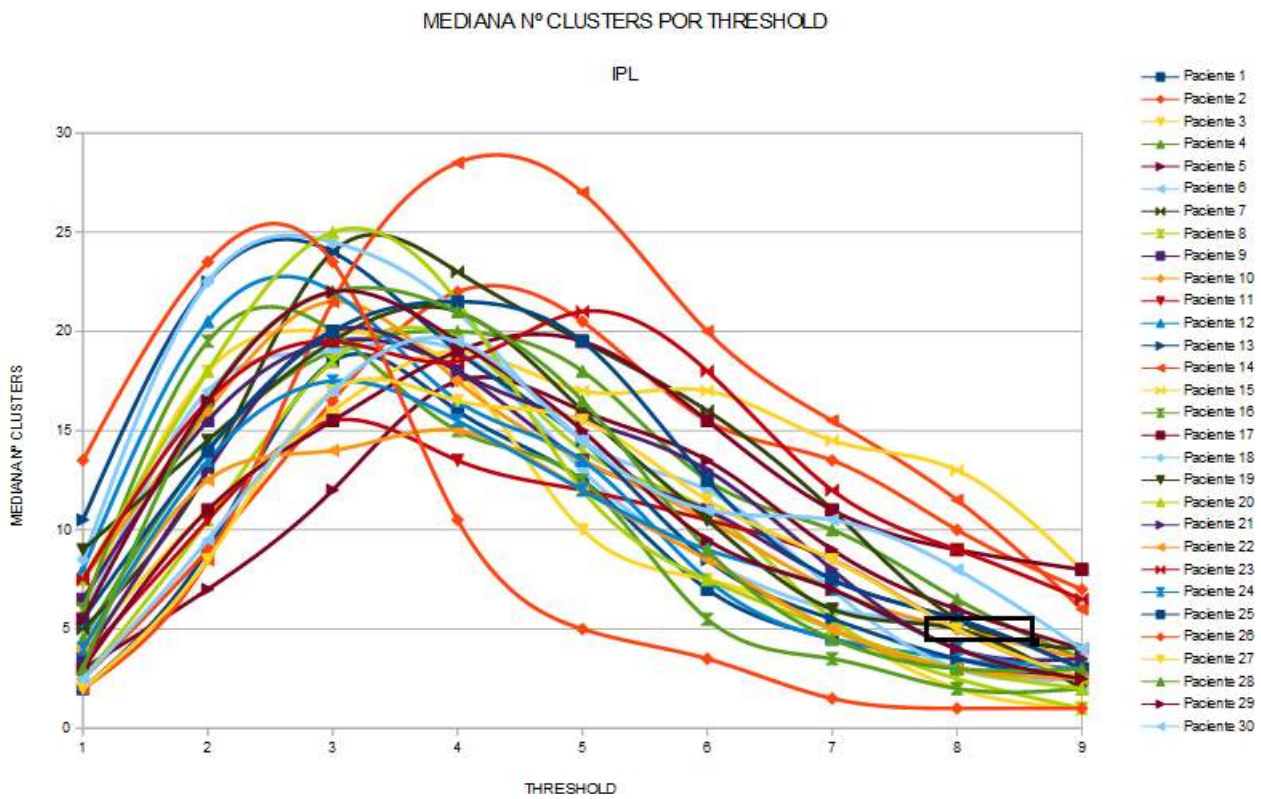


Figura 10: Mediana del número de clusters por semilla del IPL en función del límite.

■ **HC:**

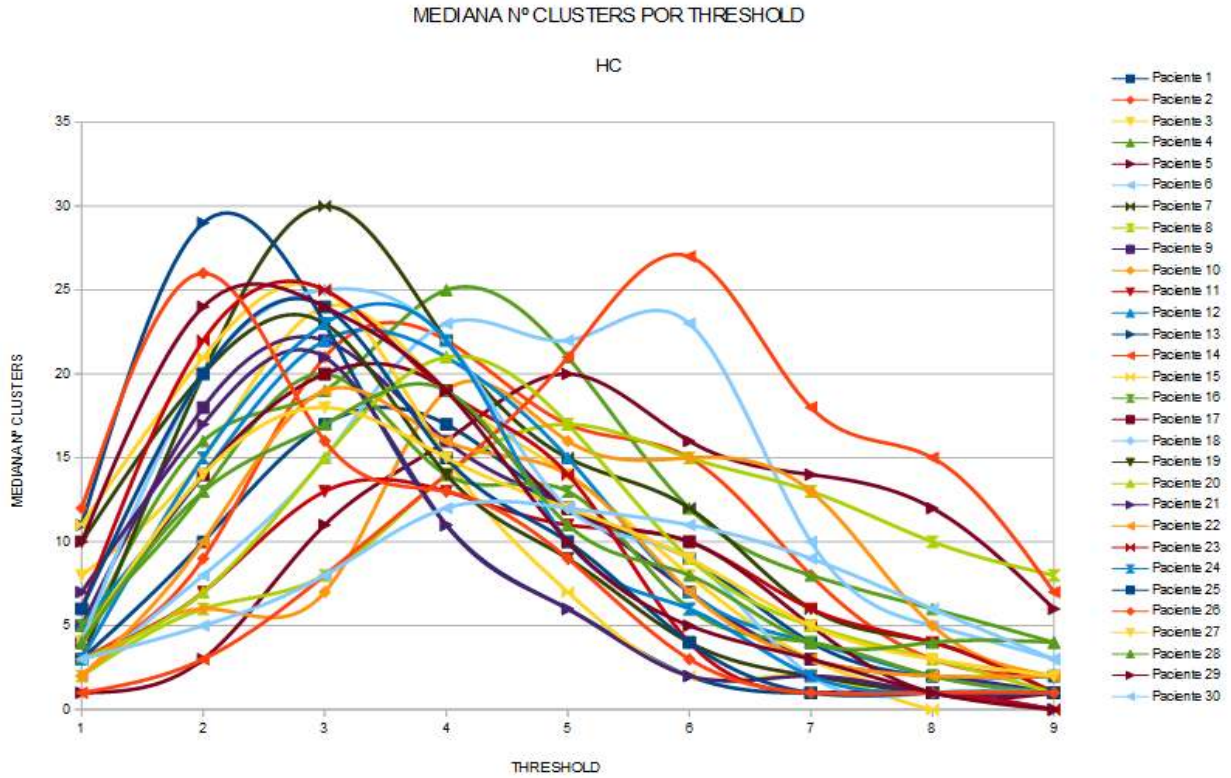


Figura 11: Mediana del número de clusters por semilla del HC en función del límite.

■ **ACC:**

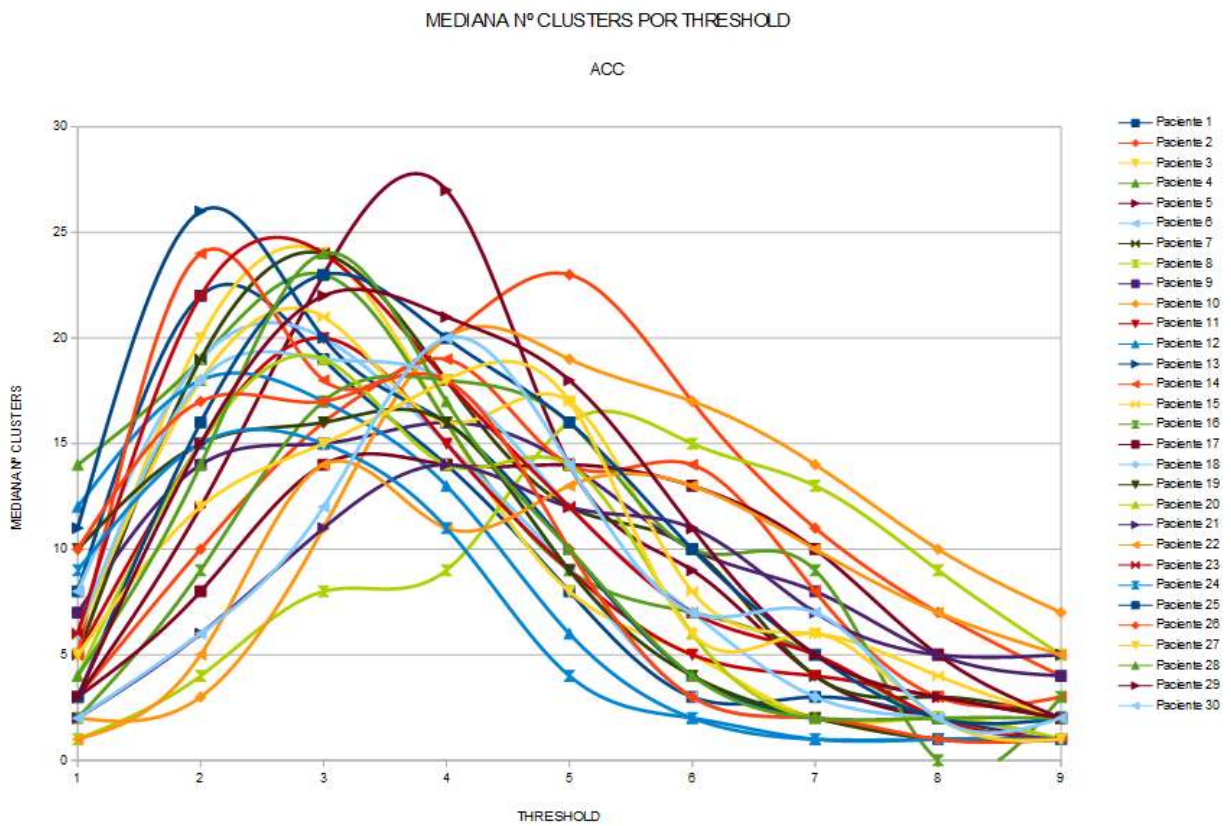


Figura 12: Mediana del número de clusters por semilla del ACC en función del límite.

En todas las regiones, se puede observar que el máximo está en torno al tercer límite, es decir, en 0.25. Por tanto, éste será el valor z mínimo que se considerará en REST para la extracción de los clusters.

4.4.- Selección de las variables a considerar

El siguiente paso es obtener las variables que se desean considerar en cada uno de los dos escenarios del presente estudio. Para ello empleamos la toolbox REST, tal y como se muestra en la figura 2. Esta herramienta dispone de muchas opciones para el procesado y el análisis estadístico de estudios R-fMRI. En el presente estudio se utiliza el visor, que dispone de una opción para generar informes de los clusters encontrados en los mapas z de conectividad introducidos. Sin embargo, la información aportada por esta opción no nos permite extraer toda la información que se requiere, por lo que se ha modificado el código del visor de REST. En las simulaciones realizadas con esta herramienta para obtener las variables deseadas, tanto en el escenario 1 como en el 2, se han utilizado los parámetros en el apartado anterior.

■ Escenario 1: En este escenario, la extracción de las variables tiene dos fases. En la primera de ellas, se debe identificar qué semilla de cada región conecta más regiones de la DMN en todo el conjunto de pacientes ("mejor semilla") y en la segunda fase, se genera un fichero con los valores de la que se ha elegido como "mejor semilla" por cada uno de los pacientes y por cada región de la DMN.

Para la selección de la "mejor semilla", por cada paciente, región y semilla, se extrae la siguiente información:

- Identificación de la semilla.
- Número de clusters obtenidos con esa semilla.
- Suma del tamaño, en voxeles, de todos los clusters.
- Valor z máximo del mapa de conectividad.
- Valor z medio ponderado por el tamaño del cluster del mapa z de conectividad.
- Varianza del parámetro anterior.
- Número de clusters detectados con esa semilla en la región PCC.
- Número de clusters detectados con esa semilla en la región Precuneus.
- Número de clusters detectados con esa semilla en la región mPFC.
- Número de clusters detectados con esa semilla en la región IPL.

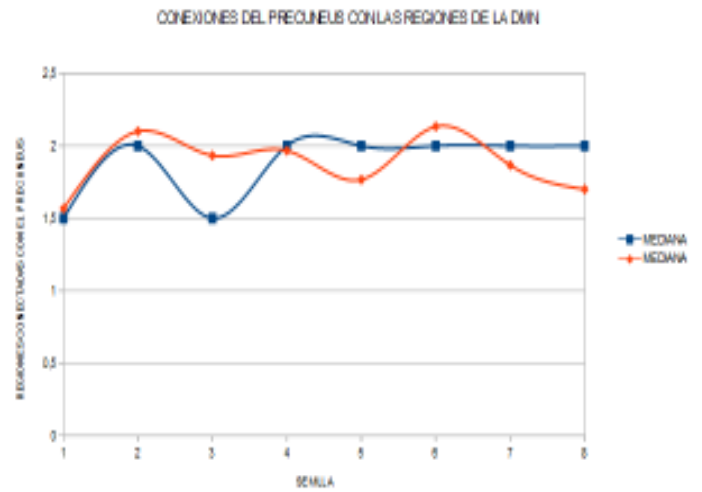
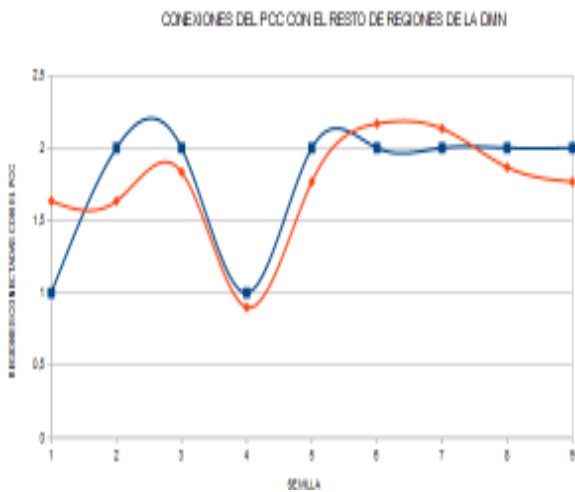
- Número de clusters detectados con esa semilla en la región HC.
- Número de clusters detectados con esa semilla en la región ACC.
- Número de regiones de la DMN que están conectadas utilizando esta semilla.

Por ejemplo, para el paciente 17, los datos serían:

Seed	Nº Clusters	Tamaño	Valor z max	Valor z medio	Varianza	Seed-PCC	Seed-Precuneus	Seed-mPFC	Seed-IPL	Seed-HC	Seed-ACC	Conexiones
PCC												
101	23	6271	2.05067	0.40745	0.02946	1	0	1	0	2	0	3
102	16	6576	1.41635	0.38982	0.02532	1	0	1	0	0	0	2
103	14	8614	1.74054	0.38123	0.01649	1	0	2	0	0	0	2
104	9	9622	2.03460	0.42574	0.00234	0	0	1	0	0	0	1
105	14	8607	1.43570	0.18838	0.03289	1	1	0	0	1	0	3
106	17	4032	1.74376	0.36678	0.03455	1	0	0	0	0	1	2
107	19	9328	1.66280	0.21419	0.08189	1	0	1	0	0	0	2
108	20	6025	1.73791	0.35590	0.01762	1	1	0	0	1	0	3
109	18	14391	1.89082	0.17814	0.04236	0	1	0	0	1	0	2
Precuneus												
201	12	8264	1.98259	0.40708	0.05150	1	1	0	0	2	0	3
202	17	10482	1.83657	0.21103	0.03020	0	1	0	0	2	0	2
203	17	6591	1.48642	0.37923	0.03430	0	1	0	0	0	0	1
204	30	7430	2.07371	0.30884	0.03477	0	1	0	1	2	0	3
205	15	6321	1.96038	0.38214	0.03389	0	1	1	0	0	0	2
206	20	4897	1.51491	0.36314	0.02782	0	1	1	0	0	0	2
207	22	5662	1.76284	0.38710	0.02655	0	1	0	1	1	0	3
208	21	5953	2.07255	0.40791	0.03935	0	1	0	0	1	0	2
mPFC												
301	10	22114	1.71260	0.01704	0.04232	0	0	1	0	0	0	1
302	18	7026	1.81026	0.43756	0.06739	1	0	1	0	0	0	2
303	11	20951	1.96521	0.03629	0.05764	0	0	1	0	0	0	1
304	23	7239	1.94577	0.42328	0.03574	1	0	1	0	0	0	2
305	23	5969	1.64584	0.37400	0.04085	0	0	1	0	0	0	1
306	19	15118	2.10959	0.08564	0.08213	1	0	1	0	1	0	3
307	8	22428	2.62329	0.00661	0.02264	0	1	1	0	0	0	2
308	13	5590	1.75731	0.43418	0.02695	1	0	1	0	0	0	2
309	13	21336	2.19067	0.01685	0.07396	0	0	1	0	0	0	1
IPL												
401	19	7717	2.16108	0.41696	0.05852	0	0	0	1	0	0	1
402	24	8141	1.41374	0.40723	0.03023	1	0	0	1	0	0	2
403	15	10533	1.70497	0.40019	0.03703	0	0	0	1	0	0	1
404	22	8776	1.69643	0.43913	0.04794	1	0	0	1	1	1	4
405	13	17112	1.72610	0.08510	0.11824	0	0	1	1	0	0	2
406	14	13304	1.50782	0.17207	0.09436	0	0	0	1	0	0	1
407	16	10690	1.69014	0.32459	0.03732	0	0	0	1	0	0	1
408	11	17087	1.83446	0.11698	0.06450	0	0	0	1	0	0	1
HC												
501	20	9632	1.67996	0.05317	0.05727	0	0	0	0	1	0	1
502	21	4685	1.40241	0.35443	0.02103	0	0	0	0	0	0	0
503	20	9632	1.67996	0.05317	0.05727	0	0	0	0	1	0	1
504	19	5287	1.68267	0.33176	0.00535	0	1	0	0	1	0	2
505	16	11511	1.54686	0.10699	0.04524	0	0	1	0	0	0	2
506	23	5345	1.29852	0.26803	0.02791	0	1	0	1	0	0	2
507	25	4071	1.55196	0.33213	0.00287	0	0	0	0	0	0	0
508	16	14429	1.22906	0.05333	0.06693	0	1	0	0	1	0	2
509	21	6343	1.59175	0.36947	0.02489	0	0	1	0	0	0	1
ACC												
601	13	8515	1.49842	0.40049	0.01851	0	0	0	0	0	1	1
602	5	175	0.59050	0.31584	0.00558	1	1	0	0	0	0	2
603	17	13569	1.51416	0.15307	0.01239	0	0	1	0	0	0	1
604	17	6612	1.29955	0.37208	0.01688	0	0	0	0	0	1	1
605	14	4436	1.34261	0.33704	0.01268	0	1	0	0	1	1	3
606	16	2411	0.67185	0.34776	0.01539	1	0	1	2	0	0	3
607	10	15413	1.51812	0.04341	0.11073	0	2	0	1	0	1	3

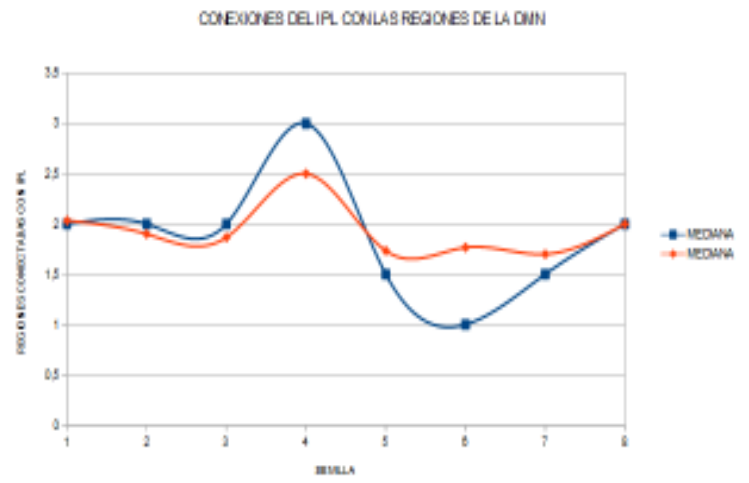
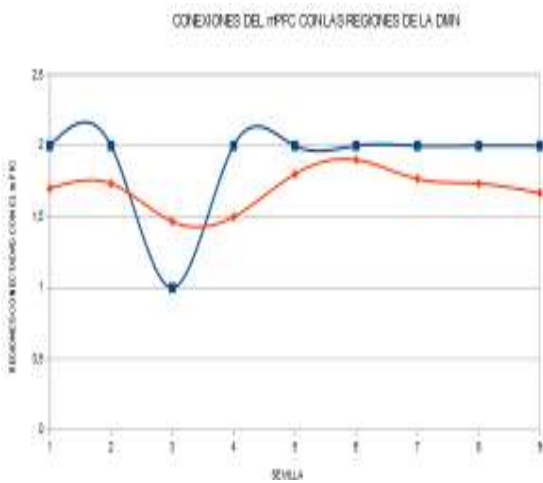
Tabla 11 : Variables extraídas por cada paciente para la selección de la mejor semilla.

Se generó esta misma tabla para todos los pacientes de la muestra. A continuación, se realizó la media y mediana del número de regiones de la DMN conectadas con cada semilla en el conjunto de todos los pacientes obteniéndose, de forma gráfica, los siguientes resultados:



Figuras 13 y 14: Media y mediana de las regiones conectadas con las semillas del PCC y precuneus respectivamente.

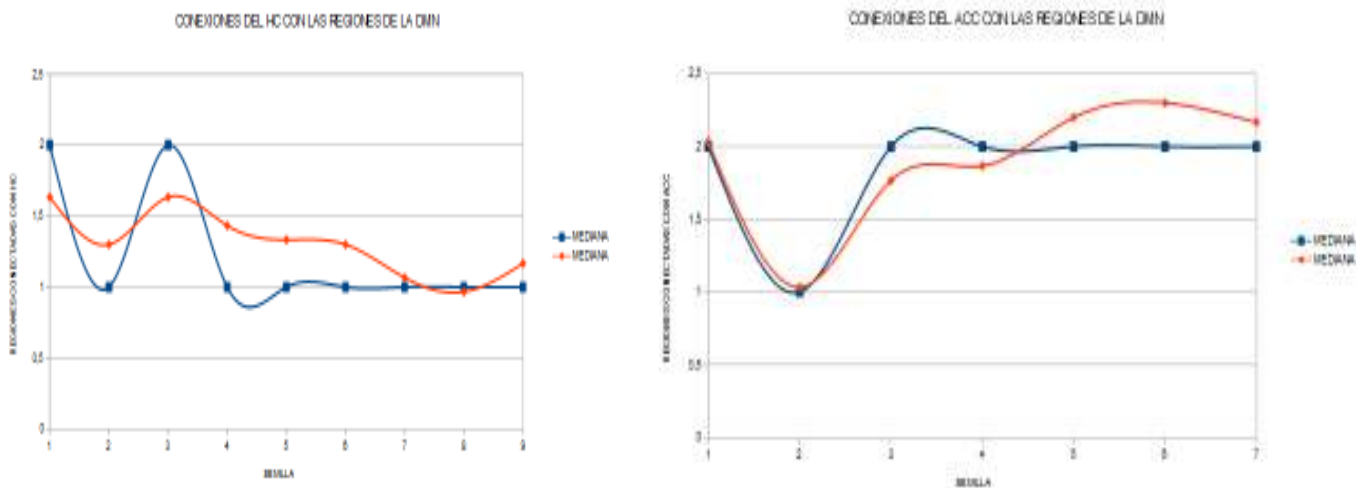
En el caso del PCC, se ha considerado como mejor semilla la seis [0 -27 30] , ya que a igual valor de mediana que las semillas 2,3, 5,7,8 y 9 tiene una media ligeramente superior. La semilla de la región precuneus con la que se obtienen los mejores valores es también la 6.



Figuras 15 y 16: Media y mediana de las regiones conectadas con las semillas del mPFC e IPL respectivamente.

En la región mPFC se selecciona la semilla 6 y en el IPL, la 4. En este último caso, la semilla cuatro tiene una mediana y media superior al del resto de semillas.

En el HC la mejor semilla puede ser la 1 o 3 (estas dos semillas están muy próximas, [-27 -24 -9] y [-27 -24 -8] respectivamente). Se selecciona la semilla 1 mientras que, en el ACC, se selecciona la seis.



Figuras 17 y 18: Media y mediana de las regiones conectadas con las semillas del HC y ACC respectivamente.

En la siguiente tabla se muestran las semillas seleccionadas de cada una de las regiones:

REGIÓN	SEMILLA	REGIÓN_ROI	COORDENADAS MNI
PCC	6	106	[0 -27 30]
Precuneus	6	206	[-9 -63 27]
mPFC	6	306	[-2 60 10]
IPL	4	404	[54 -48 48]
HC	1	501	[-27 -24 -9]
ACC	6	606	[9 33 18]

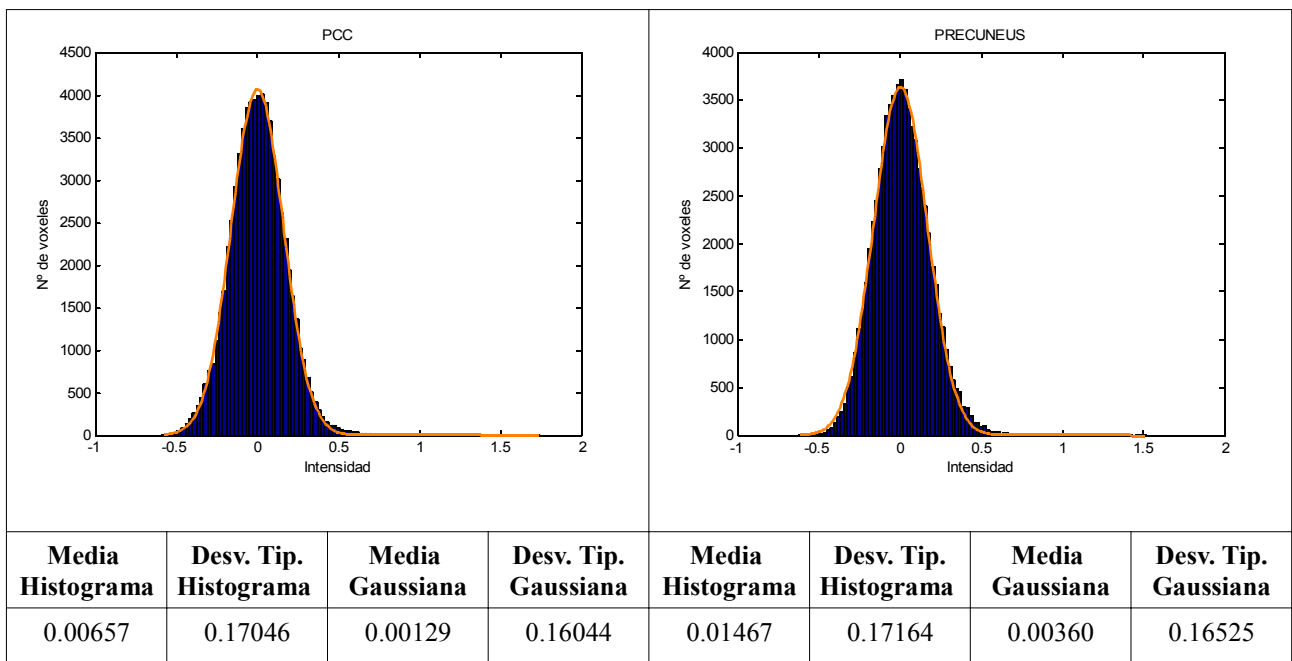
Tabla 12: Mejores semillas de cada una de las regiones de la DMN.

En este escenario, con las modificaciones realizadas en REST, también se han obtenido datos relacionados con el histograma del mapa z de conectividad funcional. Eliminando los valores nulos (más numerosos), se calcula el histograma, su media, su desviación típica y su aproximación por una gaussiana. La media y la desviación típica del histograma generado por cada una de las "mejores semillas" también se han incluido como variables. El fichero de datos que se introduce en R, tiene la siguiente forma:

Paciente	Desviación típica del histograma	Valor z máximo	Varianza	seed - PCC	seed - mPFC	seed - HC	Regiones DMN conectadas
1	0.00657	0.17496	3489	1.73563	0.32211	0.09260	1 1 0 1 1 0 4 0.00680 0.21884 4835 2.03103
2	0.01524	0.24257	19218	2.20070	0.06520	0.05273	1 0 0 0 0 0 1 0.00313 0.19962 6251 2.00180
3	0.01054	0.18826	7178	1.54122	0.20290	0.08270	1 0 0 0 1 0 2 -0.00014 0.20506 6319 1.98839
4	-0.00519	0.16504	3339	2.21644	0.39002	0.07838	1 0 0 2 0 0 2 -0.01585 0.29920 2161 0.89668
5	-0.00212	0.17261	2302	1.73127	0.34320	0.04081	1 0 1 1 0 0 3 0.00928 0.30094 28463 1.49678
6	-0.00707	0.17874	4144	1.55727	0.36545	0.02616	1 0 0 0 1 0 2 -0.00156 0.16959 3469 1.75269
7	-0.00031	0.17111	4115	1.70550	0.36086	0.09632	1 0 1 0 1 0 3 -0.00174 0.30246 72442 2.13396
8	0.00965	0.18507	5572	1.93726	0.36825	0.01565	1 0 0 0 0 0 1 0.00038 0.19590 5686 1.99295
9	-0.00998	0.22446	5243	1.91409	0.16236	0.15471	1 0 0 0 0 0 1 -0.00824 0.20084 4752 2.11941
10	-0.01745	0.21225	5602	1.98411	0.37206	0.03682	1 0 0 0 0 0 1 0.00264 0.24333 3583 0.82389
11	0.00963	0.17235	4944	1.55597	0.35553	0.03198	1 0 0 0 1 0 2 0.00397 0.21702 11858 1.79729
12	0.00226	0.17775	4905	2.01665	0.36293	0.04937	1 0 0 1 1 1 4 0.00382 0.19049 5449 1.74299
13	0.00256	0.17436	3767	2.06138	0.36791	0.04669	1 0 0 1 0 1 3 0.01178 0.19844 6589 1.56656
14	-0.00988	0.17436	3573	1.77010	0.36686	0.05149	1 0 0 0 0 1 2 -0.00544 0.19457 6170 1.47062
15	-0.00376	0.15324	2471	2.03112	0.38408	0.07094	2 1 0 2 0 0 3 -0.01073 0.25659 8664 1.96466
16	-0.00545	0.18287	3440	1.79735	0.36292	0.05635	1 1 0 1 0 1 4 -0.00408 0.18684 5141 1.30616
17	0.00657	0.17046	4032	1.74376	0.36678	0.04355	1 0 0 0 0 1 2 0.01467 0.17164 4897 1.51491
18	-0.00003	0.17915	4875	2.03992	0.23034	0.15469	1 0 1 2 0 0 3 0.00560 0.23042 6424 1.84312
19	0.00139	0.19213	5349	2.32870	0.39917	0.04387	1 0 0 1 0 0 2 0.00101 0.20574 6583 1.79850
20	0.00318	0.18656	6676	1.75106	0.10191	0.09980	1 0 1 0 0 1 3 -0.00258 0.22328 2080 0.75271
21	0.00472	0.21143	7956	2.31092	0.36886	0.02178	1 0 0 2 0 0 2 -0.00146 0.18879 4628 1.55650
22	-0.00861	0.20814	4942	1.64119	0.36132	0.02317	1 0 0 0 0 0 1 0.00289 0.26258 12346 2.15129
23	0.00094	0.21861	5122	0.89746	0.40166	0.03492	0 0 0 1 0 0 1 0.00882 0.21888 7744 2.04057
24	-0.01108	0.22664	5259	1.85971	0.39253	0.01395	1 1 0 0 0 0 2 0.00158 0.18100 4073 1.73971
25	-0.00413	0.18297	5216	2.30682	0.13718	0.12615	2 0 0 1 0 0 2 0.00125 0.20398 6550 1.77262
26	0.01749	0.21686	11454	1.67052	0.09878	0.03280	1 0 0 0 0 0 1 0.00235 0.16036 3436 1.43899
27	0.01131	0.22562	10360	1.93589	0.12260	0.05567	1 0 1 0 0 0 2 0.01262 0.22843 13496 1.87961
28	-0.00062	0.18043	4088	1.85509	0.31518	0.01692	1 2 0 0 1 0 3 0.00420 0.16927 4051 1.62442
29	-0.01098	0.19702	4820	1.98329	0.33897	0.05626	1 0 0 0 0 1 2 -0.00411 0.25597 8605 2.31394
30	0.00485	0.23360	14820	2.45632	0.10145	0.03302	1 0 0 0 0 0 1 0.01446 0.26010 18663 1.95772

Figura 19: Representación parcial del fichero de datos introducido en R en el primer escenario.

A modo de ejemplo, se incluyen los histogramas de los mapas z de conectividad obtenido con las semillas 6 del PCC, 6 del precuneus, 6 del mPFC, 4 del IPL, 1 del HC y 6 del ACC del paciente 17. La línea naranja representa la aproximación del histograma por una función gaussiana, cuyos parámetros también podrían haber sido explotados, pero se han considerado más realistas los valores del histograma.



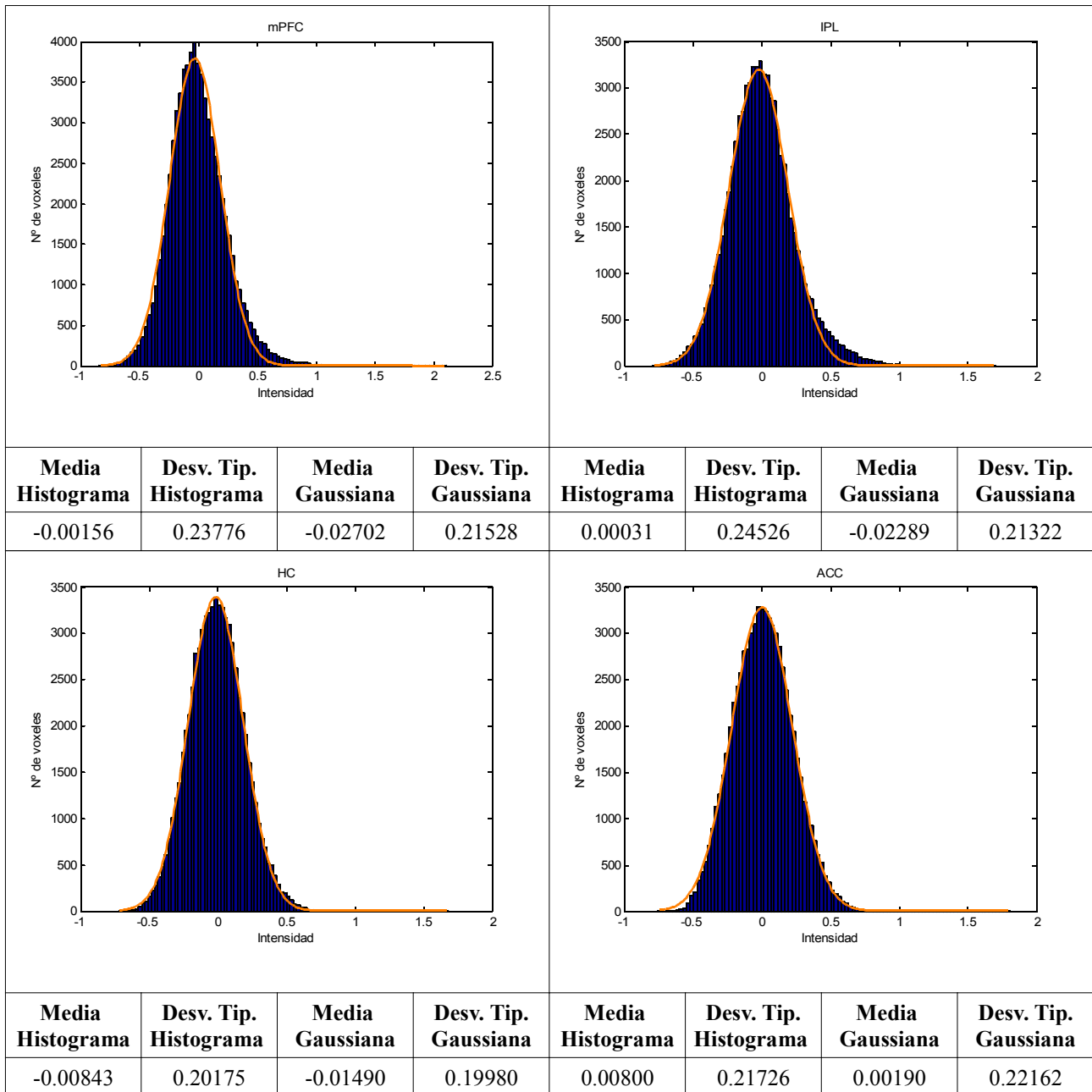


Tabla 13: Representación del histograma del mapa z de conectividad funcional generado por la mejor semilla de cada región y su aproximación por una función gaussiana (paciente 17).

■ Escenario 2: Como ya se ha comentado, en este escenario se calcula el mapa z acumulado de todas las semillas de una misma región. Para calcular el valor z de aquellos voxeles presenten en clusters de varias semillas, se ha empleado tanto la media como la mediana. En este caso, las variables seleccionadas serán:

- Tamaño, en voxeles, del mapa z acumulado de cada región. Este tamaño, obviamente, será mayor que los tamaños contemplados en el escenario 1.

- Valor z máximo.
- Valor z medio, utilizando la media y la mediana cuando el mismo voxel pertenece a clusters generados con distintas semillas de la misma región.
- Número de semillas de la región que presentan conectividad con el resto de regiones de la DMN dividido por el número total de semillas de la región en cuestión.

Como ejemplo, se ha incluido la siguiente tabla que muestra los datos del escenario 2 del paciente 17, cuando se utiliza la media para calcular el valor z de los voxeles que se superponen (primera tabla) y cuando se emplea la mediana (segunda tabla).

Región	Tamaño	Valor z máx	Valor z medio	Región-PCC	Región-Precuneus	Región-mPFC	Región-IPL	Región-HC	Región-ACC
PCC	17569	2.05067	0.28488	0.77778	0.33333	0.55556	0.00000	0.44444	0.11111
Precuneus	14931	2.07371	0.34274	0.12500	1.00000	0.25000	0.25000	0.62500	0.00000
mPFC	22912	2.62329	0.04093	0.44444	0.11111	1.00000	0.00000	0.11111	0.00000
IPL	18350	2.16108	0.22653	0.25000	0.00000	0.12500	1.00000	0.12500	0.12500
HC	15844	1.68267	0.06409	0.00000	0.33333	0.22222	0.11111	0.55556	0.00000
ACC	17086	1.51812	0.15949	0.28571	0.42857	0.28571	0.28571	0.14286	0.57143

Tabla 14: Datos del escenario 2 del paciente 17 cuando se calcula el valor z de los voxeles que se superponen con la media.

Región	Tamaño	Valor z máx	Valor z medio	Región-PCC	Región-Precuneus	Región-mPFC	Región-IPL	Región-HC	Región-ACC
PCC	17569	2.05067	0.30949	0.77778	0.33333	0.55556	0.00000	0.44444	0.11111
Precuneus	14931	2.07371	0.35310	0.12500	1.00000	0.25000	0.25000	0.62500	0.00000
mPFC	22912	2.62329	0.04953	0.44444	0.11111	1.00000	0.00000	0.11111	0.00000
IPL	18350	2.16108	0.24102	0.25000	0.00000	0.12500	1.00000	0.12500	0.12500
HC	15844	1.68267	0.07480	0.00000	0.33333	0.22222	0.11111	0.55556	0.00000
ACC	17086	1.51812	0.20126	0.28571	0.42857	0.28571	0.28571	0.14286	0.57143

Tabla 15: Datos del escenario 2 del paciente 17 cuando se calcula el valor z de los voxeles que se superponen con la mediana.

Se puede observar que la única diferencia entre las dos tablas, es el valor del valor z medio, el resto de valores son exactamente iguales. Nos quedaremos con la segunda variante que emplea la mediana, al considerar esta medida más robusta.

El fichero a introducir en R de este escenario presenta la siguiente forma:

paciente	Tamaño acumulado	Valor z máximo acumulado	Valor z medio acumulado	Semillas que presentan conectividad con el Precuneus dividido entre el número de semillas de la región	Semillas que presentan conectividad con el PCC dividido entre el número de semillas de la región	Semillas que presentan conectividad con el mPFC dividido entre el número de semillas de la región	Semillas que presentan conectividad con el IPL dividido entre el número de semillas de la región	Semillas que presentan conectividad con el HC dividido entre el número de semillas de la región	Semillas que presentan conectividad con el ACC dividido entre el número de semillas de la región	Se repiten los mismos datos para el Precuneus, etc						
1	1	11720	2.13148	0.33824	0.77778	0.22222	0.11111	0.11111	0.66667	0.11111	17324	2.08103	0.22940	0.00000	1.00000	0.12500
2	2	21303	2.21042	0.16480	0.66667	0.22222	0.33333	0.00000	0.00000	0.55556	20616	2.23901	0.24689	0.00000	0.75000	0.25000
3	3	12180	1.84635	0.29872	0.55556	0.22222	0.22222	0.00000	0.55556	0.00000	10928	2.15217	0.33960	0.12500	0.87500	0.25000
4	4	15770	2.21644	0.37947	0.33333	0.11111	0.11111	0.11111	0.33333	0.00000	15652	2.73548	0.41673	0.00000	0.62500	0.25000
5	5	28865	2.08294	0.11801	0.66667	0.22222	0.22222	0.22222	0.44444	0.00000	28463	2.41593	0.10636	0.25000	0.87500	0.37500
6	6	13059	2.23044	0.31873	0.66667	0.11111	0.11111	0.55556	0.77778	0.33333	13628	2.67039	0.34145	0.00000	1.00000	0.25000
7	7	21198	2.26365	0.17471	0.55556	0.11111	0.11111	0.11111	0.44444	0.22222	24815	2.13396	0.11045	0.12500	0.75000	0.12500
8	8	24465	1.95918	0.13352	0.77778	0.11111	0.00000	0.00000	0.33333	0.00000	27459	2.23798	0.22297	0.00000	0.87500	0.00000
9	9	15216	2.15576	0.20037	0.55556	0.22222	0.11111	0.00000	0.55556	0.11111	9146	2.11941	0.37672	0.12500	0.87500	0.25000
10	10	30142	2.31371	0.11911	0.55556	0.11111	0.00000	0.00000	0.00000	0.00000	27635	2.21535	0.09519	0.00000	0.50000	0.12500
11	11	18092	2.00150	0.26745	0.66667	0.11111	0.11111	0.00000	0.44444	0.00000	18784	2.28367	0.30901	0.00000	0.87500	0.12500
12	12	9451	2.36195	0.36229	0.55556	0.33333	0.55556	0.11111	0.33333	0.44444	13518	2.71238	0.34841	0.00000	1.00000	0.00000
13	13	11442	2.10092	0.36083	0.66667	0.33333	0.11111	0.11111	0.33333	0.55556	15129	2.43314	0.33968	0.12500	1.00000	0.50000
14	14	17194	2.10973	0.35035	0.77778	0.11111	0.11111	0.11111	0.11111	0.22222	13084	2.26752	0.29796	0.50000	0.87500	0.00000
15	15	15854	2.03112	0.36347	0.55556	0.22222	0.11111	0.22222	0.66667	0.00000	25028	2.15238	0.26567	0.12500	0.87500	0.00000
16	16	13787	2.03438	0.32082	0.77778	0.33333	0.11111	0.22222	0.11111	0.33333	16017	2.27712	0.38023	0.12500	0.62500	0.25000
17	17	17569	2.05067	0.30949	0.77778	0.33333	0.55556	0.00000	0.44444	0.11111	14931	2.07371	0.35310	0.12500	1.00000	0.25000
18	18	13858	2.34588	0.36139	0.66667	0.22222	0.66667	0.44444	0.44444	0.00000	14269	2.23148	0.36429	0.12500	0.87500	0.37500
19	19	20101	2.33927	0.37878	0.77778	0.11111	0.11111	0.11111	0.22222	0.00000	18371	2.12598	0.35656	0.12500	1.00000	0.25000
20	20	25961	2.31358	0.08559	0.44444	0.22222	0.22222	0.00000	0.22222	0.11111	19819	2.56621	0.08025	0.00000	0.75000	0.25000
21	21	14208	2.31092	0.24559	0.77778	0.66667	0.33333	0.44444	0.44444	0.00000	15872	2.45442	0.38099	0.25000	0.87500	0.25000
22	22	15981	2.23928	0.19472	0.55556	0.22222	0.33333	0.11111	0.33333	0.22222	14839	2.77053	0.15346	0.00000	0.87500	0.25000
23	23	16917	2.65005	0.41080	0.66667	0.11111	0.44444	0.11111	0.00000	0.22222	15426	2.54583	0.40723	0.00000	1.00000	0.37500
24	24	15270	2.36275	0.37528	0.77778	0.22222	0.22222	0.33333	0.44444	0.11111	17042	2.28211	0.26840	0.12500	0.62500	0.37500
25	25	22968	2.30682	0.21474	0.55556	0.11111	0.44444	0.66667	0.00000	0.00000	17084	2.56278	0.27729	0.00000	0.87500	0.50000
26	26	12083	2.07931	0.24800	0.66667	0.33333	0.00000	0.00000	0.55556	0.00000	11630	2.25899	0.28426	0.12500	1.00000	0.00000
27	27	12897	1.93589	0.25289	0.77778	0.11111	0.11111	0.11111	0.44444	0.00000	17749	2.05363	0.24326	0.00000	1.00000	0.12500
28	28	18693	1.85509	0.25599	0.77778	0.33333	0.00000	0.00000	0.55556	0.00000	13493	2.50076	0.35557	0.25000	1.00000	0.12500
29	29	15650	2.05794	0.26998	0.55556	0.33333	0.33333	0.22222	0.44444	0.11111	15306	2.43816	0.37511	0.00000	1.00000	0.62500
30	30	17435	2.45632	0.17006	0.77778	0.11111	0.55556	0.33333	0.11111	0.00000	19401	2.15577	0.19720	0.00000	1.00000	0.00000

Figura 20: Representación parcial del fichero de datos introducido en R en el segundo escenario.

4.5.- Aplicación de técnicas de minería de datos

En este apartado se exponen las técnicas de minería de datos empleadas en el estudio. En primer lugar, se deberá realizar un preprocesado sobre los datos (estandarización), ya que existen variables con rangos de valores muy diferentes. En segundo lugar, será necesario aplicar técnicas de reducción de la dimensionalidad, puesto que el número de instancias es bastante menor que el número de atributos. Finalmente, aplicaremos una técnica de agrupamiento para obtener grupos de pacientes.

4.5.1.- Preprocesado del conjunto de datos.

Aunque en algunos métodos de minería de datos no es necesario realizar un preprocesado de los datos, como en el caso de los árboles de decisión o los sistemas de reglas, en otros sí. Tanto en el caso de los mapas de difusión como en el caso del análisis de componentes principales (PCA), es necesario realizar un escalado de los datos que debe ser no lineal, puesto que en nuestro conjunto de datos existen unas variables (el tamaño de los clusters de los mapas z de conectividad funcional) que tienen unos valores mucho más elevados que el resto. Además, y solo en el caso de PCA, se realizará, a posteriori, un reescalado de los datos previamente estandarizados, ya que esta técnica requiere un rango de datos entre cero y uno.

4.5.2.- Reducción de la dimensionalidad.

Antes de aplicar el algoritmo de agrupamiento debemos disminuir el número de variables a contemplar, ya que es muy superior al número de instancias del conjunto de datos. En el escenario 1, manejamos 79 variables y en el escenario dos, 55. El conjunto de datos está constituido, en ambos casos, por 30 instancias, una por cada uno de nuestros pacientes. Nos enfrentamos a los que se ha denominado "*la maldición de la dimensionalidad*" (Hernández et al., 2004).

Para disminuir el número de variables, se puede recurrir a los algoritmos de reducción de la dimensionalidad, que se dividen en técnicas de selección y técnicas de extracción o transformación de variables (Hernández et al., 2004).

■ Selección de atributos: En este caso, el objetivo es considerar únicamente el mejor subconjunto de atributos del conjunto de datos original, es decir, se eliminan los atributos menos relevantes. Utilizando WEKA, se ha aplicado esta técnica a los conjuntos de datos de los escenarios 1 y 2, seleccionando como evaluador de atributos "*CfsSubsetEval*" y como método de búsqueda los siguientes:

- "*BestFirst*": el cual busca en el espacio de los subconjuntos de atributos utilizando la estrategia greedy hillclimbing con backtracking (Hall, M.).
- "*GeneticSearch*": Realiza la búsqueda utilizando algoritmos genéticos sencillos (Hall, M.).

- "*GreedyStepwise*": Puede realizar la búsqueda hacia adelante o hacia atrás en el espacio formado por el subconjunto de atributos y puede comenzar con todos o con ningún atributo o bien, en cualquier punto del espacio de búsqueda. El algoritmo para, cuando añadir o eliminar atributos tiene como resultado un descenso en la evaluación (Hall, M.).
- "*Exhaustive Search*": La simulación con este método de búsqueda requería demasiado tiempo. La búsqueda exhaustiva prueba todas las combinaciones posibles de atributos (Hall, M.).

El evaluador de atributos "*CfsSubsetEval*" utilizado, mide el valor de un subconjunto de atributos considerando la capacidad de predicción individual de cada característica junto con el grado de redundancia entre ellas (Hall, M.).

■ Extracción o transformación de atributos: Permite disminuir el número de atributos a considerar generando nuevas variables que no se corresponden con las variables del conjunto de datos original. Se probaron dos técnicas:

- Análisis de Componentes Principales (PCA): Es una técnica lineal que reduce el conjunto de atributos o variables original en otras, que son una combinación lineal de las primeras. Las nuevas variables son independientes entre sí y se construyen por orden de importancia en función de la variabilidad total que recogen de la muestra (Hernández et al., 2004). Para poder aplicar el análisis de componente principal es necesario reescalar los datos ya que el valor de éstos debe oscilar entre 0 y 1 (Hernández et al., 2004).

- Mapas de difusión: Es una técnica no lineal que estima la distancia real entre dos puntos a través de un proceso ficticio de difusión. Para ello, nos basamos en que la distancia euclídea entre dos puntos en el espacio transformado se puede aproximar por la distancia de difusión entre estos dos puntos en el espacio de datos original. Las variaciones de los parámetros en el manifold original se conservan prácticamente en el nuevo espacio de datos (Coifman et al., 2006).

4.5.3.- Algoritmo de agrupación.

A priori, no se conoce el diagnóstico (clase) de los pacientes de la muestra, por lo que debemos aplicar un algoritmo de aprendizaje no supervisado. Lo que se desea obtener son grupos de pacientes, por lo que las técnicas de agrupamiento o clustering se ajustan adecuadamente a las características de nuestro problema. En concreto, se ha utilizado el algoritmo K medias (K means, en inglés) que es un método de agrupamiento por vecindad. Es el método más popular de los métodos de agrupamiento denominados "por partición" en contraposición a los métodos jerárquicos. K medias determina unos prototipos o centroides en el espacio, de tal forma que las instancias que pertenecen al mismo grupo tengan características similares (Hernández et al., 2004). Los grupos se definen minimizando la suma de las distancias cuadráticas entre cada instancia de entrada y el centro de su correspondiente clase, representado por el centroide al que está asociada esta instancia. Para aplicar este algoritmo, es preciso indicar el número de grupos que se desean formar, tres en nuestro caso (control, QSM y DCL).

El funcionamiento del algoritmo es el siguiente (Hernández et al., 2004):

- Se determinan tantos centroides como grupos se deseen formar.
- Por cada instancias de entrada, se calcula cuál es el centroide más próximo (generalmente se utiliza la distancia euclídea) y esa instancia pasa a formar parte de los ejemplos pertenecientes a ese centroide.
- Una vez introducidas todas las instancias, cada una de ellas estará asociada a un centroide.
- El centroide se desplaza hasta el centro de masas de su conjunto de instancias.
- Este proceso se repite hasta que los centroides ya no varíen.

Cuando reducimos la dimensionalidad empleando mapas de difusión, el algoritmo de agrupamiento que se ha utilizado es "*diffusionKmeans*", que realiza un Kmeans sobre el mapa de difusión permitiendo seleccionar el número de nuevas variables a considerar (Richards et al., 2014).

5.- RESULTADOS

A continuación, se exponen los resultados obtenidos con las simulaciones realizadas. Empezaremos mostrando el resultado que genera cada una de las técnicas de reducción de la dimensionalidad utilizadas.

- Reducción de la dimensionalidad: Como se mencionó en el apartado anterior, se aplican tanto algoritmos de selección de atributos como de transformación de los mismos.
 - Selección de atributos: En las siguientes tablas se muestran los resultados obtenidos aplicando como evaluador de atributos "*CfsSubsetEval*" y diferentes métodos de búsqueda.

Escenario 1:

EVALUADOR DE ATRIBUTOS	MÉTODO DE BÚSQUEDA	ATRIBUTOS SELECCIONADOS
CfsSubsetEval	BestFirst	4, 17, 25, 68, 73, 74, 75, 76, 77
CfsSubsetEval	GeneticSearch	4, 17, 25, 68, 73, 74, 75, 76, 77
CfsSubsetEval	GreedyStepwise	4, 17, 25, 68, 73, 74, 75, 76, 77
CfsSubsetEval	LinearForwardSelection	4, 17, 25, 68, 73, 74, 75, 76, 77

Tabla 16: Selección de atributos empleando diferentes métodos de búsqueda en el escenario 1.

Como se puede observar en la tabla, todos los métodos han seleccionado los mismos atributos, que son los que consideraremos como el nuevo conjunto de variables.

Escenario 2:

EVALUADOR DE ATRIBUTOS	MÉTODO DE BÚSQUEDA	ATRIBUTOS SELECCIONADOS
CfsSubsetEval	BestFirst	13, 22, 33, 34, 39, 48, 53
CfsSubsetEval	GeneticSearch	13, 22, 33, 34, 39, 48, 53
CfsSubsetEval	GreedyStepwise	13, 22, 33, 34, 39, 48, 53
CfsSubsetEval	LinearForwardSelection	13, 22, 33, 34, 39, 48, 53

Tabla 17: Selección de atributos empleando diferentes métodos de búsqueda en el escenario 2.

En este caso, ocurre lo mismo que en el escenario 1. El nuevo conjunto de variables estará compuesto por los atributos seleccionados por los métodos de búsqueda.

En la siguiente tabla, se ha incluido el nuevo subconjunto de atributos de cada uno de los escenarios, incluyendo la descripción de cada uno de ellos:

ESCENARIO 1		ESCENARIO 2	
Atributo	Descripción	Atributo	Descripción
4	Tamaño (voxels) del mapa z de conectividad funcional de la "mejor semilla" del PCC.	13	Valor z medio del mapa de conectividad funcional acumulado del Precuneus
17	Tamaño (voxels) del mapa z de conectividad funcional de la "mejor semilla" del Precuneus.	22	Valor z medio del mapa de conectividad funcional acumulado del mPFC
25	Existencia de conectividad funcional entre el Precuneus y el HC	33	Nivel de conectividad funcional del IPL con el Precuneus
68	Desviación típica del histograma del mapa z de conectividad funcional del ACC	34	Nivel de conectividad funcional del IPL con el mPFC
73	Existencia de conectividad funcional entre el ACC y el PCC	39	Valor z máximo del mapa de conectividad funcional acumulado del HC
74	Existencia de conectividad funcional entre el ACC y el Precuneus	48	Valor z máximo del mapa de conectividad funcional acumulado del ACC
75	Existencia de conectividad funcional entre el ACC y el mPFC	53	Nivel de conectividad funcional del ACC con el IPL
76	Existencia de conectividad funcional entre el ACC y el IPL		
77	Existencia de conectividad funcional entre el ACC y el HC		

Tabla 18: Enumeración y descripción de los atributos seleccionados en cada uno de los escenarios.

Analizando los atributos seleccionados en cada escenario, cabe destacar:

- Escenario 1: En este escenario la agrupación de pacientes se va a realizar, en gran medida, en función de la conectividad del ACC con el resto de regiones que constituyen la DMN.
- Escenario 2: En el escenario 2, la selección de atributos no está tan focalizada en una región como ocurre en el escenario 1. Se han incluido valores z medios de dos regiones (Precuneus y mPFC), valores z máximos de otras dos regiones (HC y ACC) y el nivel de conectividad entre IPL y Precuneus, IPL y mPFC y ACC con IPL.

Solo un atributo ha sido seleccionado en los dos escenarios, el que mide la conectividad funcional entre el ACC y el IPL.

La agrupación de pacientes obtenida en este caso es la siguiente:

Paciente	Escenario 1	Escenario 2	
Paciente 1	3	3	1
Paciente 2	2	2	1
Paciente 3	1	1	1
Paciente 4	1	1	1
Paciente 5	1	2	0
Paciente 6	1	1	1
Paciente 7	1	2	0
Paciente 8	3	3	1
Paciente 9	1	1	1
Paciente 10	2	2	1
Paciente 11	3	3	1
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	3	1	0
Paciente 15	1	1	1
Paciente 16	3	3	1
Paciente 17	3	3	1
Paciente 18	1	1	1
Paciente 19	3	1	0
Paciente 20	1	2	0
Paciente 21	3	1	0
Paciente 22	1	2	0
Paciente 23	1	1	1
Paciente 24	1	3	0
Paciente 25	3	3	1
Paciente 26	1	3	0
Paciente 27	1	2	0
Paciente 28	3	3	1
Paciente 29	1	3	0
Paciente 30	2	3	0
Pacientes igualmente agrupados			18
Cluster 1	17	12	
Cluster 2	3	7	
Cluster 3	10	11	

Figura 21: Agrupación de pacientes aplicando la selección de atributos.

- Técnicas de transformación de las variables: Se han utilizado dos algoritmos, PCA y los mapas de difusión:
 - **Análisis de Componentes Principales:** para aplicar este algoritmo, se realiza un reescalado de los datos posterior a la estandarización que permita trabajar con rangos de valores de los atributos entre 0 y 1. Una vez hecho esto se aplica, en el programa R, un análisis de componentes independientes. Sin embargo, los resultados obtenidos no nos permiten poder utilizar esta técnica para la transformación de atributos, ya que en el escenario 1 se generan 29 componentes, mientras que en el escenario 2 se generan 28. Para cubrir el 90% de los datos, necesitaríamos trabajar con 18 y 19 componentes principales respectivamente en cada uno de los casos, y son demasiadas considerando el número de instancias de las que disponemos (30).
 - **Mapas de difusión:** para generar los mapas de difusión recurrimos al paquete "*diffusionMap*" de R (Richards et al., 2014). Al ejecutar el comando "*diffuse*" se debe especificar cuantas variables de desean generar. Hemos trabajado con un rango de 2 a 6 variables, para comparar los resultados obtenidos.

2 VARIABLES			
Paciente	Escenario 1	Escenario 2	
Paciente 1	1	2	0
Paciente 2	3	3	1
Paciente 3	3	2	0
Paciente 4	1	1	1
Paciente 5	3	3	1
Paciente 6	1	2	0
Paciente 7	1	1	1
Paciente 8	3	3	1
Paciente 9	1	1	1
Paciente 10	3	3	1
Paciente 11	3	1	0
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	1	1	1
Paciente 15	1	1	1
Paciente 16	1	1	1
Paciente 17	1	1	1
Paciente 18	1	1	1
Paciente 19	1	1	1
Paciente 20	3	1	0
Paciente 21	3	1	0
Paciente 22	2	2	1
Paciente 23	2	1	0
Paciente 24	2	2	1
Paciente 25	1	1	1
Paciente 26	3	2	0
Paciente 27	2	2	1
Paciente 28	1	1	1
Paciente 29	1	1	1
Paciente 30	3	3	1
Pacientes igualmente agrupados	22		
Cluster 1	16	18	
Cluster 2	4	7	
Cluster 3	10	5	

3 VARIABLES			
Paciente	Escenario 1	Escenario 2	
Paciente 1	1	2	0
Paciente 2	3	1	0
Paciente 3	1	2	0
Paciente 4	1	1	1
Paciente 5	1	3	0
Paciente 6	1	2	0
Paciente 7	1	1	1
Paciente 8	3	3	1
Paciente 9	3	1	0
Paciente 10	3	3	1
Paciente 11	1	1	1
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	1	1	1
Paciente 15	1	1	1
Paciente 16	3	1	0
Paciente 17	3	1	0
Paciente 18	1	1	1
Paciente 19	1	1	1
Paciente 20	1	1	1
Paciente 21	3	1	0
Paciente 22	2	2	1
Paciente 23	2	1	0
Paciente 24	2	2	1
Paciente 25	3	1	0
Paciente 26	1	2	0
Paciente 27	2	2	1
Paciente 28	3	1	0
Paciente 29	1	1	1
Paciente 30	3	3	1
Pacientes igualmente agrupados	17		
Cluster 1	16	19	
Cluster 2	4	7	
Cluster 3	10	4	

4 VARIABLES			
Paciente	Escenario 1	Escenario 2	
Paciente 1	1	2	0
Paciente 2	3	1	0
Paciente 3	1	2	0
Paciente 4	1	1	1
Paciente 5	1	3	0
Paciente 6	1	2	0
Paciente 7	1	1	1
Paciente 8	3	3	1
Paciente 9	3	1	0
Paciente 10	3	3	1
Paciente 11	1	1	1
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	1	1	1
Paciente 15	1	1	1
Paciente 16	3	1	0
Paciente 17	3	1	0
Paciente 18	1	1	1
Paciente 19	1	1	1
Paciente 20	1	1	1
Paciente 21	3	1	0
Paciente 22	2	2	1
Paciente 23	2	1	0
Paciente 24	2	2	1
Paciente 25	3	1	0
Paciente 26	1	2	0
Paciente 27	2	2	1
Paciente 28	3	1	0
Paciente 29	1	1	1
Paciente 30	3	3	1
Pacientes igualmente agrupados	17		
Cluster 1	16	19	
Cluster 2	4	7	
Cluster 3	10	4	

5 VARIABLES			
Paciente	Escenario 1	Escenario 2	
Paciente 1	1	2	0
Paciente 2	3	1	0
Paciente 3	1	2	0
Paciente 4	1	1	1
Paciente 5	1	3	0
Paciente 6	2	2	1
Paciente 7	1	1	1
Paciente 8	3	3	1
Paciente 9	2	1	0
Paciente 10	3	3	1
Paciente 11	1	1	1
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	1	1	1
Paciente 15	2	1	0
Paciente 16	2	1	0
Paciente 17	2	1	0
Paciente 18	1	1	1
Paciente 19	1	1	1
Paciente 20	1	1	1
Paciente 21	3	1	0
Paciente 22	2	2	1
Paciente 23	2	1	0
Paciente 24	2	2	1
Paciente 25	2	1	0
Paciente 26	1	2	0
Paciente 27	2	2	1
Paciente 28	2	1	0
Paciente 29	2	1	0
Paciente 30	3	3	1
Pacientes igualmente agrupados	16		
Cluster 1	13	19	
Cluster 2	12	7	
Cluster 3	5	4	

6 VARIABLES			
Paciente	Escenario 1	Escenario 2	
Paciente 1	1	2	0
Paciente 2	3	1	0
Paciente 3	3	2	0
Paciente 4	3	1	0
Paciente 5	3	3	1
Paciente 6	1	2	0
Paciente 7	3	1	0
Paciente 8	3	3	1
Paciente 9	1	1	1
Paciente 10	3	3	1
Paciente 11	3	1	0
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	1	1	1
Paciente 15	1	1	1
Paciente 16	1	1	1
Paciente 17	1	1	1
Paciente 18	1	1	1
Paciente 19	3	1	0
Paciente 20	3	1	0
Paciente 21	3	1	0
Paciente 22	2	2	1
Paciente 23	2	1	0
Paciente 24	2	2	1
Paciente 25	1	1	1
Paciente 26	3	2	0
Paciente 27	2	2	1
Paciente 28	1	1	1
Paciente 29	1	1	1
Paciente 30	3	3	1
Pacientes igualmente agrupados	18		
Cluster 1	13	19	
Cluster 2	4	7	
Cluster 3	13	4	

Figura 22: Agrupación de pacientes aplicando la transformación de atributos con mapas de difusión (2 a 6 variables).

6.- DISCUSIÓN

En la agrupación obtenida reduciendo la dimensionalidad con la selección de variables, 18 pacientes son asignados al mismo grupo en el escenario 1 y 2. Sin embargo, con el conocimiento del dominio adquirido hasta el momento, no se encuentra mucha justificación para la selección de variables que se ha realizado en cada uno de los escenarios. Especialmente llamativa es la selección de atributos realizada en el escenario 1, en el que la mayor parte de las variables seleccionadas están relacionados con el ACC.

Utilizando los mapas de difusión, se han realizado simulaciones seleccionando de 2 a 6 variables. Se puede observar que el escenario 2 es más estable, ya que desde el inicio (dos variables) prácticamente proporciona la misma agrupación de pacientes. Desde el punto de vista de la minería de datos, el escenario 2 contiene más información que el 1, ya que maneja los mapas z de conectividad funcional acumulados de todas las semillas de una misma región, mientras que en el escenario 1 se selecciona solo una, la que según la mediana, es la semilla que está más conectada (en el conjunto de pacientes), con el resto de las regiones de la DMN. El escenario 2 contiene gran parte de la información del escenario 1, más la información aportada por otras muchas semillas, lo que podría justificar su estabilidad.

En la siguiente tabla, se muestra un resumen de los resultados obtenidos considerando distinto número de variables al generar los mapas de difusión. Como se puede observar, los pacientes 8,10 y 30, en todos los casos son incluidos en el grupo de pacientes 3. Los pacientes 12, 13, 14 y 18 siempre se asocian con el centroide del grupo 1 y los pacientes 22, 24 y 27 se asignan al grupo 2 independientemente del número de variables de los mapas de difusión consideradas y del escenario. En total, diez pacientes son siempre igualmente agrupados independientemente del número de variables de las mapas de difusión consideradas y del escenario en el que nos encontremos.

Paciente	2 VARIABLES	3 VARIABLES	4 VARIABLES	5 VARIABLES	6 VARIABLES	COINCIDENCIAS
Paciente 1	0	0	0	0	0	0
Paciente 2	1	0	0	0	0	1
Paciente 3	0	0	0	0	0	0
Paciente 4	1	1	1	1	0	4
Paciente 5	1	0	0	0	1	2
Paciente 6	0	0	0	1	0	1
Paciente 7	1	1	1	1	0	4
Paciente 8	1	1	1	1	1	5
Paciente 9	1	0	0	0	1	2
Paciente 10	1	1	1	1	1	5
Paciente 11	0	1	1	1	0	3
Paciente 12	1	1	1	1	1	5
Paciente 13	1	1	1	1	1	5
Paciente 14	1	1	1	1	1	5
Paciente 15	1	1	1	0	1	4
Paciente 16	1	0	0	0	1	2
Paciente 17	1	0	0	0	1	2
Paciente 18	1	1	1	1	1	5
Paciente 19	1	1	1	1	0	4
Paciente 20	0	1	1	1	0	3
Paciente 21	0	0	0	0	0	0
Paciente 22	1	1	1	1	1	5
Paciente 23	0	0	0	0	0	0
Paciente 24	1	1	1	1	1	5
Paciente 25	1	0	0	0	1	2
Paciente 26	0	0	0	0	0	0
Paciente 27	1	1	1	1	1	5
Paciente 28	1	0	0	0	1	2
Paciente 29	1	1	1	0	1	4
Paciente 30	1	1	1	1	1	5

Tabla 19: Pacientes igualmente agrupados utilizando los mapas de difusión, independientemente del número de variables del mapa y del escenario.

Comparando los resultados obtenidos tras realizar la agrupación utilizando técnicas de transformación de variables y las técnicas de selección de atributos, se puede observar que, los resultados del escenario 2 con selección de atributos son más similares a los obtenidos con transformación de variables hasta el punto de que, de los diez pacientes que se clasifican siempre de la misma manera con transformación de atributos, ocho de ellos también son igualmente clasificados en el escenario 2 con selección de atributos. No ocurre lo mismo con el escenario 1 con selección de atributos, que solo coincide con los resultados generados con los mapas de difusión en 4 de los 10 pacientes. Se puede afirmar que, el escenario 2 con selección de variables, se asemeja más a los resultados obtenidos utilizando los mapas de difusión. Como ya se ha comentado, esto podría deberse a que el escenario 2 (independientemente del método empleado para reducir la dimensionalidad) contiene más información que el 1.

Paciente	Mapas Difusión	Sel. Variables Escenario 1	Sel. Variables Escenario 2
Paciente 8	3	3	3
Paciente 10	3	2	2
Paciente 12	1	1	1
Paciente 13	1	1	1
Paciente 14	1	3	1
Paciente 18	1	1	1
Paciente 22	2	1	2
Paciente 24	2	1	3
Paciente 27	2	1	2
Paciente 30	3	2	3

Tabla 20: Comparación de los resultados obtenidos utilizando los mapas de difusión y la selección de variables.

Según la bibliografía (Sorg et al., 2007; Gili et al., 2011), los pacientes con DCL sufren las siguientes alteraciones en la DMN, ya mencionadas anteriormente:

- La conectividad funcional entre el HC y el PCC, presente en los pacientes sanos, no se observa en pacientes con DCL (Sorg et al., 2007).
- Se reduce la conectividad entre las siguientes regiones de la DMN (Gili et al., 2011):
 - PCC y ACC bilateralmente.
 - mPFC y PCC unilateralmente.
 - mPFC y ACC bilateralmente.

Aunque no se dispone del diagnóstico de los pacientes para valorar los grupos obtenidos, vamos a analizar en qué medida los pacientes incluidos en la tabla 19 han acusado los efectos que según la bibliografía sufre la DMN en pacientes con DCL. La intención es analizar si la clasificación obtenida tiene sentido clínico. Para ello, solo vamos a considerar los pacientes incluidos en la tabla 19 que son los que, aplicando los mapas de difusión, e independientemente del número de variables consideradas y del escenario que contemplemos, en todos los casos se incluyen en el mismo grupo. Es decir, del cluster 1 se han considerado solo 4 pacientes (12, 13, 14 y 18) del cluster 2, 3 pacientes (22, 24 y 27) y del cluster 3, 3 pacientes (8, 10 y 30). Consultando la tabla que se generó por cada paciente para determinar la mejor semilla en el escenario 1, comprobamos la existencia o no de conectividad funcional entre las regiones especificadas en el tabla 21 obteniéndose los resultados incluidos en la misma. Se ha considerado que existe conectividad si con alguna semilla existe un cluster en las regiones estudiadas. Los pacientes pertenecientes al grupo 3, que además es el menos numeroso, es el que presenta menor conectividad entre las

regiones especificadas, por lo que se podría pensar que este cluster es el que corresponde a los pacientes con DCL. Comparando las columnas de los clusters 1 y 2, se observa que éstos no se diferencian tanto entre sí como con la columna correspondiente al cluster 3, lo que nos lleva a pensar que estos dos clusters corresponden a los pacientes con QSM (cluster 2) y pacientes sin desórdenes detectados (cluster 1). El cluster 1, que es el que presenta una conectividad funcional mayor entre las regiones especificadas, es el más numeroso. Cabe destacar, que la conectividad de la DMN de estos pacientes también está deteriorada debido al envejecimiento.

FC	CLUSTER 1	CLUSTER 2	CLUSTER 3
PCC - HC	0 (4/4)	1 (3/3)	0 (2/3)
HC - PCC	1 (4/4)	1 (3/3)	1 (2/3)
PCC - ACC	1 (3/4)	1(2/3)	0 (3/3)
ACC - PCC	1 (3/4)	1(2/3)	0 (2/3)
mPFC - PCC	1 (4/4)	0 (2/3)	1 (2/3)
mPFC - ACC	0 (2/4)	0 (2/3)	0 (3/3)
ACC - mPFC	0 (2/4)	0 (2/3)	0 (2/3)

Tabla 21: Análisis para la identificación de los grupos.

Nota: Los números entre paréntesis indican en cuantos pacientes del grupo considerado se cumple que hay conectividad , si lo que precede al paréntesis es un "1"o que no, si lo que hay delante del paréntesis es un "0".

7.- CONCLUSIÓN

Se ha diseñado una herramienta de ayuda al diagnóstico de pacientes con DCL a partir de los estudios de R-fMRI de 30 pacientes, entre los que se encuentran personas mayores sin ningún desorden detectado, personas con QSM y pacientes con DCL. A partir de esos estudios, se calcularon los mapas z de conectividad, se extrajeron las variables a explotar y, finalmente, se aplicaron técnicas de minería de datos obteniéndose una agrupación de pacientes. Se han planteado dos escenarios posibles, el primero de ellos en el que se considera una única semilla de cada una de las regiones de la DMN (aquella que muestra más conexiones con el resto de regiones de la DMN en el conjunto de pacientes) y el segundo, en el que se acumulan los mapas z de conectividad funcional de todas las semillas de cada una de las regiones. En base a los resultados obtenidos, se puede afirmar que el segundo escenario es más estable que el primero, por lo que se podría considerar como una forma más óptima de abordar el problema.

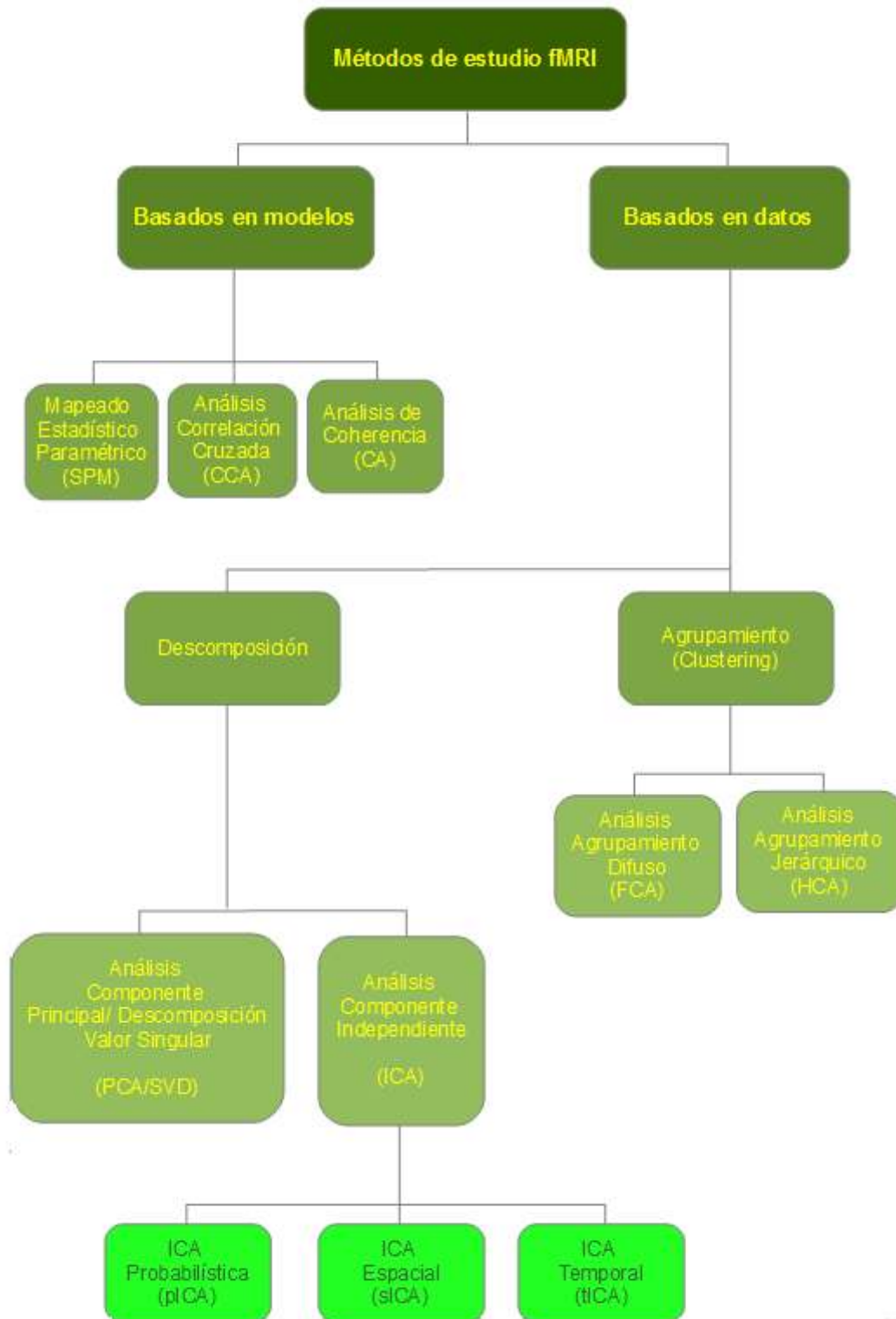
Tras el análisis realizado en base a los datos de la tabla 21, se puede afirmar que la clasificación obtenida tiene sentido desde el punto de vista clínico, aunque esto no se podrá confirmar hasta tener el diagnóstico de los pacientes.

Cabe mencionar que, las muestras en este tipo de estudios, no pueden ser pequeñas, ya que el conjunto de variables con el que se trabaja es elevado por la cantidad de información que se puede extraer de los mapas z de conectividad y que, de antemano, no debemos despreciar. En nuestro caso, se han aplicado técnicas de reducción de la dimensionalidad para solventar este inconveniente.

No se ha podido valorar la bondad de nuestra herramienta al no disponer del diagnóstico de los pacientes incluidos en la muestra. Sin embargo, tras el estudio realizado, considero que la aplicación de técnicas de minería de datos a estudios R-fMRI puede suponer un gran beneficio a este campo. Combinando toda la información que se puede extraer de los mapas z de conectividad funcional con las numerosas técnicas de minería de datos disponibles hasta el momento, se dispone de un gran campo para la investigación, no solo en el ámbito de las enfermedades neurodegenerativas, sino en el estudio de cualquier desorden neurológico. No hay que olvidar que la R-fMRI es una técnica que se puede aplicar desde que el feto se encuentra en el útero, ya que el desarrollo de las redes funcionales en reposo comienzan en la semana 20 de gestación (Schopf et al., 2012) y no supone ningún perjuicio para el paciente al ser una técnica no invasiva y que no produce radiación.

8.- ANEXOS

8.1.- Métodos de análisis de imágenes fMRI



8.2.- Comparativa métodos de análisis de imágenes fMRI

TIPO	MÉTODO	VENTAJAS	INCONVENIENTES
MODELOS	Mapeado Paramétrico Estadístico (SPM)	- Relativamente simple e intuitivo	- Selección de ROIs a priori
	Análisis de Correlación Cruzada (CCA)		- Resultados obtenidos dependientes de la semilla seleccionada
	Análisis de Coherencia (CA)		- Preprocesado complejo para minimizar la influencia de las fuentes no neuronales - Dificultad para establecer patrones de conexiones funcionales del cerebro completo
DATOS	Análisis de la Componente Principal (PCA) Descomposición de valor particular (SVD)	- No requiere la selección de ROIs	- Malos resultados con CNR baja - Requiere seleccionar un número de componentes - Se suele emplear como método de preprocesado para reducir la dimensionalidad y después realizar el análisis con el método ICA
	Análisis de Componente Independiente (ICA)	- No requiere la selección de ROIs - Facilita la comparación entre los resultados obtenidos para diferentes sujetos	- El usuario debe seleccionar manualmente las componentes relevantes del análisis diferenciando los componentes de ruido fisiológico - Asume que las componentes son espacial (sICA) o temporalmente (tICA) independientes - Menos intuitivo que los métodos por modelos
	Análisis Agrupamiento Difuso (FCA)	- No requiere la selección de ROIs	- Elección del número de grupos a constituir - Distancias contaminadas por ruido
	Análisis por Agrupamiento Jerárquico (HCA)	- Se pueden eliminar fácilmente los ruidos - No se requiere la selección de ROIs	- Requiere mucha potencia computacional

9.- BIBLIOGRAFÍA

- Agosta, F., Pievani, M., Geroldi, C., Copetti, M., Frisoni, G. B., & Filippi, M. (2012). *Resting state fMRI in Alzheimer's disease: beyond the default mode network*. *Neurobiology of aging*, 33(8), 1564-1578.
- Anchisi, D., Borroni, B., Franceschi, M., Kerrouche, N., Kalbe, E., Beuthien-Beumann, B., ... & Perani, D. (2005). *Heterogeneity of brain glucose metabolism in mild cognitive impairment and clinical progression to Alzheimer disease*. *Archives of neurology*, 62(11), 1728-1733.
- Bai, F., Watson, D. R., Yu, H., Shi, Y., Yuan, Y., & Zhang, Z. (2009). *Abnormal resting-state functional connectivity of posterior cingulate cortex in amnesic type mild cognitive impairment*. *Brain research*, 1302, 167-174.
- Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. *Investigations into resting-state connectivity using independent component analysis*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360 (1457), 1001–1013.
- Binnewijzend, M. A., Schoonheim, M. M., Sanz-Arigita, E., Wink, A. M., van der Flier, W. M., Tolboom, N., ... & Barkhof, F. (2012). *Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment*. *Neurobiology of aging*, 33(9), 2018-2028.
- Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S., 1995. *Functional connectivity in the motor cortex of resting human brain using echo-planar MRI*. *Magn. Reson. Med.* 34 (4), 537–541.
- Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., ... & Windischberger, C. (2010). *Toward discovery science of human brain function*. *Proceedings of the National Academy of Sciences*, 107(10), 4734-4739.
- Biswal, B. B. (2012). *Resting state fMRI: a personal history*. *Neuroimage*, 62(2), 938-944.
- Blakemore, S. J. (2008). *The social brain in adolescence*. *Nature Reviews Neuroscience*, 9(4), 267-277.
- Bluhm, R. L., Williamson, P. C., Osuch, E. A., Frewen, P. A., Stevens, T. K., Boksman, K., ... & Lanius, R. A. (2009). *Alterations in default network connectivity in posttraumatic stress disorder related to early-life trauma*. *Journal of psychiatry & neuroscience: JPN*, 34(3), 187.
- Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y. Y., Weber, J., & Kober, H. (2011). *Meditation experience is associated with differences in default mode network activity and connectivity*. *Proceedings of the National Academy of Sciences*, 108(50), 20254-20259.

- Bryant, R. A., Felmingham, K. L., Whitford, T. J., Kemp, A. H., Hughes, G., Peduto, A., & Williams, L. M. (2008). *Rostral anterior cingulate volume predicts treatment response to cognitive-behavioural therapy for posttraumatic stress disorder*. Journal of Psychiatry and Neuroscience, 33(2), 142-146.
- Calhoun, V. D., Adali, T., Hansen, L. K., Larsen, J., & Pekar, J. J. (2003). *ICA of functional MRI data: an overview*.
- Chai, X. J., Castañón, A. N., Öngür, D., & Whitfield-Gabrieli, S. (2012). *Anticorrelations in resting state networks without global signal regression*. Neuroimage, 59(2), 1420-1428.
- Chao-Gan, Y., & Yu-Feng, Z. (2010). *DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI*. Frontiers in systems neuroscience, 4.
- Cieslik, E. C., Zilles, K., Caspers, S., Roski, C., Kellermann, T. S., Jakobs, O., ... & Eickhoff, S. B. (2012). *Is there "one" DLPFC in cognitive action control? Evidence for heterogeneity from co-activation-based parcellation*. Cerebral Cortex, bhs256.
- Coifman, R. R., & Lafon, S. (2006). *Diffusion maps. Applied and computational harmonic analysis*, 21(1), 5-30.
- Coricelli, G., & Nagel, R. (2009). *Neural correlates of depth of strategic reasoning in medial prefrontal cortex*. Proceedings of the National Academy of Sciences, 106(23), 9163-9168.
- Crone, J. S., Ladurner, G., Höller, Y., Golaszewski, S., Trinka, E., & Kronbichler, M. (2011). *Deactivation of the default mode network as a marker of impaired consciousness: an fMRI study*. PLoS One, 6(10), e26373.
- Damoiseaux, J.S., Rombouts, S.A., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. *Consistent restingstate networks across healthy subjects*. Proc. Natl. Acad. Sci. U. S. A. 103 (37), 13,848–13,853.
- De Luca, M., Beckmann, C.F., De Stefano, N., Matthews, P.M., Smith, S.M., 2006. *fMRI resting state networks define distinct modes of long-distance interactions in the human brain*. Neuroimage 29 (4), 1359–1367.
- Duncan, K., Curtis, C., & Davachi, L. (2009). *Distinct memory signatures in the hippocampus: intentional States distinguish match and mismatch enhancement signals*. The Journal of Neuroscience, 29(1), 131-139.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). *The human brain is intrinsically organized into dynamic, anticorrelated functional networks*. Proceedings of the National Academy of Sciences of the United States of America, 102(27), 9673-9678.
- Gili, T., Cercignani, M., Serra, L., Perri, R., Giove, F., Maraviglia, B., ... & Bozzali, M. (2011). *Regional brain atrophy and functional disconnection across Alzheimer's disease evolution*. Journal of Neurology, Neurosurgery & Psychiatry, 82(1), 58-66.

- Greicius, M. D., Krasnow, B., Reiss, A. L., Menon, V. (2002). *Functional connectivity in the resting brain: a network analysis of the default mode hypothesis*. Proceeding of the National Academy of Sciences, vol 100 no 1, 253-258].
- Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). *Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI*. Proceedings of the National Academy of Sciences of the United States of America, 101(13), 4637-4642.
- Groen, G., Sokolov, A. N., Jonas, C., Roebing, R., & Spitzer, M. (2011). *Increased resting-state perfusion after repeated encoding is related to later retrieval of declarative associative memories*. PloS one, 6(5), e19985.
- Hafkemeijer, A., van der Grond, J., & Rombouts, S. A. (2012). *Imaging the default mode network in aging and dementia*. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1822(3), 431-441.
- Happé, F., & Frith, U. (Eds.). (2010). *Autism and talent* (Vol. 364, No. 1522). Oxford University Press.
- Harrison, B. J., Pujol, J., Ortiz, H., Fornito, A., Pantelis, C., & Yücel, M. (2008). *Modulation of brain resting-state networks by sad mood induction*. PLoS One,3(3), e1794.
- Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). *Introducción a la Minería de Datos*. Editorial Pearson Educación SA, Madrid.
- Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M., & Engel, A. K. (2012). *Large-scale cortical correlation structure of spontaneous oscillatory activity*. Nature neuroscience, 15(6), 884-890.
- Lee, M. H., Smyser, C. D., & Shimony, J. S. (2012). *Resting-State fMRI: A Review of Methods and Clinical Applications*. American Journal of Neuroradiology.
- Lemogne, C., Gorwood, P., Bergouignan, L., Pélioso, A., Lehericy, S., & Fossati, P. (2011). *Negative affectivity, self-referential processing and the cortical midline structures*. Social cognitive and affective neuroscience, 6(4), 426-433.
- Lemogne, C., Delaveau, P., Freton, M., Guionnet, S., & Fossati, P. (2012). *Medial prefrontal cortex and the self in major depression*. Journal of affective disorders, 136(1), e1-e11.
- Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). *Review of methods for functional brain connectivity detection using fMRI*. Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society,33(2), 131.
- Li, X., Cao, M., Zhang, J., Chen, K., Chen, Y., Ma, C., ... & Zhang, Z. (2014). *Structural and Functional Brain Changes in the Default Mode Network in Subtypes of Amnesic Mild Cognitive Impairment*. Journal of geriatric psychiatry and neurology, 0891988714524629.
- Lynch, C. J., Uddin, L. Q., Supekar, K., Khouzam, A., Phillips, J., & Menon, V. (2013). *Default mode network in childhood autism: posteromedial cortex heterogeneity and relationship with social deficits*. Biological psychiatry, 74(3), 212-219.

- Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). *The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?*. *Neuroimage*, 44(3), 893-905.
- Otsuka, Y., Osaka, N., Morishita, M., Kondo, H., & Osaka, M. (2006). *Decreased activation of anterior cingulate cortex in the working memory of the elderly*. *Neuroreport*, 17(14), 1479-1482.
- Pizoli, C. E., Shah, M. N., Snyder, A. Z., Shimony, J. S., Limbrick, D. D., Raichle, M. E., ... & Smyth, M. D. (2011). *Resting-state activity in development and maintenance of normal brain function*. *Proceedings of the National Academy of Sciences*, 108(28), 11638-11643.
- Prieto, C., Eimil, M., López de Silanes, C., Llanero, M. (2011). *Impacto social de la Enfermedad de Alzheimer y otras demencias 2011*. La fundación del cerebro.
- Qi, Z., Wu, X., Wang, Z., Zhang, N., Dong, H., Yao, L., & Li, K. (2010). *Impairment and compensation coexist in amnesic MCI default mode network*. *Neuroimage*, 50(1), 48-55.
- Richards, J., & Richards, M. J. (2014). Package 'diffusionMap'.
- Rombouts, S. A., Damoiseaux, J. S., Goekoop, R., Barkhof, F., Scheltens, P., Smith, S. M., & Beckmann, C. F. (2009). *Model-free group analysis shows altered BOLD FMRI networks in dementia*. *Human brain mapping*, 30(1), 256-266.
- Salgado-Pineda, P., Fakra, E., Delaveau, P., McKenna, P. J., Pomarol-Clotet, E., & Blin, O. (2011). *Correlated structural and functional brain abnormalities in the default mode network in schizophrenia patients*. *Schizophrenia research*, 125(2), 101-109.
- Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., Bullmore, E., 2005a. *Neurophysiological architecture of functional magnetic resonance images of human brain*. *Cereb. Cortex* 15 (9), 1332–1342.
- Schöpf, V., Kasprian, G., Brugger, P. C., & Prayer, D. (2012). *Watching the fetal brain at 'rest'*. *International Journal of Developmental Neuroscience*, 30(1), 11-17.
- Snyder, A. Z., & Raichle, M. E. (2012). *A brief history of the resting state: the Washington University perspective*. *Neuroimage*, 62(2), 902-910.
- Sorg, C., Riedl, V., Mühlau, M., Calhoun, V. D., Eichele, T., Läer, L., ... & Wohlschläger, A. M. (2007). *Selective changes of resting-state networks in individuals at risk for Alzheimer's disease*. *Proceedings of the National Academy of Sciences*, 104(47), 18760-18765.
- Van Den Heuvel, M. P., Mandl, R. C., Kahn, R. S., Pol, H., & Hilleke, E. (2009). *Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain*. *Human brain mapping*, 30(10), 3127-3141.
- Van Den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). *Exploring the brain network: a review on resting-state fMRI functional connectivity*. *European Neuropsychopharmacology*, 20(8), 519-534.

- Wang, L., Zang, Y., He, Y., Liang, M., Zhang, X., Tian, L., ... & Li, K. (2006). *Changes in hippocampal connectivity in the early stages of Alzheimer's disease: evidence from resting state fMRI*. *Neuroimage*, 31(2), 496-504.
- Wang, Z., Liang, P., Jia, X., Jin, G., Song, H., Han, Y., ... & Li, K. (2012). *The baseline and longitudinal changes of PCC connectivity in mild cognitive impairment: a combined structure and resting-state fMRI study*. *PloS one*, 7(5), e36838.
- Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., ... & Masuda, N. (2013). *A pairwise maximum entropy model accurately describes resting-state human brain networks*. *Nature communications*, 4, 1370.
- Yan, H., Zhang, Y., Chen, H., Wang, Y., & Liu, Y. (2013). *Altered effective connectivity of the default mode network in resting-state amnesic type mild cognitive impairment*. *Journal of the International Neuropsychological Society*, 19(04), 400-409.
- Zang, Y., Jiang, T., Lu, Y., He, Y., & Tian, L. (2004). *Regional homogeneity approach to fMRI data analysis*. *Neuroimage*, 22(1), 394-400.
- Zang YF, He Y, Zhu CZ, Cao QJ, Sui MQ, Liang M, Tian LX, Jiang TZ, Wang YF. (2007) Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev* 2007;29:83–91.
- Zhang, H. Y., Wang, S. J., Xing, J., Liu, B., Ma, Z. L., Yang, M., ... & Teng, G. J. (2009). *Detection of PCC functional connectivity characteristics in resting-state fMRI in mild Alzheimer's disease*. *Behavioural brain research*, 197(1), 103-108.
- Zuo, X. N., Di Martino, A., Kelly, C., Shehzad, Z. E., Gee, D. G., Klein, D. F., ... & Milham, M. P. (2010). *The oscillating brain: complex and reliable*. *Neuroimage*, 49(2), 1432-1445.

9.1.- Fuentes electrónicas

- Calhoun, V. D. (2004). *Group ICA of fMRI toolbox (GIFT)*. Online at <http://icatb.sourceforge.net>.
- Devlin, H. (2007). What is Functional Magnetic Resonance Imaging (fMRI)?. *Psych Central*. Retrieved on August 31, 2014, from <http://psychcentral.com/lib/what-is-functional-magnetic-resonance-imaging-fmri/0001056>
- Fundación del cerebro. (n. d.). Consultado el 4 de agosto de 2014. <http://http://www.fundaciondelcerebro.es/>
- Hall, M., Consultado el 23 de julio de 2014. <http://weka.sourceforge.net/doc.dev/weka/attributeSelection/BestFirst.html>.
- Wikipedia. (2014). Definición de demencia. Consultado el 4 de agosto de 2014. <http://es.wikipedia.org/wiki/Demencia>
- Wikipedia. (2014). Definición de Enfermedad de Alzheimer. Consultado el 4 de agosto de 2014. http://es.wikipedia.org/wiki/Enfermedad_de_Alzheimer
- Metereau, E., & Dreher, J. C. (2013). Consultado el 15 de julio de 2014. http://cercor.oxfordjournals.org/content/suppl/2012/01/28/bhs037.DC1/BHS037_Supplementary_material.pdf