

UNIVERSITA' DEGLI STUDI DI PADOVA

CORSO DI LAUREA IN BIOINGEGNERIA

TESI DI LAUREA MAGISTRALE

The seal of the University of Padua is a large, circular emblem in the background. It features a central shield with two figures: a woman on the left and a man on the right. Above the shield are three stars. The shield is flanked by two columns. The entire seal is surrounded by a circular border containing the Latin text 'UNIVERSITAS STUDII PADOVAE' and the year 'MCCXXII' at the bottom.

***CARATTERIZZAZIONE DELLE
METRICHE DI GRAFO APPLICATE
ALLE MAPPE DI CONNETTIVITA'
FUNZIONALE IN RESTING-STATE***

Relatore: ***Chiar.ma Prof.ssa Alessandra Bertoldo***

Correlatori: ***Ing. Erica Silvestri***

Ing. Alessandro Palombit

Laureanda: ***Francesca Digiacomo***

Anno accademico 2016/2017

Indice

1. INTRODUZIONE	1
2. STATO DELL'ARTE.....	5
2.1 BASI DI RISONANZA MAGNETICA FUNZIONALE	5
2.2 RESTING-STATE NETWORKS.....	7
2.3 COSTRUZIONE DELLE NETWORK CEREBRALI	9
2.4 TEORIA DEI GRAFI	11
2.5 METRICA DI GRAFO	12
3. MATERIALI E METODI.....	19
3.1 DATA-SET E PREPROCESSING	19
3.2 TEST PRELIMINARI SULLA CORRELAZIONE FUNZIONALE	21
3.3 METRICHE DI GRAFO	24
3.3.1 Correlazioni negative	24
3.3.2 Caratterizzazione metriche di grafo con indice di Entropia	24
3.3.3 Caratterizzazione metriche di grafo tramite survival plot.....	26
3.3.4 Specifiche tecniche	28
4. RISULTATI.....	31
4.1. TEST PRELIMINARI	31
4.1.1 Analisi delle distribuzioni individuali.....	31
4.1.2 Connettività Funzionale: impatto dei valori negativi di correlazione	35
4.1.3.CONNETTIVITÀ FUNZIONALE: IMPATTO DEI VALORI DI CORRELAZIONE POSITIVI	37
4.1.4 Connettività Funzionale al variare della soglia	39
4.1.5 Considerazioni Complessive sui Test Preliminari.....	41
4.2. METRICHE DI GRAFO	42
4.2.1 Entropia di metrica di grafo vs Entropia Connettività Funzionale	43
4.2.2 Survival Plot.....	46
4.2.3 Analisi topologica dei grafi.....	52
4.2.4 Analisi robustezza e sensibilità complessiva.....	55
5. DISCUSSIONE.....	59
5.1 VARIABILITÀ INTRA-SOGGETTO E INTER-SOGGETTO.....	59
5.2 IMPATTO DELLA SCALA NUMERICA	60
5.3 IMPATTO DELLA SOGLIA PER GRAFI PESATI E NON	61
5.4 IMPATTO DEI VALORI NEGATIVI DI CORRELAZIONE	63
5.5 IMPATTO DEI SOFTWARE	63
6. CONCLUSIONE	65

<i>6.6 SVILUPPI FUTURI</i>	67
<i>APPENDICE</i>	69
<i>ELENCO DEGLI ACRONIMI</i>	69
<i>ELENCO DELLE FIGURE</i>	71
<i>ELENCO DELLE TABELLE</i>	73

Capitolo 1

Introduzione

La risonanza magnetica funzionale, *fMRI*, in condizioni di *resting-state*, come pratica clinica, nasce all'incirca 15 anni dopo l'invenzione della *fMRI* in risposta a task specifici [4]. La condizione di *resting-state* è un particolare stato mentale che prevede il completo rilassamento dell'attività neurale durante la quale le connessioni cerebrali non sono impegnate in alcuna funzione. Proprio in questo particolare status è possibile esaminare l'attività cerebrale spontanea che si verifica nelle varie aree neurali [5]. Ad ora, infatti, gli studi di *fMRI* in *resting-state* hanno lo scopo di mappare la connettività funzionale tra le aree neuronali di ogni soggetto, attraverso la misura del livello delle loro dinamiche spontanee. La connettività funzionale, infatti, è definita come la dipendenza temporale tra eventi neurofisiologici distanti a livello spaziale [4]. Nei vari esperimenti che sono stati svolti negli anni, ad esempio dal gruppo di ricerca Biswal *et al.* [5], si è rilevata la presenza di reti fortemente connesse durante la condizione di riposo. Le network, finora identificate, sono otto tra cui, la più nota ed indagata, *default mode network*, DMN, contenente: il precuneo, la regione mediale frontale, parietale inferiore e temporale e la corteccia temporale bilaterale e anteriore cingolata[6]. Tale rete costituisce il fulcro per l'insorgenza di molteplici malattie neurodegenerative. In particolare, è noto da letteratura che le disfunzioni dell'attività cerebrale in cui sia coinvolto il DMN interessano patologie come Parkinson e Alzheimer e disturbi neuropsichiatrici come demenza, depressione e schizofrenia[7]. Risulta, quindi, di fondamentale importanza la ricerca a riguardo.

Le modalità di studio delle reti neurali sono molteplici ma in questo lavoro si è scelto di affrontare l'analisi attraverso la teoria dei grafi. Sebbene l'introduzione del concetto matematico di grafo sia nato nel 1736 grazie ad Eulero[8], la sua applicazione nei confronti dell'area cerebrale è di recente sviluppo e di notevole interesse nell'ambito della ricerca. In particolare, la teoria dei grafi si presta bene a descrivere le rete cerebrali se associata

all'analisi dei dati di *fMRI* basata sulle regioni di interesse, altrimenti note come ROI. Le posizioni delle ROI sono individuate tramite parcellizzazione delle aree cerebrali secondo l'utilizzo di atlanti neurali specifici e statistici. Infatti, in questo modo è possibile identificare i nodi del grafo che rappresentano le ROI e i link che descrivono le connessioni cerebrali tra le ROI, pesati sulla base della forza delle connessioni all'interno della network.

Da letteratura sono numerose le metriche, fino adesso sviluppate e indagate, con cui caratterizzare tali network ma manca una chiara e definitiva caratterizzazione topologica delle stesse come anche una tecnica decisiva per determinare robuste mappe di connettività funzionale. Motivo per cui, durante il lavoro di tesi, si è cercato di apportare un contributo allo stato attuale della ricerca attraverso lo studio delle reti neurali con l'utilizzo della teoria dei grafi. L'obiettivo della tesi, infatti, è stato di caratterizzare le metriche di grafo applicate alle mappe di connettività funzionale in *resting-state* per soggetti di controllo, al variare dei valori di soglia modulata sulla significatività delle mappe a livello fisiologico e sull'attendibilità delle misure calcolate.

In primo luogo, dunque, si è svolta l'analisi preliminare del contenuto informativo delle mappe di connettività funzionale dei soggetti con lo scopo di comprendere quale fosse l'informazione contenuta nelle matrici e se la struttura fondamentale della rete fosse condivisa tra i soggetti. Sono state, dunque, usate le tecniche standard di analisi dei dati relativi ad *fMRI*, come approfondito in par.3.3.

Successivamente si è testata e modellata una prima metodologia per la caratterizzazione delle metriche di grafo basata sull'indice di entropia calcolato sia per le mappe di connettività sia per le metriche di grafo e sotto diverse ipotesi di lavoro, approfondito in par 3.4.2.

Le misure di grafo sono state caratterizzate attraverso *survival plot*[9], grazie al confronto diretto del trend entropico relativo alla connettività e al comportamento delle singole misure. E' stato applicato un intervallo di soglia assoluta operando tagli netti sulle singole reti cerebrali dei soggetti, al fine di osservare il loro trend in relazione al continuo depauperamento delle network. Inoltre, l'implementazione delle metriche è stato svolto utilizzando due software: *Brain Connectivity Toolbox*(BCT) [10], più usato in letteratura e *GraPh thEoretical Network Analysis* (GRETNA), di recente sviluppo [1]. Sono stati valutati, quindi, l'impatto del cambiamento dei valori di soglia applicati alle singole metriche entro criteri fisiologici, l'impatto della scala numerica sulle mappe individuali, l'impatto del software utilizzato e l'influenza della variabilità intra-soggetto, esposti nel par.3.3.3.

Le analisi sono state, in definitiva, riproposte attraverso un numero maggiore di soggetti valutandone l'impatto a livello di rete neurale e rappresentando visivamente l'aspetto topologico della rete al variare della soglia, come proposto nel par.4.2

Capitolo 2

Stato dell'arte

Nel 1736 Eulero introduce per primo il concetto di *grafo*, durante la risoluzione del problema del ponte di Königsberg. Successivamente nasce la disciplina di teoria dei grafi, come parte integrante della matematica discreta. Da allora, ogni tipologia di rete presente nel mondo reale è descritta da grafi, perfino la rete cerebrale. In questo capitolo, infatti, si vuole sviluppare il percorso teorico che stabilisce le fondamenta per capire come dalla risonanza magnetica funzionale *in resting-state*, attraverso la misura di connettività funzionale tra regioni cerebrali specifiche, si possa giungere a rappresentare un rete cerebrale con il concetto di grafo e delinearla mediante metriche descrittive di facile valutazione ed applicazione.

2.1 *Basi di risonanza magnetica funzionale*

La risonanza magnetica funzionale, *fMRI* è una tecnica di *neuroimaging*, non invasiva, volta a misurare la risposta emodinamica associata all'attività neurale di specifiche aree cerebrali. In dettaglio, una funzione cognitiva, quando è in corso, richiede un alto consumo di ossigeno che si riflette in cambiamenti locali dei livelli di ossigenazione dell'emoglobina nel flusso sanguigno e, di conseguenza, comporta l'aumento della concentrazione di emoglobina deossigenata. La risposta emodinamica, nota come segnale BOLD, *Blood Oxygenation Level Dependent*, è indicatrice delle variazioni di concentrazione di emoglobina ossigenata e deossigenata. Sebbene non sia una misura diretta, riflette il consumo di ossigeno durante un'attività neurale, di conseguenza, evidenzia l'attivazione e il coinvolgimento di specifiche aree cerebrali. La *fMRI*, dunque, sfrutta l'effetto BOLD per ricavare immagini funzionali pesate via parametro T_2^* ed è in grado di mappare l'attività

cerebrale senza l'utilizzo di mezzi di contrasto esogeni. Il parametro T_2^* rappresenta il tempo di rilassamento spin-spin comprensivo sia dei contributi dovuti alle interazioni molecolari sia alle disomogeneità del campo magnetico. La presenza di emoglobina deossigenata, infatti, genera disomogeneità del campo magnetico con conseguenti diminuzione del parametro T_2^* . Tale diminuzione, infine, è strettamente legata all'attività cerebrale. Tramite *fMRI*, quindi, è possibile misurare la coattivazione di specifiche regioni in termini di correlazione temporale.

Nasce, così, la connettività funzionale definita come dipendenza temporale tra eventi neurofisiologici distanti a livello spaziale[4]. Nel contesto della *neuroimaging* funzionale, essa descrive la relazione tra pattern di attivazione neuronale di regioni cerebrali anatomicamente separate. Abitualmente, l'analisi della correlazione di Pearson permette di esaminare se diverse regioni variano simultaneamente il loro stato di attivazione e in che modo avviene il coinvolgimento delle aree cerebrali nello svolgere una certa funzione. Di norma, i valori di correlazione sono compresi nel range [-1,1], dove il segno positivo rappresenta l'attivazione simultanea delle aree interessate, il segno negativo indica l'attivazione inversa delle regioni, eventuali correlazioni nulle, invece, identificano aree temporalmente indipendenti le une dalle altre. In generale, lo studio della connettività funzionale consiste nel misurare la correlazione dell'andamento del segnale nel tempo in diverse regioni cerebrali.

Un metodo molto comune di analisi statistica delle immagini funzionali si basa sulla scelta di una o più Regioni cerebrali Di Interesse, ROI, tra cui misurare le correlazioni delle serie temporali. Solitamente le ROI possono essere definite in termini di caratteristiche funzionali e strutturali[11]. Dal punto di vista strutturale, sono individuate anatomicamente a livello macroscopico. Infatti, per eseguire tale tipologia di analisi è necessario effettuare la localizzazione e la mappatura delle aree cerebrali di interesse sulla superficie encefalica, utilizzando atlanti anatomici precostituiti. In particolare, si parla di parcellizzazione cerebrale, suddivisione intera o parziale della superficie cerebrale, in aree specifiche. In molti casi il modo migliore è di definire le ROI per ogni soggetto, sulla base della propria struttura anatomica, dal momento che esiste una sostanziale variabilità tra gli individui a livello macroscopico[11]. L'utilizzo di atlanti anatomici basati sul singolo soggetto come *AAL atlas* o *Talairach atlas* richiede particolare attenzione data l'incapacità di far coincidere perfettamente la struttura cerebrale tra individui diversi e di eseguire un'appropriata normalizzazione spaziale. Tuttavia la mancanza di sovrapposizione tra ogni gruppo di soggetti e questi atlanti non è sostanziale[11]. Recentemente, si sono preferiti

atlanti probabilistici di anatomia macroscopica o atlanti probabilistici basati sulle aree di *Brodmann* che sono disponibili nei toolbox per l'analisi delle ROI, tra cui si ricordano SPM e FSL. In conclusione, l'analisi basata su ROI risulta vantaggiosa perchè permette di esplorare i dati di una singola regione riducendo la complessità delle analisi, per determinare le regioni anche in mancanza di attivazione e risulta adatta per la *fMRI* in *resting state*, descritta successivamente.

2.2 *Resting-state networks*

All'incirca 15 anni dopo l'invenzione della risonanza magnetica funzionale, molti ricercatori hanno iniziato ad esaminare la connettività funzionale tra regioni cerebrali come il livello di coattivazione delle serie temporali spontanee, registrate durante il riposo[4]. Questo tipo di analisi ha preso il nome di risonanza magnetica funzionale *in resting-state*, *R-fMRI*[12]. Durante gli esperimenti, ai volontari è stato richiesto di rilassarsi senza pensare a qualcosa in particolare, mentre il loro livello di attività cerebrale spontanea veniva misurata. Biswal *et al.* [5] sono stati i primi a dimostrare che le regioni emisferiche destra e sinistra della corteccia motoria primaria non erano silenti a riposo, invece, mostravano un'alta correlazione tra le loro serie temporali dei segnali BOLD, suggerendo la presenza di attività cerebrale anche a riposo. Anche altri studi hanno replicato gli stessi risultati pionieristici mostrando una forte attività cerebrale anche nella rete uditiva e visiva[13]. Ad oggi, gli studi di *fMRI in resting-state* hanno l'obiettivo di mappare i canali di comunicazione funzionale tra le aree neuronali misurando il livello delle loro dinamiche spontanee. E' ben noto, infatti, che i neuroni mostrano un alto livello di *firing* spontaneo in assenza di task esterni, trasportando continuamente informazione ad altri neuroni[14].

Durante i vari esperimenti, alcuni gruppi di studio hanno rilevato la presenza di reti fortemente connesse durante la condizione di riposo note come *resting-state network* *DMD* [6]. Consistono in regioni interconnesse, integrate funzionalmente ma separate anatomicamente che evidenziano un alto livello di connettività funzionale. Fin ad ora, sono state identificate all'incirca otto reti funzionalmente connesse. Esse includono la rete motoria, visiva, le regioni superiori parietali e frontali e la cosiddetta *default mode network*, *DMN*, contenente: il precuneo, la regione mediale frontale, parietale inferiore e temporale e la corteccia temporale bilaterale e anteriore cingolata (**Figura 1**). Di rilievo è il *default mode network* in quanto evidenzia un livello elevato di attività neuronale a riposo, piuttosto che in

risposta a task esterni, suggerendo che l'attività di questa rete riflette uno stato di default dell'attività neurale nel cervello umano[15]. L'attività e la connettività del DMN formano il cuore della cognizione umana e includono l'integrazione dei processi emozionali e cognitivi, di monitoraggio del mondo circostante e di divagazione con la mente. Tali attività e connettività sono di speciale interesse per esaminare disfunzioni neurologiche cognitive in disordini cerebrali e psichiatrici[16].

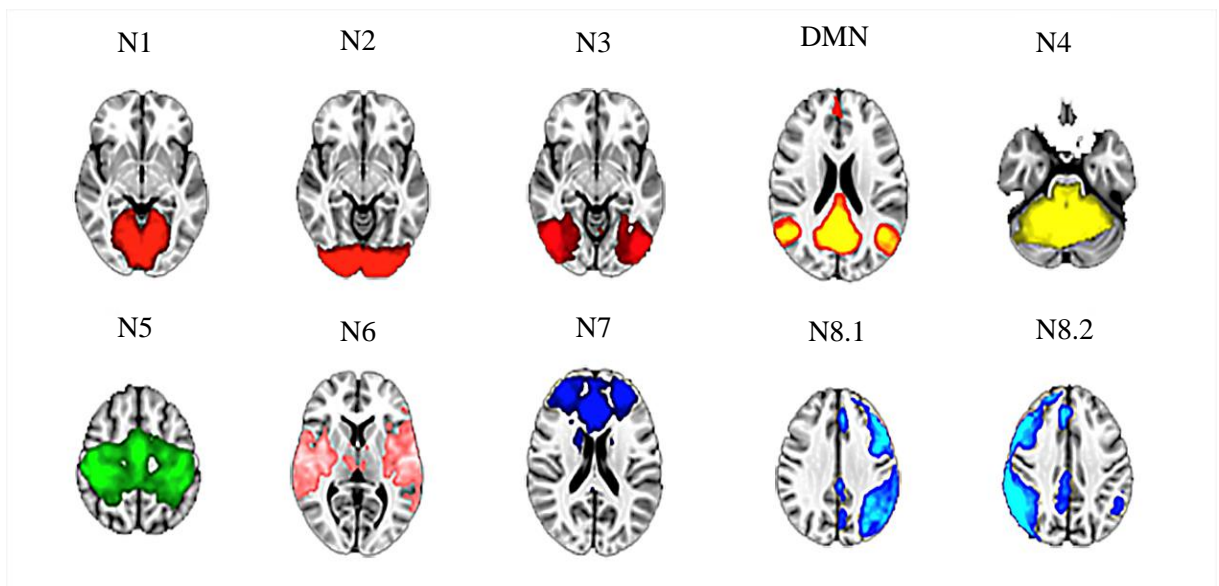


Figura 1 Resting-state network: *N1_network_visivo_1; N2_network_visivo_2; N3_network_visivo_3; DMN_default_mode_network; N4_network_cerebellare; N5_network_sensomotorio; N6_network_uditivo; N7_network_controllo_esecutivo; N8.1_network_frontoparietale_sinistro; N8.2_network_frontoparietale_destro (modificata[3]).*

Come descritto in precedenza, la *fMRI* offre la possibilità di registrare all'unisono l'attivazione di più regioni cerebrali, permettendo lo studio di reti neurali. Generalmente, è possibile esaminare la rete cerebrale con molteplici tecniche, tra le più recenti ed avanzate si annovera l'analisi dei grafi[5]. E' necessario, quindi, ricavare le network cerebrali a partire dalle dinamiche temporali di connettività funzionale. Di seguito è descritta la procedura di costruzione delle reti.

2.3 *Costruzione delle network cerebrali*

Una network è una rappresentazione matematica di un sistema reale e complesso definito dalla collezione di nodi e collegamenti o link tra di essi [10]. I nodi all'interno delle network, di solito, rappresentano le regioni cerebrali, mentre i collegamenti possono rappresentare le connessioni funzionali, anatomiche o effettive, in dipendenza al dataset da analizzare. Nel caso di connessioni funzionali, esse corrispondono al valore delle correlazioni temporali in termini di attività cerebrale e possono verificarsi tra coppie di regioni anatomicamente non collegate. La costruzione della network a partire dalle immagini di R-fMRI segue quattro fondamentali stadi[17]. Come mostra la **Figura 2.2** in primo luogo, dall'esame di R-fMRI, si estraggono le immagini funzionali[2]. Successivamente, si esegue l'analisi delle immagini basate sulle ROI. E' necessaria, quindi, la parcellizzazione anatomica dell'encefalo in linea con atlanti anatomici in modo da mappare le ROI. In questo modo si possono estrarre le serie temporali delle dinamiche cerebrali relative alle aree cerebrali. Di conseguenza è generata la matrice di connettività associata alla rete cerebrale, in quanto qualsiasi sia la tipologia della network essa è rappresentata dalla propria matrice di connettività o adiacenza. Mentre le matrici di connettività binaria o adiacenza indicano soltanto la presenza o assenza di connessione tra le ROI, le matrici di connettività per grafi pesati rappresentano il pattern di cross-correlazione tra i segnali BOLD stimati dalle dinamiche temporali[17]. Le righe e le colonne delle matrici denotano i nodi, i valori di correlazione, costituenti le matrici invece, denotano i collegamenti[17]. A questo livello della costruzione è importante la scelta di una soglia perché permette di generare grafi con densità di connessioni differenti[10]. Una volta costruito il grafo, quindi, si può eseguire la quantificazione di molteplici metriche e parametri per l'analisi della network.

In definitiva la natura dei nodi e dei link nelle reti neurali individuali è determinata dalla combinazione dei metodi di mappatura cerebrale, gli schemi di parcellizzazione e le misure di connettività[18]. Di solito la scelta di una specifica combinazione deve essere attentamente motivata, dal momento che determina la natura dei nodi, dei link e l'interpretazione neurobiologica della topologia della network [19]. La parcellizzazione, ad esempio, può influenzare le misure calcolate e anche la possibilità di confrontare tra loro le reti funzionali[17], dal momento che soltanto nel caso in cui esse condividono lo stesso schema di mappatura possono essere comparate tra loro. Inoltre, in molti studi di elettroencefalografia e di magnetoencefalografia, i nodi sono equivalenti alle posizioni dei sensori o degli elettrodi usati durante le acquisizioni, ma le network possono basarsi su

sorgenti anatomiche ricostruite[17]. Durante l'implementazione delle misure neurali l'ordine dei nodi nelle matrici di connettività non ha effetto ma è molto importante per la visualizzazione delle network[10].

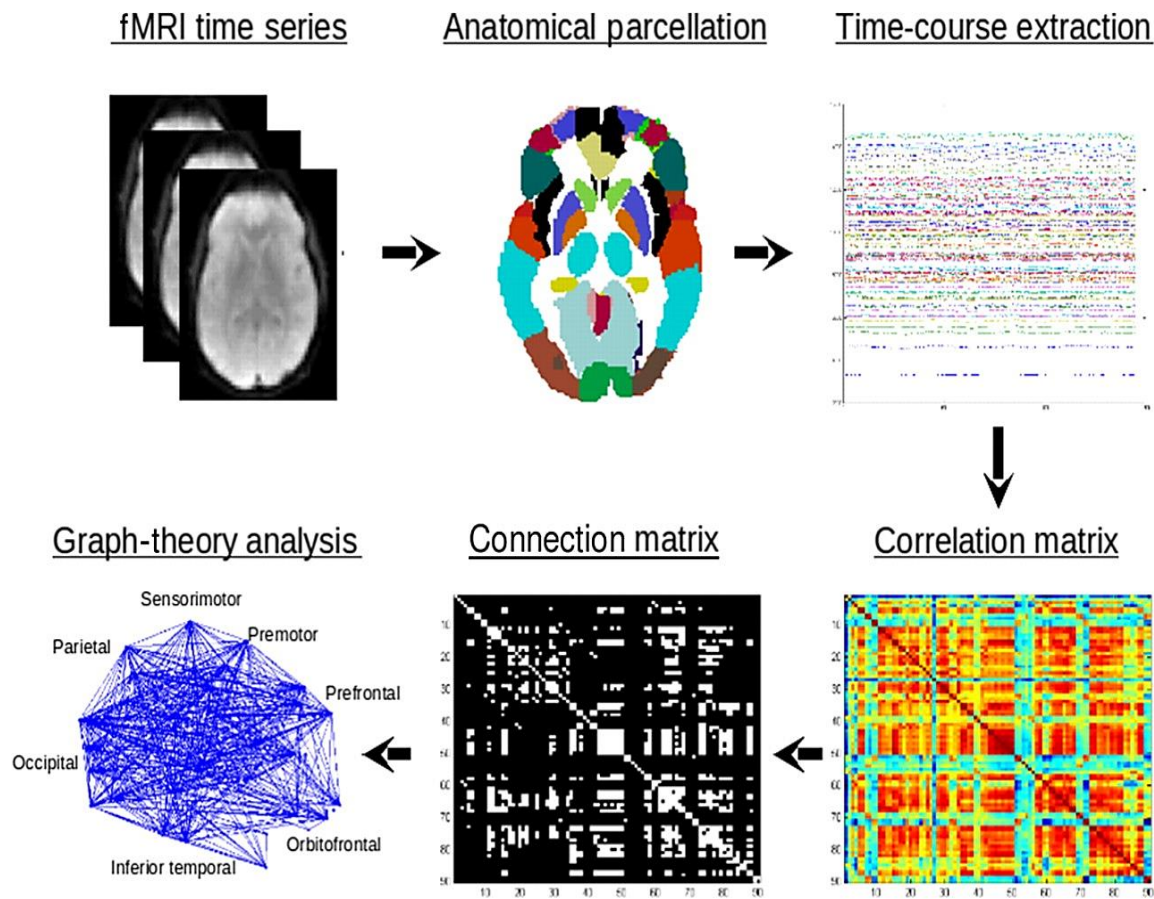


Figura 2.2 Diagramma di flusso relativo alla costruzione della network cerebrale: 1.Acquisizione immagini di R-fMRI, 2.Parcellizzazione anatomica per analisi basata su ROI, 3.Estrazione delle serie temporali di connettività funzionale tra le regioni cerebrali, 4.Generazione delle matrici di connettività, 5.Analisi del grafo tramite metriche descrittive [2]

2.4 Teoria dei grafi

Formalmente un grafo rappresenta in termini matematici una rete complessa del mondo reale. Matematicamente un grafo $G = (N, L)$ consiste in una coppia di due insiemi: N , l'insieme dei nodi e L , l'insieme degli archi che connettono coppie di nodi. Gli elementi di $N = \{n_1, n_2, \dots, n_N\}$ sono detti nodi (o vertici) del grafo G e rappresentano le unità fondamentali del sistema, con $N \neq \emptyset$, mentre gli elementi di $L = \{l_1, l_2, \dots, l_K\}$ sono detti archi o links e rappresentano le connessioni tra coppie di nodi ossia le relazioni tra gli elementi del sistema [8]. N è il numero di nodi ed indica la dimensione del grafo mentre K è il numero degli elementi di L ossia il numero di archi. Un nodo è indicato con la sua posizione i -esima nell'insieme N , mentre un arco, definito dalla coppia di nodi i e j che connette, si indica come l_{ij} . Quando due nodi sono connessi da un link sono detti adiacenti o vicini.

Generalmente, gli archi si distinguono a seconda della direzione in cui connettono i nodi e a seconda della loro intensità. Per quanto riguarda la direzione degli archi, si distinguono:

- *arco orientato*: se gli elementi di L sono connessioni tra coppie di nodi ordinati e specifici che non possono essere inter-scambiati. In questo caso, infatti, l_{ij} indica un arco diretto da i e j e $l_{ij} \neq l_{ji}$,
- *arco non orientato*: se gli elementi di L sono link tra coppie non ordinate di nodi: in questo caso, non esiste un'orientazione privilegiata e $l_{ij} = l_{ji}$.

A seconda dell'intensità delle connessioni tra coppie di nodi, si riconoscono gli archi:

- *Non pesato o binario*: indica lo stato di connessione tra due nodi, espressa come presenza o assenza del link. Il valore pari a 1 esprime la presenza della connessione, lo 0 individua l'assenza del link tra due nodi. Di conseguenza, non esiste una scala di importanza tra i vari link ma, quando presenti, sono tutti equivalenti tra loro.
- *Pesato*: a ciascun arco è attribuito un numero reale w_{ij} che determina il suo peso ossia la forza della connessione tra due nodi, rendendo possibile la distinzione tra legami deboli e legami più forti.

Le combinazioni delle caratteristiche, sopra riportate, identificano quattro tipologie di grafi possibili:

- *Grafo non pesato e non orientato*
- *Grafo non pesato e orientato*

- *Grafo pesato e non orientato*
- *Grafo pesato e orientato*

In dettaglio, si definisce grafo pesato l'unione di tre insiemi $G = (N, L, W)$ con N l'insieme degli N nodi e non vuoto, L è l'insieme degli archi e $W = \{w_1, w_2, \dots, w_K\}$, l'insieme dei valori reali dei pesi associati a ciascun arco. Un grafo pesato è denotato con G^W [8]. Solitamente la sua rappresentazione grafica mostra i link con diverso spessore, in modo da evidenziare l'entità dei loro pesi. Inoltre, il segno del valore del peso descrive la tipologia di interazione tra i nodi. Il segno negativo identifica coattivazione di natura distruttiva dei nodi, il segno positivo individua l'interazione costruttiva. Applicando il concetto di grafo alle reti cerebrali, le singole ROI costituiscono i nodi del grafo, interconnesse tramite link pesati, i cui pesi sono rappresentati dai valori della connettività funzionale.

2.5 *Metrica di grafo*

Per ogni singolo grafo è possibile quantificare diverse misure, che in base alle proprietà che quantificano si distinguono in quattro macro-categorie:

- ***Misure di centralità:*** Analizzano il ruolo preponderante che alcuni nodi assumono all'interno della rete, classificandoli a seconda della centralità che detengono nella gestione della trasmissione del segnale tra una regione e un'altra della network.
- ***Misure di integrazione:*** Quantificano l'efficienza del grafo nel trasmettere informazioni a livello globale, anche tra nodi con lunghi cammini, definiti distanti.
- ***Misure di segregazione:*** Stabiliscono l'influenza sulla connettività della rete da parte di insiemi di nodi adiacenti rispetto ad un particolare nodo e quantificano la tendenza alla formazione di regioni popolate da nodi fortemente interconnessi tra loro e poco con gli altri nodi della rete.
- ***Misure di elasticità:*** Quantificano le caratteristiche anatomiche che riflettono la vulnerabilità della rete ad attacchi esterni. Molti studi, infatti, hanno evidenziato la capacità di una neuropatologia di ledere l'attività delle funzioni cerebrali, ad esempio l'ictus[10].

Per ogni categoria si sono scelte alcune metriche, oggetto di approfondimento durante il lavoro di tesi[10]. Come misura di centralità si è scelta la *Betweenness Centrality*, tra le misure di integrazione: il *Char Path* e la *Global Efficiency*, delle misure di segregazione: il *Clustering Coefficient*, la *Modularity* e la *Local Efficiency* e come misura di elasticità: la *Strength*. Di seguito sono riportate le loro definizioni e descrizioni in riferimento a grafi non orientati e non pesati e sono menzionati gli opportuni cambiamenti per grafi pesati.

Betweenness Centrality

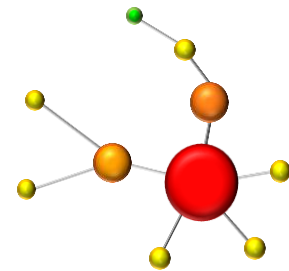
$$b = \frac{1}{(n-1)(n-2)} \sum_{\substack{h,j \in N \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}}$$

n : numero dei nodi

N : set dei nodi

ρ_{hj} : percorsi più brevi tra h e j

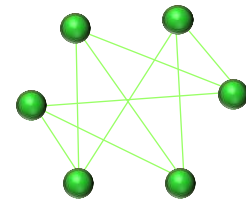
$\rho_{hj}(i)$: numero di percorsi più brevi tra h e j ,
passanti per i



Global Efficiency

$$E = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n-1}$$

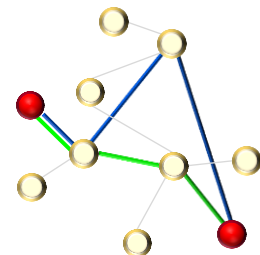
$$E_i = \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n-1} : \text{efficienza del nodo } i$$



Char Path

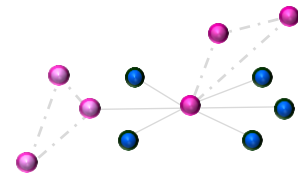
$$L = \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n-1}$$

$$L_i = \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n-1} : \text{efficienza del nodo } i$$



Clustering
Coefficient

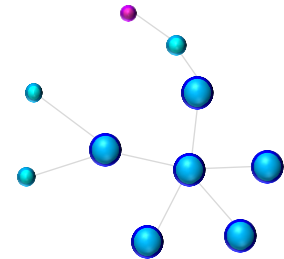
$$C = \sum_{i \in N} C_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)}$$



C_i : clustering coefficient del nodo i

Local
Efficiency

$$E_{loc} = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j, h \in N, j \neq i} a_{ij} a_{ih} [d_{jh}(N_i)]^{-1}}{k_i(k_i - 1)}$$



$$E_{loc,i} = \frac{\sum_{j, h \in N, j \neq i} a_{ij} a_{ih} [d_{jh}(N_i)]^{-1}}{k_i(k_i - 1)}$$

$E_{loc,i}$: efficienza locale del nodo i

$d_{jh}(N_i)$: è la lunghezza del percorso più corto
che contiene solo i vicini di i

Modularity

$$Q = \frac{1}{l} \sum_{i, j \in N} \left(a_{ij} - \frac{k_i k_j}{l} \right) \delta_{m_i, m_j}$$

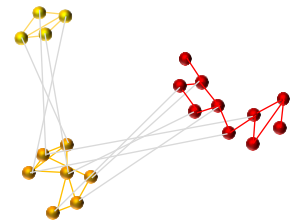
m_i : è il modulo che contiene il nodo i

$\delta_{m_i, m_j} = 1$ se $m_i = m_j$, 0 altrimenti

a_{ij} : stato della connessione

l : numero dei link

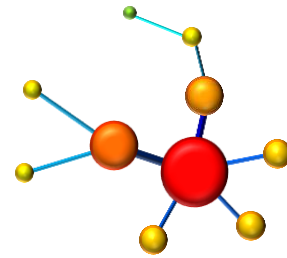
k_i : grado del nodo i



Strength

$$S = \sum_{j \in N} w_{ij}$$

w_{ij} : peso del link tra i nodi i, j



Rispettivamente:

- **Betweenness Centrality:** appartiene alle misure di centralità e stima la rilevanza di un nodo in base al numero di cammini di minima lunghezza che lo attraversano. E' definita, infatti, come la frazione di cammini di minima lunghezza che connettono due nodi arbitrari j e h che passano per il nodo i rispetto al numero totale di cammini di minima lunghezza che legano j e h . Nel caso di grafo pesato, inoltre, la definizione non subisce cambiamenti.
- **Global Efficiency:** è la media inversa della lunghezza del percorso più corto. E è un indice della capacità della rete di sopportare e gestire un flusso ingente di informazioni e segnali: tanto più grande è E , tanto minore è il cammino caratteristico e tanto più efficiente la rete, in quanto consente all'informazione di viaggiare tra un nodo e un altro nel modo più rapido possibile. Normalmente, dunque, si prediligono network strutturalmente efficienti e organizzati in quanto sono ad alta efficienza e i nodi sono molto ravvicinati tra loro. Nella condizione di grafo pesato, la definizione subisce modifiche nei confronti di d_{ij}^{-1} che diventa $(d_{ij}^w)^{-1}$, dal momento che è opportuno considerare i pesi dei link della rete.
- **Char Path:** è la media della lunghezza dei cammini più corti tra coppie di nodi all'interno della network. E' una metrica influenzata dai cammini lunghi, poiché la presenza di cammini distanti rende l'efficienza del grafo nulla. Per l'analisi di grafi pesati la definizione assume cambiamenti esattamente identici alla *Global Efficiency*, come già descritto.
- **Clustering Coefficient:** è la media della stima della probabilità che due nodi j e h adiacenti ad un nodo i di un grafo G siano adiacenti tra loro; in altri termini stima la media del numero di triangoli in cui il nodo i è uno dei vertici, su tutti gli N nodi del grafo G . I suoi valori, di solito, sono compresi tra $[0,1]$. Per i grafi pesati, inoltre,

è possibile modificare la definizione tenendo conto dei pesi dei nodi adiacenti al nodo i -esimo e della sua forza; la modifica riguarda soltanto t_i in t_i^W .

- **Local Efficiency:** è la stima della media dell'efficienza di sotto-grafi. E' strettamente correlata alla misura di *Clustering Coefficient* perché rivela quanto la rete è *fault tolerant*, in altri termini, mostra quanto efficiente sia la comunicazione tra i nodi vicini all' i -esimo quando l' i -esimo nodo è rimosso[20]. Anche E_{loc} si modifica come la E in caso di grafi pesati.
- **Modularity:** quantifica la qualità della divisione del grafo in moduli i cui nodi sono fortemente intra-connessi con numerosi archi e debolmente inter-connessi con i nodi degli altri moduli della rete. Di norma il suo valore varia tra $[-1/2, 1]$. Con segno negativo si descrive una rete con numero di archi presenti minore di quelli attesi, viceversa il segno positivo illustra un grafo con numero di archi presenti maggiori di quelli attesi. Gli archi attesi, difatti, suggeriscono l'entità di modularità nel momento in cui i link siano distribuiti in maniera random. Per grafi pesati, invece, la definizione diventa la seguente:

$$Q = \frac{1}{l^w} \sum_{i,j \in N} \left(w_{ij} - \frac{k_i^w k_j^w}{l^w} \right) \delta_{m_i, m_j}$$

- **Strength:** costituisce la somma dei pesi degli archi connessi al nodo i -esimo. Più è forte la connessione al nodo i -esimo, maggiore è l'elasticità del nodo e minore la sua vulnerabilità ad attacchi esterni.

Inoltre, per caratterizzare meglio le singole network in esame, è di rilevanza anche la *Node Degree* definita come:

$$S = \sum_{j \in N} a_{ij}$$

a_{ij} : stato di connessione del link tra i nodi i, j

Ha lo scopo di conteggiare la quantità delle connessioni in un grafo, analizzando la presenza o l'assenza di un link tra coppie di nodi.

Per il riconoscimento della collocazione cerebrale del quantitativo entropico di un grafo è necessaria annoverare l'entropia topologica espressa come:

$$E(G) = -\frac{A}{M} \sum_i^N \frac{C_i}{A} \ln\left(\frac{C_i}{A}\right)$$

A : numero di collegamenti del grafo

N : numero dei vertici del grafo

M : numero di cicli di pruning (sogliatura) necessari per disconnettere completamente (o quasi) il grafo

C_i : grado di connettività di ogni vertice

$\frac{C_i}{A}$: probabilità che un generico nodo C_j con $j \neq i$ sia direttamente connesso al nodo C_i

Da letteratura[21], il trend di tale parametro è associato ineluttabilmente all'andamento del diametro del grafo, in quanto rende possibile la comprensione della struttura del grafo a livello dimensionale e fisico. Nello specifico il diametro del grafo è definito come:

$$diam(G) = \max_{u,v \in V(G)} \min_{p \rightarrow u,v} \{|p|\}$$

Costituisce il valore massimo della distanza tra una coppia di vertici di un grafo G , dove il termine distanza tra due vertici di un grafo identifica la lunghezza del cammino più breve che li collega e $|p|$ è il modulo della lunghezza del cammino p tra i vertici u e v .

2.6 Correlazioni negative

Durante l'analisi, relativa a tale studio, delle matrici FC dei soggetti a disposizione per la caratterizzazione delle metriche di grafo su network cerebrali è emerso che tutte contengono valori di correlazione non solo positivi ma anche negativi. Mentre i valori positivi hanno un chiaro significato fisiologico a livello cerebrale e il loro studio segue le procedure standard di analisi, i valori negativi non godono della stessa caratteristica. Il primo passo dell'indagine, quindi, è stato comprendere come trattare i pesi negativi nelle reti neurali.

Come si evince da letteratura, il gruppo di ricerca di *Sporns* et al. ha consigliato la rimozione dei pesi negativi prima di affrontare l'analisi delle reti cerebrali [10]. Molti ricercatori affrontando la stessa problematica [22] hanno optato per azzerare le connessioni negative, denominate *link spuri* [23], individuando la necessità di affrontare lo studio delle stesse con tecniche completamente differenti da quelle standard e già esistenti per le positive[24]. Dopo solo due anni i medesimi ricercatori, *Sporns* et al. hanno segnalato l'utilità dell'informazione prodotta dalle connessioni negative, hanno marcato la disuguaglianza del loro apporto informativo rispetto ai pesi positivi e introdotto metriche in grado di gestirle [25]. L'attenzione sui pesi negativi, infatti, è derivata dallo studio svolto

dai ricercatori *Chang et al.* [26] i quali hanno notato la presenza di anti-correlazioni all'interno del DMN nelle matrici di connettività per R-fMRI. Sebbene la loro comparsa sia influenzata dalle tecniche di pre-processing attuate sui dati, in altre situazioni, invece, dopo l'eliminazione di fonti di rumore, i pesi negativi hanno incrementato il loro peso e hanno assunto significatività fisiologica. Motivo per cui *Sporns et al.* hanno intrapreso lo studio per ridefinire alcune metriche di grafo al fine di includere i pesi negativi nelle analisi delle network, in particolare la *Modularity* e altre misure di centralità[25]. In aggiunta, studi successivi hanno associato i pesi negativi alla non omogeneità della risposta emodinamica cerebrale [27] o all'attività neurale che connette le aree corticali e non, ciononostante senza la capacità di contestualizzarli a livello fisiologico [28]. Il loro ruolo, dunque, rimane tuttora controverso. Pertanto per quanto riguarda la fase dei test preliminari, si è scelto di affrontare l'analisi dei valori di FC individuando quattro casi distinti: FC con segno, FC solo positivi, FC con applicazione di una soglia crescente per valori sia negativi sia positivi, FC in modulo con soglia crescente, descritti in dettaglio successivamente.

Per ciò che concerne l'identificazione di un criterio per sogliare in maniera adeguata i dati, tuttora la letteratura non è in grado di fornire una strategia robusta. Dai ricercatori *Simpson et al.* sono consigliate vari tipi di soglie quali: soglia fissa, altrimenti nota come assoluta, soglia di grado medio e soglia legata alla densità del link tra i nodi [23]. Mentre le prime due costituiscono soglie uniformi applicate a tutte le reti, la terza ha lo scopo di minimizzare il numero di connessioni con la garanzia che si mantengano le connessioni alle componenti principali. Ad ogni modo rappresentano soglie applicate in maniera omogenea all'intera network, ossia tutti i link cerebrali sono recisi allo stesso modo. Un altro gruppo di ricerca, in particolare *Wang et al.* ha aggiunto oltre alla soglia assoluta, una soglia relativa alla sparsità dei valori di correlazione della matrice, ossia una percentuale di connessioni, scelta dall'utente, è eliminata con la sicurezza del mantenimento della compresenza di pesi sia positivi sia negativi [1].

Capitolo 3

Materiali e Metodi

Nella seguente sezione è descritto il data-set esaminato durante lo studio e con quali tecniche di pre-processing è stato ottenuto. Sono approfondite in dettaglio le procedure applicate ai dati. Esse comprendono i test preliminari svolti sull'informazione contenuta nelle matrici di correlazione funzionale, la caratterizzazione delle metriche di grafo e l'applicazione delle procedure scelte su un numero maggiore di soggetti rispetto agli stadi precedenti.

3.1 Data-set e Preprocessing

Per il lavoro di tesi il data-set studiato è costituito da matrici di connettività funzionale di soggetti di controllo, definiti sani rispetto all'area anatomica di interesse: la regione cerebrale. In totale i soggetti esaminati sono stati 25, ripartiti in 12 donne e 13 uomini con età media $39,44 \pm 11,27$ anni. I dati sono stati acquisiti nel periodo compreso tra il 02/02/2015 e 16/01/2016. Nella **Tabella 3.1** sono riportati i 25 soggetti con la loro età, l'identificativo interno e la data di acquisizione.

Durante le fasi iniziali, riguardanti i test preliminari e l'analisi delle metriche di grafo, solo i dati derivanti da 10 soggetti sono stati sottoposti alle metodologie sperimentali, per lo stadio finale della sperimentazione, invece, sono stati aggiunti i restanti 15. Ogni soggetto è stato sottoposto all'esame *fMRI* in condizioni di *resting-state*. Sono state, quindi, ottenute le immagini relative a ciascun individuo da cui si sono estrapolate le metriche temporali di connettività, basate sull'analisi di correlazione. Le matrici di connettività ricavate, dunque, contengono per ogni soggetto i valori di correlazione *pairwise* tra le singole regioni di interesse, con valori compresi nel range $[-0.5545: 1]$. Le aree cerebrali sono state suddivise in 99 regioni volumetriche di interesse (ROI) (Tabb. 2.2 e 2.3), secondo la segmentazione di

immagini T1w effettuata con il *multi-atlas segmentation algorithm with joint label fusion*(MALF)[29].

<i>ID Soggetto</i>	<i>Data di Acquisizione</i>	<i>Sesso</i>	<i>Età</i>
01	20150202	F	42
02	20150309	M	34
03	20150316	M	35
04	20150316	F	65
05	20150316	M	24
06	20150323	M	50
07	20150411	M	24
08	20150411	M	31
09	20150411	M	27
10	20150524	M	35
11	20150524	F	53
12	20150525	F	31
13	20150525	F	34
14	20150525	F	42
15	20150926	F	47
16	20150926	F	60
17	20150926	F	47
18	20150926	F	51
19	20150926	F	48
20	20160116	M	45
21	20160116	M	39
22	20160116	M	36
23	20160116	F	26
24	20160116	M	36
25	20160116	M	24

Tabella 3.1 *Soggetti di Controllo*

Le immagini sono state sottoposte ad operazioni di *pre-processing* avanzato quali:

- *slice timing correction*, per allineare temporalmente le immagini[30],
- *readout distortion correction*, necessario per eliminare le distorsioni dovute agli eco per la costruzione delle immagini[31],

- *motion correction* con una trasformazione a 12 gradi di libertà, utile per ridurre eventuali elementi di rumorosità dovuti ad involontari spostamenti del capo del paziente[32],
- *spatial normalization* con *template* MNI 152, al fine di rendere confrontabili i soggetti rilevati[33].

Inoltre, alle serie temporali *fMRI* sono stati applicati un filtro passa-alto con frequenza di taglio fissata a 1/128 Hz e un filtro spaziale con ampiezza del kernel pari a 5mm, in corrispondenza della metà del suo massimo (*FWHM*). Infine, la parcellizzazione in ROI, anch'essa normalizzata, è stata eseguita per l'estrazione delle dinamiche relative alle singole regioni. Inoltre, altre sorgenti di rumore sono state rimosse dalle dinamiche temporali delle ROI quali: il movimento, segnali relativi a materia bianca e liquido cerebrospinale attraverso l'utilizzo di una regressione lineare. Le matrici di connettività funzionale sono state ottenute con la correlazione di *Pearson*, definita dalla seguente espressione:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

con ρ_{XY} , la correlazione, σ_{XY} , la covarianza di X e Y, σ_X e σ_Y , le deviazioni standard di X e Y.

3.2 *Test preliminari sulla correlazione funzionale*

La prima fase di studio ha curato l'analisi statistica dei dati con l'obiettivo di discernere l'informazione utile da elementi di rumore contenuta nelle matrici FC, individuare un criterio per scegliere opportunamente le matrici al fine di eliminare le connessioni meno robuste, garantire metriche di grafo resistenti agli outlier e al contempo sensibili. In particolare, si è quantificata la variabilità della connettività funzionale su tre livelli di indagine: intra-soggetto, inter-soggetto e di popolazione. Precisamente, intra-soggetto indica la valutazione della distribuzione dei valori di FC considerando le correlazioni cerebrali presenti all'interno dell'encefalo del singolo soggetto (ricerca orizzontale); inter-soggetto si riferisce all'analisi multivariata tra più campioni esaminati in simultanea (ricerca verticale); l'analisi di popolazione riguarda lo studio dell'informazione relativa all'insieme dei valori di connettività di tutti i soggetti considerati come un unico campione, quindi, non è più possibile identificarli singolarmente. I soggetti sono accomunati dallo stesso range di misura

riguardante il peso della connettività e la stessa parcellizzazione. Per ciascun livello di indagine le distribuzioni statistiche di ciascun soggetto relative ai valori di correlazione delle matrici di connettività sono state esaminate singolarmente attraverso ispezione visiva dei relativi istogrammi.

Le tecniche statistiche, ritenute adeguate allo scopo dello studio, sono state: l'analisi della componente principale (PCA), test di Student (t-test), test della somma dei ranghi, l'indice di entropia e gli indici di media, mediana e varianza per l'analisi di popolazione. In relazione alla trattazione, lo scopo della PCA è stato di quantificare la variabilità intra-soggetto della connettività funzionale, tramite decorrelazione dei dati relativi ai singoli soggetti. Inoltre, ha permesso di comprendere il potere discriminante della scala numerica applicata alle matrici di connettività. La quantità di informazione è stata, infatti, descritta dalla varianza spiegata. In particolare, la PCA è stata implementata in due ipotesi di lavoro: matrice di connettività funzionale in condizioni originali e la stessa in condizioni di normalizzazione. Dopo l'implementazione della PCA, solo la prima componente principale è confrontata con il campione medio FC dei soggetti, ottenuto come media dei valori FC inter-soggetto. Il confronto è eseguito sia con il t-test sia con la somma dei ranghi. Per entrambi i test l'intento è di verificare se la prima componente principale sia in grado di rappresentare il campione medio FC dei soggetti, ossia se i due campioni siano non statisticamente significativi e, dunque, i valori medi delle due distribuzioni uguali.

Infine, l'indice di entropia indica il contenuto informativo delle matrici di connettività in esame. In questo elaborato è stata utilizzata la definizione di entropia secondo Shannon. Tale indice è stato selezionato in quanto costituisce un utile indicatore dello stato di attivazione dell'area cerebrale e un potenziale marcatore per malattie cerebrali proprio nell'analisi in resting-state. Dal momento che è un utile indicatore del contenuto informativo delle matrici di connettività al variare della soglia, il suo andamento è stato usato come parametro di confronto per stabilire la robustezza delle metriche di grafo al cambiamento della soglia applicata, in linea allo stato fisiologico della rete cerebrale. In questo modo, quindi, si è cercato di caratterizzare le singole misure ma anche di identificare un intervallo di soglia adeguato da applicare alle mappe di connettività. Di norma, i valori possibili sono soltanto positivi e il valore minimo assumibile è zero, corrispondente al minimo disordine e massima uniformità dell'informazione. Si è scelto di monitorare tale indice sui tre livelli di indagine e al variare della soglia. La tipologia di soglia ritenuta adeguata, in base a quelle proposte in letteratura, è stata quella assoluta in quanto effettua un taglio netto ed omogeneo sulle connessioni delle network dei soggetti. In parallelo, al variare della soglia, è monitorata

anche la quantità di connessioni rimanenti sui grafi dei soggetti, per marcare come la variabilità intra-soggetto abbia un ruolo non trascurabile durante la sperimentazione non solo sull'informazione ricavabile dalle matrici di connettività ma anche sull'analisi delle metriche di grafo in termini di stabilità e affidabilità.

L'evidente compresenza di pesi negativi[24] e positivi in tutti grafi e la difficoltà di caratterizzare il significato fisiologico della loro unione, riscontrata dalla letteratura[28], hanno determinato la scelta di condurre le prove iniziali su quattro casi di studio per valutare l'effetto topologico delle network:

- *FC sia positivi sia negativi*: relativo ai dati originali senza alcuna connessione rimossa e senza l'applicazione di una soglia,
- *FC positivi*: indica il mantenimento delle sole correlazioni positive e l'azzeramento delle negative[10],
- *FC con soglia assoluta crescente [-0.5745: 0.01: 0.92]*: rappresenta il caso in cui i dati originali sono gradualmente eliminati in base al valore della soglia imposta e in maniera omogenea; il valore minimo e il valor massimo sono dovuti a scelte progettuali al fine di semplificare l'analisi dei dati, approfondita in *Specifiche tecniche*, il passo di un centesimo permette di avere un'ampia e puntuale visione dell'effetto del cambiamento della network
- *FC in modulo con soglia crescente [0: 0.01: 0.92]*: riferito ai dati originali presi in modulo con la cancellazione delle connessioni con peso al di sotto della soglia

La scelta progettuale di effettuare tale distinzione ha permesso di comprendere a fondo sia il contenuto informativo prodotto dalla presenza contemporanea dei pesi di entrambi i segni sia quello derivato dalla presenza dei pesi considerati singolarmente in segno, seguendo il principio di sovrapposizione degli effetti. Specificamente, il range di valori, adattato ai casi in esame, è oggetto di approfondimento e la sua descrizione segue nella sezione 3.3.4. Inoltre, i dati sono impiegati in due tipologie di distribuzioni: originale, senza l'attuazione di alcuna modifica e standardizzata, con media nulla e deviazione standard unitaria. La distinzione tra le due tipologie di distribuzioni si mantiene solo per la PCA. Per i test statistici, invece, al fine di stabilizzare la valutazione delle distribuzioni dei campioni comparati, i valori sono successivamente standardizzati. I test, infatti, sono implementati in due ipotesi di lavoro. La prima prevede di standardizzare i dati solo prima di eseguire i test, la seconda standardizza i dati prima ancora di ottenere i campioni da confrontare e di eseguire i test statistici.

3.3 *Metriche di grafo*

La seconda fase dello studio concerne la vera e propria caratterizzazione delle metriche di grafo, relativamente alle loro proprietà di sensibilità e robustezza, e l'individuazione di un valore di soglia identificato come limite superiore, ricavato tramite l'incrocio delle informazioni derivanti dagli andamenti delle misure stesse.

3.3.1 *Correlazioni negative*

In prima istanza, al fine di definire in maniera permanente se mantenere o meno il quantitativo di correlazioni negative, sono state analizzate tre metriche di grafo quali: *betweenness centrality*, *clustering coefficient* e *local efficiency*, in quanto da letteratura sono meglio caratterizzati i loro andamenti e sono le più usate per comprendere le reti cerebrali in un'analisi preliminare. Per ciascuna metrica si sono ispezionate le loro distribuzioni statistiche per valori FC integrali, solo positivi, solo negativi e con applicazione di una soglia fissa in due tipologie di intervalli: asimmetrico con range di valori $[-0.02, 1]$, simmetrico con intervallo $[-0.12, 0.12]$ e infine si è anche applicato un taglio netto a 0.3.

3.3.2 *Caratterizzazione metriche di grafo con indice di Entropia*

Dai test preliminari si è ricavato l'andamento dell'indice entropico del contenuto informativo delle matrici FC di partenza. Si è voluto, quindi, sfruttare questa conoscenza e metterla a confronto con l'entropia calcolata sulle singole metriche di grafo. In questo modo si sono potute caratterizzare le misure in termini di sensibilità e robustezza al variare della soglia applicata alle mappe, ma anche delineare un valore di soglia adeguato ai dati FC. In questa fase dell'indagine, l'insieme delle metriche di grafo si è arricchito di ulteriori misure, quali: *strength* e *global efficiency*, in quanto possiedono la particolarità di distinguere i singoli soggetti grazie ad una lettura più semplice delle informazioni derivanti dai grafi e perché entrano più in profondità nell'analisi degli stessi. Per ciascuna metrica è stato implementato l'indice di entropia a vari livelli di analisi: intra-soggetto, inter-soggetto, popolazione e matrice, come già descritto in precedenza e le matrici FC sono state considerate solo positive, nel caso di grafi non orientati e pesati. Per le modalità implementative si fa riferimento al

paragrafo *Specifiche tecniche*. Inoltre, il trend entropico relativo a livello inter-soggetto ha mostrato un ginocchio significativo al variare della soglia, come evidenziato in **Figura 3.3**.

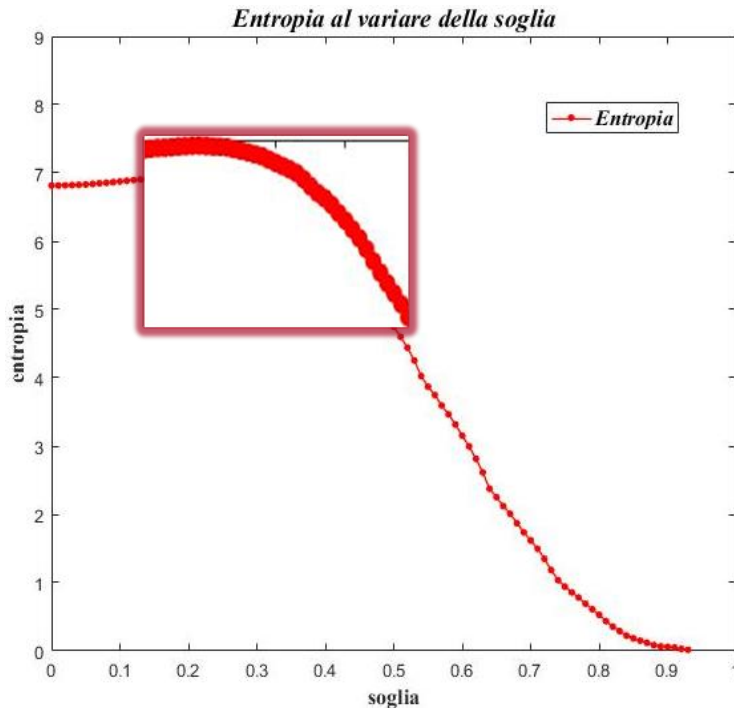


Figura 3.3 Entropia Inter-Soggetto al variare della soglia: punto di ginocchio evidenziato nel riquadro significativo per la caratterizzazione delle metriche di grafo

L'individuazione del punto critico del ginocchio si è ritenuto utile per un'eventuale identificazione della soglia. Proprio la presenza del ginocchio, infatti, ha marcato la necessità di scegliere opportunamente un intervallo di soglia per i dati di connettività funzionale, allo scopo di mantenere soltanto i link più informativi e con alto livello di attività cerebrale. Tale esigenza è evidenziata dalla diminuzione del valore dell'indice di entropia, una volta superato il punto di ginocchio. Si è cercato, quindi, di ricavare un modello matematico tramite linearizzazione di tratti della curva, in particolare linearizzando il plateau iniziale e la retta con pendenza negativa successiva al punto di ginocchio, per poter ottenere in maniera deterministica tale valore, tramite l'individuazione del punto di intersezione tra le due rette. Tuttavia, sia la modellizzazione sia il confronto tra le metriche di grafo e le matrici FC, basate sull'indice di entropia, non hanno fornito risultati soddisfacenti e ripetibili. Pertanto si è approcciato il problema attraverso i *survival plot*.

3.3.3 Caratterizzazione metriche di grafo tramite survival plot

Da letteratura i *survival plot* [9] sono definiti come grafici basati sul confronto tra il trend entropico, ricavato dai dati FC preliminari, e gli andamenti delle metriche di grafo. Nella stessa figura vengono rappresentate l'entropia della matrice di connettività e la metrica di grafo in funzione della variazione della soglia, come viene rappresentato in **Figura 3.4**

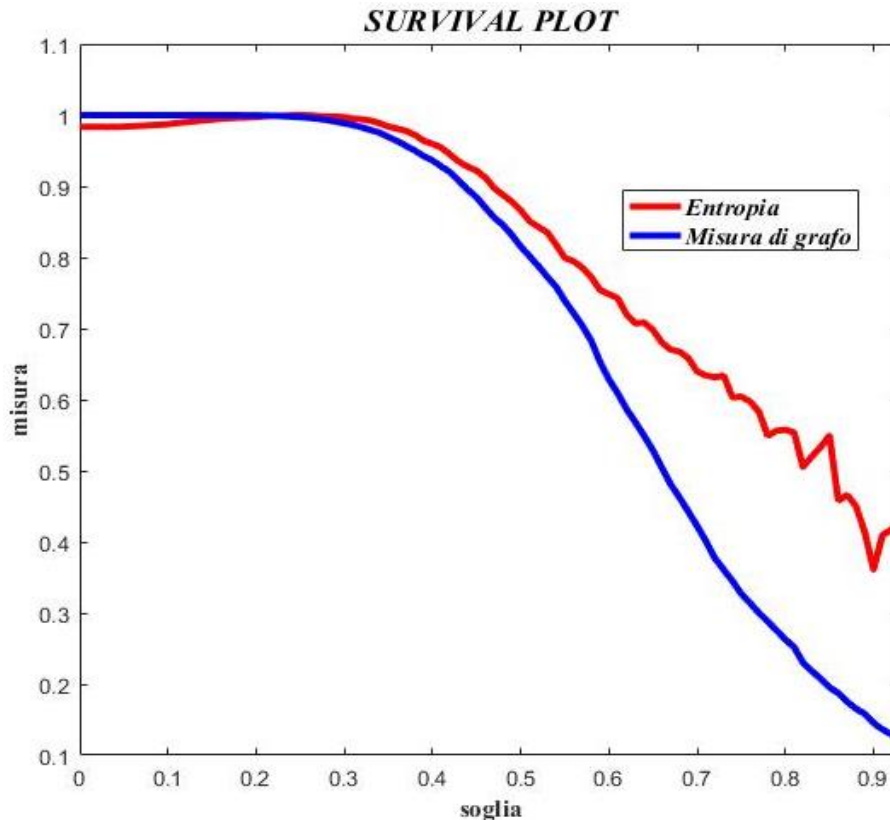


Figura 3.4 Survival Plot: la curva di entropia e la curva della metrica di grafo sono confrontate nello stesso grafico

Le curve sono state implementate a livello intra-soggetto e inter-soggetto, al fine di caratterizzare le metriche di grafo in termini di variabilità, sensibilità e stabilità al variare della soglia. Per identificare eventuali outlier, seguire il contenuto statistico dell'informazione delle metriche di grafo e stabilire definitivamente il ruolo della variabilità intra-soggetto sono state poste a confronto le singole misure, mediante opportuni boxplot e tracciando gli archi di peso non significativo al variare della soglia. Quest'ultimo aspetto ha richiesto il calcolo della metrica *Node Degree*, che rappresenta una delle misure basilari per la comprensione topologica dei grafi. A questo stadio dell'analisi l'insieme delle metriche si è esteso ad altre misure più sofisticate e in grado di caratterizzare le reti cerebrali a livello

globale, quali: *modularity* e *char path*. Mentre a livello implementativo, le metriche già evidenziate in precedenza sono state calcolate in condizioni di default rispetto alle variabili in gioco, per quanto concerne la metrica di *modularity* e la *clustering coefficient* sono state sperimentate anche opzioni non predefinite. Per la *modularity*, infatti, sono stati considerati tre diverse grandezze di moduli, noti come: *small*, *classic* e *large*, identificati dai rispettivi valori della variabile *gamma* [0.5, 1, 2], con l'obiettivo di indagare varie dimensioni delle ROI cerebrali. Inoltre, sono state utilizzate due tecniche implementative di ottimizzazione, la *Danon*[22] e la *Newman*[34], al fine di discriminare eventuali influenze dovute al metodo di calcolo. La *Danon*, sebbene sia un algoritmo più lento e la stima di solito non sia ottima, interpreta in maniera esatta la metrica non sacrificando l'accuratezza dell'informazione per la velocità di calcolo. La *Newman*, invece, è un algoritmo più accurato, veloce per reti di piccole dimensioni, fornisce una stima ottima della metrica ma comporta la perdita del grado di accuratezza della misura per favorire la velocità computazionale. Tuttavia, è necessario puntualizzare che l'ottimizzazione con la rete reale oggetto di studio permette di garantire che la modularità trovata non sia legata al metodo di misura, ma è effettivamente una proprietà intrinseca della rete[8]. Per la *Clustering Coefficient*, invece, la scelta ha riguardato la definizione della misura ed è ricaduta sulla formula di *Onnela* come riportato dal gruppo di ricerca *Rubinov et al.*[10]. Allo scopo di indagare l'impatto dell'applicazione della scala sui dati FC, la capacità delle metriche di ricavare informazioni utili in grafi pesati e non e data la scelta di considerare solo valori positivi di correlazioni, le ipotesi di lavoro applicate a tutte le metriche sono stati identificate in 6, di seguito descritte:

- **PON: Positivo Originale Non pesato**
- **POP: Positivo Originale Pesato**
- **PAF: Positivo Adiacenza Fisher**
- **PAN: Positivo Adiacenza Normale**
- **PPF: Positivo Pesato Fisher**
- **PPN: Positivo Pesato Normale**

Gli acronimi identificano le modalità di trattamento dei valori FC: positivo, indica l'esclusione di eventuali pesi negativi, non pesato/adiacenza e pesato, individuano la tipologia del grafo non pesato o pesato, originale, normale e Fisher, sottolineano l'applicazione sui dati di un'eventuale scala numerica, in particolare sono state applicate la trasformata di Fisher e la normale, o sono stati mantenuti inalterati. Gli andamenti delle singole metriche ricavati dai *survival plot* sono stati raggruppati per un'analisi comparativa

inter-metrica e tra metrica ed entropia FC, possibile attraverso la normalizzazione dei valori di ciascuna grandezza rispetto al proprio massimo.

Come immediata conseguenza si è svolta l'analisi statistica delle metriche che ha riguardato il calcolo di:

- Variabilità a soglia: la media della metrica di grafo,
- La sensibilità a soglia: $\frac{sd}{media}$ rapporto tra deviazione standard e la media,
- Il range di misura: $[min,max]$, l'intervallo in cui sono compresi i valori di ogni metrica[16].

In questo modo si sono potute distinguere le misure più robuste o meno all'applicazione di una soglia fissa sui dati FC. Infine, per poter comprendere il contenuto informativo del grafo di ogni soggetto a livello strutturale è stata implementata l'entropia topologica al variare della soglia, indice che ha permesso di semplificare le valutazioni sui grafi. Come descritto nel paragrafo 2.5, il trend entropico assume significato in concomitanza alla visione della curva del diametro di grafo, in quanto l'incrocio di tali dati rende comprensibile la posizione delle ROI più informative e la struttura del grafo a livello topologico. Tuttavia in questa ultima fase di studio i casi di studio affrontati sono stati: POP, PON, PPF e PAF, si è esclusa la normalizzazione dei dati FC. Se inizialmente queste analisi hanno riguardato soltanto il gruppo ristretto di 10 soggetti, successivamente sono stati inclusi i 15 soggetti restanti per verificare la robustezza e la ripetibilità dei risultati, i quali verranno sviluppati nel capitolo seguente.

3.3.4 *Specifiche tecniche*

Il software *MatLab R2016a* è stato usato per l'elaborazione dei dati. L'esecuzione delle tecniche standard di analisi per le matrici di connettività funzionale ha previsto l'utilizzo delle seguenti funzioni implementative:

- *Pca*: per la tecnica di PCA con opzioni tutte in default tranne la *Centered*, definita come *false*,
- *Entropy*: per l'indice di entropia, con opzioni di default,
- *Ttest*: per il test di Student, in condizioni di campioni *paired*, tipologia di test bilaterale e ipotesi di uguaglianza della deviazione standard tra i campioni,

- *Ranksum*: per la somma dei ranghi, con opzioni di default, tipologia di test bilaterale.

Il grado di significatività, per entrambi i test, è scelto pari a 5%, significativo e a 1%, molto significativo, senza l'introduzione di correzioni per test multivariati.

E' noto da letteratura che la correlazione *pairwise* è una metrica misurata per coppie di ROI il cui valore è sempre compreso tra -1 e 1. Durante il lavoro di tesi il range di valori delle soglie assolute sono modellati in base ai dati in possesso, il valore minimo contraddistingue esattamente il limite inferiore dei pesi delle connessioni, il massimo valore, invece, è arrestato a 0.92 per motivi implementativi. Il numero di connessioni rimanenti nei grafi diventa talmente esiguo da non rendere possibile l'individuazione della prima componente principale e, soprattutto, la progressione dell'analisi risulta per nulla significativa. Il passo 0.01, invece, ha permesso di scandagliare minuziosamente la variabilità delle misure al crescere della soglia e quantificare la variabilità topologica dei grafi. Oltretutto, per rendere più stabile la valutazione degli elementi, le distribuzioni sono standardizzate attraverso l'operazione di *z-score*. Essa, esattamente, è un procedimento che riconduce una distribuzione di media μ e varianza σ^2 [16] ad una distribuzione standard di media nulla e varianza unitaria, eliminando la media e dividendo per la deviazione standard ($z = \frac{X-\mu}{\sigma}$). Per l'applicazione della trasformata *z-Fisher*, invece, si ricorda che essa è definita secondo la seguente formula $z = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$ con r il coefficiente di correlazione.

Inoltre, per caratterizzare le distribuzioni dei valori di correlazione sia delle matrici di connettività sia per le metriche sono stati calcolati indici quali:

- Skewness: $\frac{m^{(4)}}{\sigma^3} = \frac{\text{momento centrale di ordine 3}}{\text{deviazione standard}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^3}{\sigma^3}$: indice di asimmetria
- Curtosi: $\frac{m^{(4)}}{[m^2]^2} = \frac{\text{momento centrale di ordine 4}}{\text{varianza}^2} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^4}{\sigma^4}$: indice di gaussianità
- Coefficiente di variazione: $\frac{\sigma}{|\mu|} = \frac{\text{deviazione standard}}{\text{modulo della media}}$: indice di variabilità
- Media: $\frac{\sum_{i=1}^n x_i}{n} = \frac{\text{sommatoria delle misure } x_i}{\text{numero dei dati}}$: indice di posizione
- Deviazione standard: $\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$: indice di dispersione

Software	R-fMRI Pre- processing	Network construction (static)	Network construction (dynamic)	Graph analysis	Statistics	File	GUI	Parallel computing	Vis	Website
GRETNA	✓	✓	✓	✓	✓	✓	✓	✓	x	http://www.nitrc.org/projects/gretna/
BCT	x	x	x	✓	x	x	x	x	x	https://sites.google.com/site/bctnet/
GAT	x	✓	x	✓	✓	x	✓	x	✓	Not available
PANDA	x	✓	x	x	x	x	✓	✓	x	http://www.nitrc.org/projects/panda/
CONN	✓	✓	x	✓	✓	x	✓	x	✓	http://www.nitrc.org/projects/conn
eConnectome	x	✓	x	x	x	x	✓	x	✓	http://econnectome.umn.edu/
BrainNet Viewer	x	x	x	x	x	x	✓	x	✓	http://www.nitrc.org/projects/brw/
GraphVar	x	✓	✓	✓	✓	✓	✓	x	✓	http://www.nitrc.org/projects/graphvar/
Brainwaver	x	✓	x	✓	x	x	x	x	✓	http://cran.r-project.org/web/packages/brainwaver/

File, flexibility; GUI, graphical user interface; Vis, visualization. Note: flexibility is determined according to whether a toolbox provides options regarding at least three of the following factors: network node, network connectivity, network connectivity member, network type and thresholding procedure.

Tabella 3.2 Sommario dei software per l'imaging connettomico: confronto tra GRETNA e BCT [1]

In conclusione, è necessario puntualizzare anche l'accurata rimozione della componente di bias al variare della soglia, ossia la presenza di una grande quantità di connessioni nulle che tendono ad alterare in maniera consistente gli esiti delle analisi. Tale operazione è effettuata in tutte le fasi dello studio, sia operando manualmente sia automaticamente durante l'utilizzo delle specifiche funzioni.

Per computare le metriche di grafo si è ricorso a due software: *Brain Connectivity Toolbox*(BCT) [10], il primo toolbox creato per lo sviluppo delle metriche di grafo per applicazioni su rete cerebrali e *GRaph thEoretical Network Analysis* (GRETNA), toolbox per l'imaging connettomico di recente diffusione[1]. Tale adozione ha permesso di analizzare a fondo le metriche in base alle ipotesi di lavoro e con una grande varietà di opzioni. La scelta di utilizzare due software contemporaneamente è stata dettata dal fatto che in GRETNA il toolbox delle metriche prevede una rosa completa di funzioni a disposizione per l'analisi dei grafi, come si può evincere dalla tabella riportata.

Il software GRETNA non solo permette l'analisi dei grafi ma anche la costruzione sia statica sia dinamica delle network, l'analisi statistica, di applicare le tecniche di pre-processing, a differenza della BCT in cui è possibile solo implementare le metriche di grafo. Relativamente alle metriche di grafo, GRETNA detiene una più ampia scelta di misure e permette di valutare un numero maggiore di metriche in presenza di pesi negativi, quali *Clustering Coefficient*, *Strength*, *Modularity* ed altre ancora.

Capitolo 4

Risultati

Nel capitolo seguente sono descritti i risultati ottenuti che giustificano quantitativamente l'impatto delle scelte progettuali sulle network dei soggetti. In prima istanza, al fine di comprendere il ruolo delle correlazioni negative all'interno delle mappe FC sono state svolte le analisi sulle matrici del data-set a disposizione attraverso il principio di sovrapposizione degli effetti, includendo tutte le ipotesi di lavoro relative al trattamento dei valori di connettività funzionale. Successivamente, si è testata una prima metodologia per la caratterizzazione delle metriche di grafo basata sull'indice di entropia. Infine, le metriche di grafo sono state caratterizzate in dettaglio grazie ai grafici contenenti tutte le metriche normalizzate, che forniscono un quadro completo degli andamenti delle misure, e all'ispezione visiva delle network individuali dei singoli soggetti presi in esame.

4.1. Test preliminari

4.1.1. Analisi delle distribuzioni individuali

Nella prima fase di studio, i test preliminari sono stati svolti su un gruppo ristretto di 10 soggetti che hanno mostrato dinamiche temporali più stabili. La componente relativa ad artefatti da movimento, infatti, è poco significativa. Per ciascun individuo sono state esaminate le distribuzioni statistiche dei valori FC con segno singolarmente, sia mantenendo i valori di correlazione inalterati sia applicando la procedura di normalizzazione.

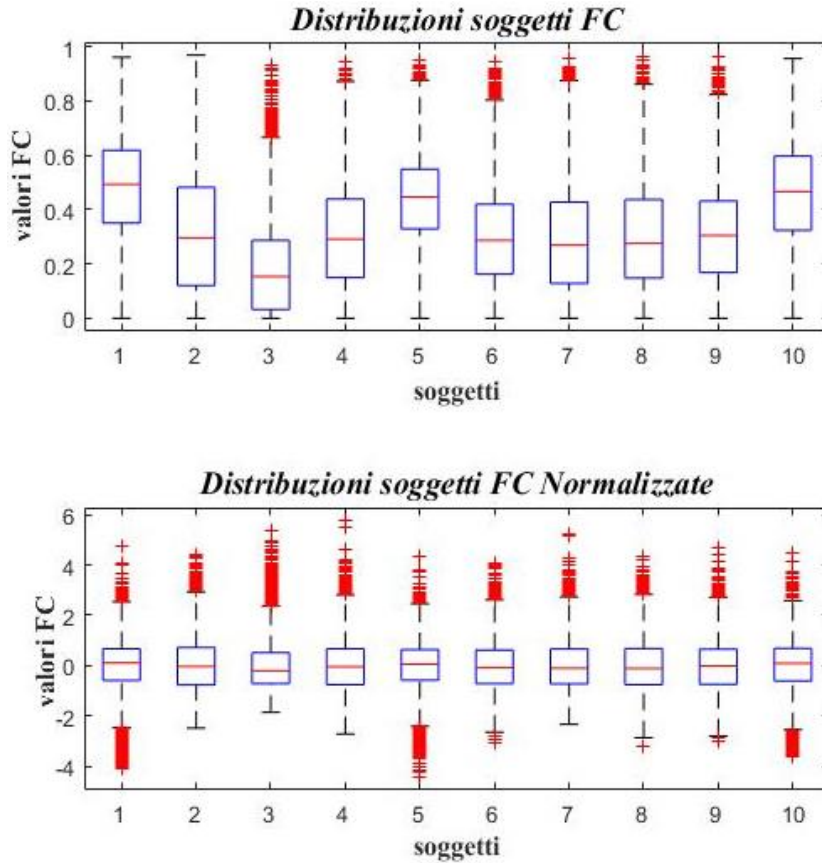


Figura 4.5 Distribuzioni individuali relativi ai 10 soggetti dei valori di connettività funzionale sia positivi sia negativi e normalizzati

Le distribuzioni dei valori di correlazione sia positivi sia negativi hanno mostrato un range di misura simile intra-soggetto, tale da individuare più individui con lo stesso trend. Nel pannello superiore della **Figura 4.5**, infatti, si possono notare dei raggruppamenti. Mentre i soggetti 1-10 costituiscono un primo gruppo, i soggetti 3 e 5 si discostano non solo dal primo gruppo ma anche da tutti gli altri che evidenziano lo stesso intervallo di valori. Il 3, infatti, è l'unico che detiene valori molto bassi di correlazione compresi tra 0.1 e 0.3. Il soggetto 5 ha una mediana inferiore rispetto ai soggetti 1 e 10 e anche un range di misura più concentrato rispetto al valor medio con deviazione standard pari a 0.173, valore minimo tra tutti i soggetti. Queste osservazioni hanno indicato la presenza di una certa variabilità intra-soggetto non trascurabile, confermata ulteriormente nelle successive indagini. A supporto di tale evidenza sono state riportate le matrici di connettività media e la deviazione standard (**Figura 4.6**). I valori bassi di standard deviation individuano il valor medio come dato consistente per i successivi test.

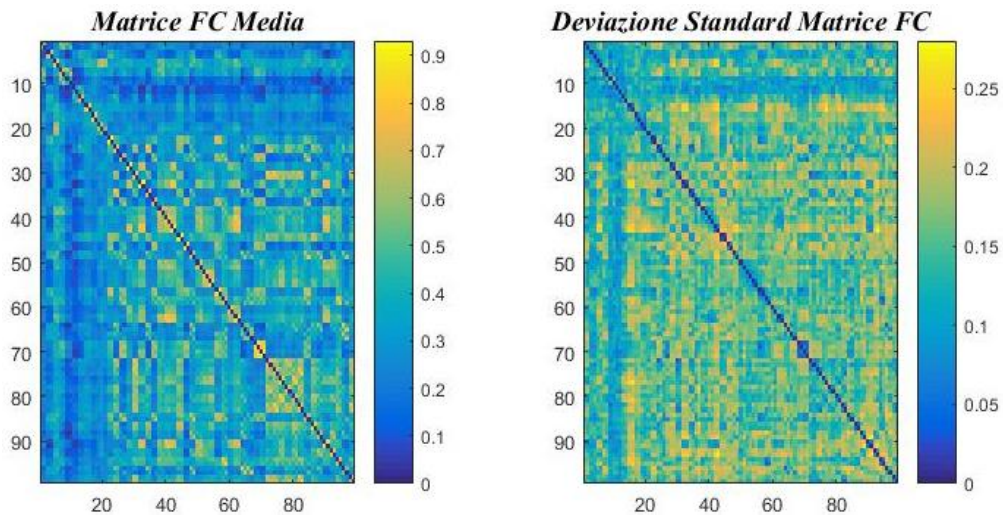


Figura 4.6 Immagini della connettività funzionale media intra-soggetto e della deviazione standard relativi ai 10 soggetti

Inoltre, l'analisi di correlazione di Pearson per le matrici di connettività ha determinato distribuzioni dei campioni non normali e con valori di skewness negativi, come per i soggetti 1, 5 e 10, pari a -0.35 per 1 e 10 e -0.33 per il soggetto 5. Nel caso di distribuzioni normalizzate, invece, il range di valori è totalmente equalizzato e le distribuzioni dei valori di correlazione sono simmetriche per tutti i soggetti, in quanto la trasformazione z-score normalizza i valori, di conseguenza, la variabilità intra-soggetto si annulla, come si osserva nel pannello inferiore della **Figura 4.5**.

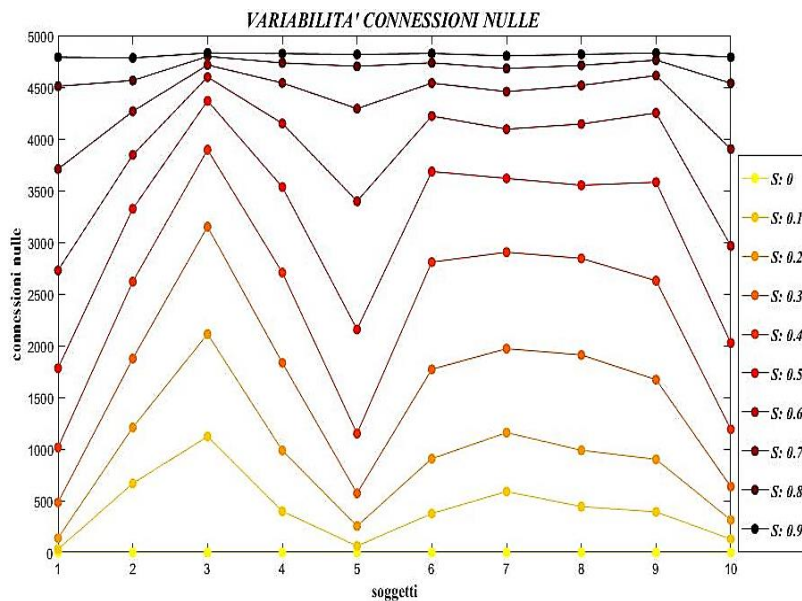


Figura 4.7 Variabilità individuale del numero di connessioni nulle al variare della soglia con passo di campionamento di un decimo

Inoltre, le distribuzioni sono state analizzate anche con l'applicazione di una soglia assoluta $[-0.5745: 0.01: 0.92]$. Il range di valori, descritto nel capitolo 3.2, è caratterizzato da un valor minimo inferiore al minimo valore di correlazione presente nelle matrici di connettività e da un valore massimo, troncato a 0.92. Sia l'uno sia l'altro valore sono stati scelti per semplificare i calcoli durante le analisi. Per cambiamenti della soglia con passo di campionamento di un decimo, molti valori di correlazione si annullano, a causa della scomparsa di numerose connessioni cerebrali. La **Figura 4.7**, infatti, mostra la variabilità dei singoli soggetti in relazione alla quantità di connessioni rimosse all'aumentare della soglia. Inoltre, il valore medio aumenta, in quanto persistono solo i link con peso maggiore, le code degli outlier si assottigliano, dal momento che resistono le connessioni più significative e la variabilità intra-soggetto decresce, dando vita a distribuzioni simmetriche e similari per tutti i soggetti. La similarità, infatti, è stata calcolata come variabilità dell'indice di skewness tra i soggetti. La variabilità dell'indice di skewness è rappresentata dal valor medio dei valori di skewness delle distribuzioni del singolo soggetto al variare della soglia. Qui si riporta l'andamento dell'indice di simmetria medio **Figura 4.8**.

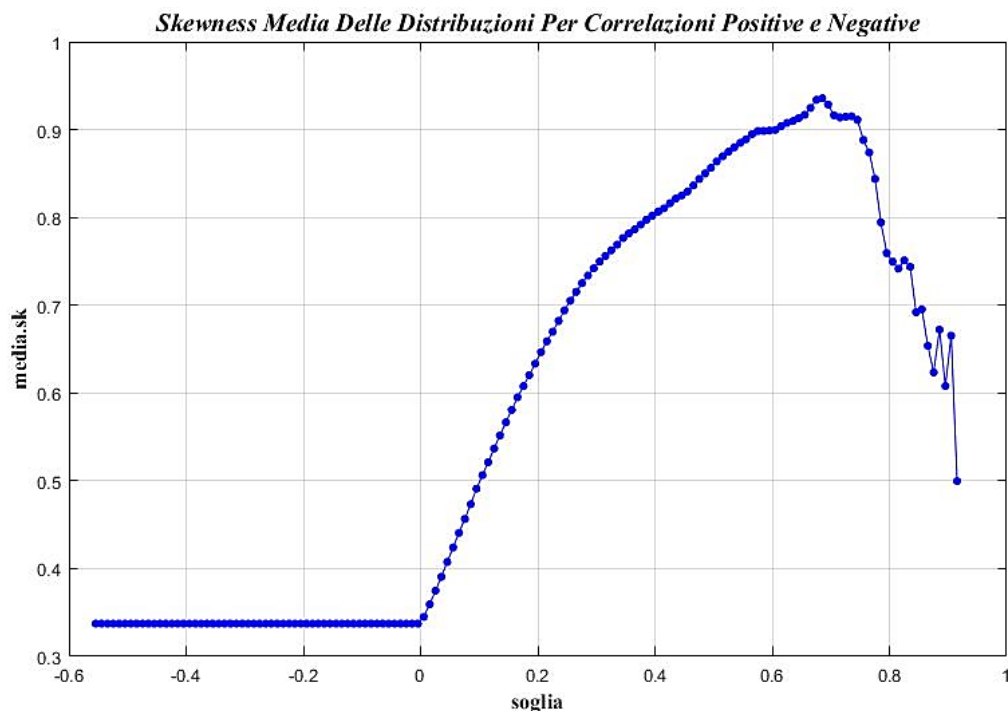


Figura 4.8 Profilo di skewness medio al variare della soglia delle distribuzioni per valori di correlazione sia positivi sia negativi

Si può notare che la simmetria delle distribuzioni dei valori di correlazione sia positivi sia negativi decresca per valori di soglia superiori a 0.7. Tale valore di soglia costituisce il punto di massimo della curva in corrispondenza del quale più del 60% delle connessioni cerebrali sono rimosse.

4.1.2. Connettività Funzionale: impatto dei valori negativi di correlazione

In primo luogo è stata esaminata la connettività funzionale con valori di correlazione sia positivi sia negativi senza l'applicazione di alcuna soglia e lasciando inalterati i valori di correlazione presenti nelle matrici di connettività funzionale. Dall'analisi svolta tramite PCA si è osservato che la varianza spiegata è maggiore nel caso in cui la distribuzione non sia stata standardizzata, come in **Figura 4.9**.

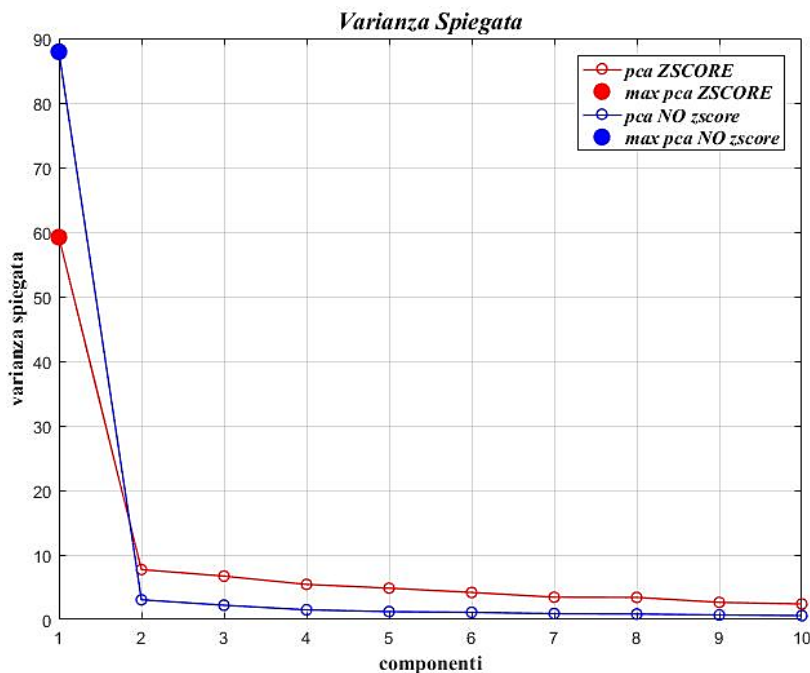


Figura 4.9 Varianza Spiegata per Matrici di Connettività senza l'applicazione di una soglia, nel caso di dati normalizzati e non

I pallini pieni costituiscono il massimo valore della varianza spiegata nelle due tipologie di scala numerica per le distribuzioni dei soggetti, normalizzate e non, in corrispondenza della prima componente principale. Per il caso di distribuzione originale la varianza si attesta all'88%, nel secondo, invece, al 59%. La differenza in percentuale tra le due modalità è di circa 30%, considerevole dal punto di vista della quantità di informazione associata ad essa, e rivela una forte variabilità intra-soggetto. Quest'ultima, infatti, è dimostrata dal peso della varianza spiegata della PCA che si rileva più elevata quando i soggetti non sono uniformati, quindi, in assenza di uniformità non è possibile distinguere l'informazione utile da quella di rumore, né è possibile usare soltanto l'informazione derivabile dai valori di connettività funzionale per la scelta di una soglia. Questo evidenzia la sensibilità della procedura di PCA

all'applicazione di un fattore di scala sui dati originali e come tale caratteristica influenzi il contenuto informativo della prima componente principale.

In relazione all'analisi dei dati tramite test statistici, dal momento che i valori di skewness calcolati per le distribuzioni individuali hanno mostrato la mancanza del carattere di gaussianità, si riportano gli esiti del test non parametrico, la somma dei ranghi. L'ipotesi nulla è accettata con $p\text{-value}>0.05$, quando i dati vengono standardizzati in una fase successiva all'ottenimento dei campioni da confrontare. Se la standardizzazione è attuata preliminarmente, invece, l'ipotesi nulla non è accettata con un $p\text{-value}<0.01$. Di conseguenza, la prima componente di PCA è in grado di rappresentare il campione medio FC solo senza previa standardizzazione delle distribuzioni delle correlazioni. La normalizzazione influenza la possibilità di decorrelare i dati di connettività funzionale per poter eliminare fonti di rumore.

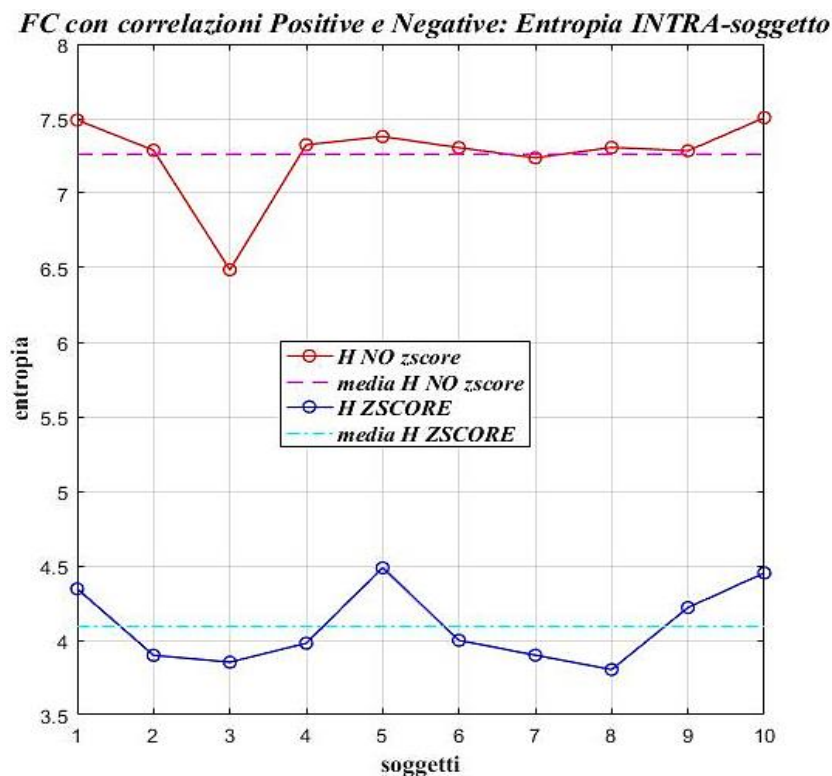


Figura4. 10 Entropia Intra-Soggetto per matrici di connettività funzionale originali per distribuzioni normalizzate e non

Infine, dal confronto dell'indice di entropia intra-soggetto, per le due tipologie di distribuzioni, si evince che tutti i soggetti mantengono un valore alto di entropia eccetto per il soggetto 3, come mostrato in **Figura4. 10**. Il valor medio si individua a circa 7.26. Al contrario, l'indice di entropia tende a decrescere quando si standardizzano le distribuzioni,

con un valore medio pari a 4.03 ed evidenziando un comportamento differente per i soggetti 1-5-10. In generale, lo stato di disordine dell'informazione, per le matrici non sogliate, è elevato e presenta una forte variabilità intra-soggetto.

4.1.3. *Connettività Funzionale: impatto dei valori di correlazione positivi*

Per le matrici di connettività funzionale contenenti solo valori di correlazione positivi, le connessioni negative sono azzerate e questo apporta lievi modifiche sui campioni rispetto al caso precedente. Considerando la procedura della PCA si riconferma una situazione analoga alla precedente, con la sola differenza che le percentuali di varianza spiegata sono più alte, per il caso di connettività funzionale normalizzata essa è pari a 60% rispetto al 59% di quello precedente, per il caso con le correlazioni negative è pari al 90%, invece di 88%, e si mantiene uno scarto del 30% tra i due casi di studio, come si può individuare nella **Figura 4.11**. La prima componente della PCA della matrice delle correlazioni, per sole correlazioni positive, racchiude un contenuto di informazione maggiore rispetto a quando si considerano anche le correlazioni negative. Tuttavia le considerazioni, svolte in precedenza, rimangono valide anche per tale situazione.

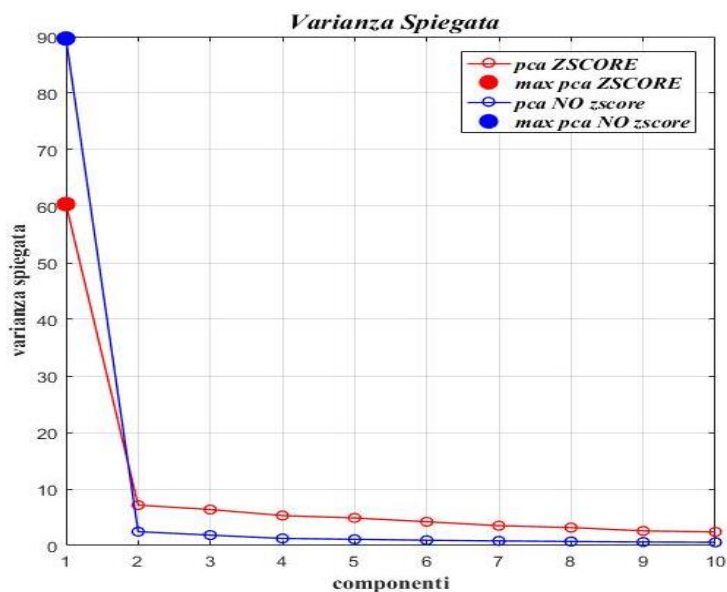


Figura 4.11 Varianza spiegata per mappe di connettività funzionale con correlazioni positive per distribuzioni normalizzate e non

In riferimento al test statistico della somma dei ranghi si ripropone l'analogo risultato del caso precedente. Motivo per cui la prima componente principale, solo se non equalizzata in principio, costituisce un adeguato sostituto del campione medio FC.

Nondimeno, l'indice di entropia registra un aumento del disordine informativo. Come evidenziato anche in precedenza, il trend entropico cambia se le distribuzioni sono standardizzate. Tuttavia, rispetto al caso precedente, i valori medi per ciascuno indice sono 7.45 e 3.95, rispettivamente più alto e più basso dei valori medi del caso precedente. Infatti, in presenza delle correlazioni negative, i valori medi sono stati 7.26 e 4.03 (**Figura 4.12**).

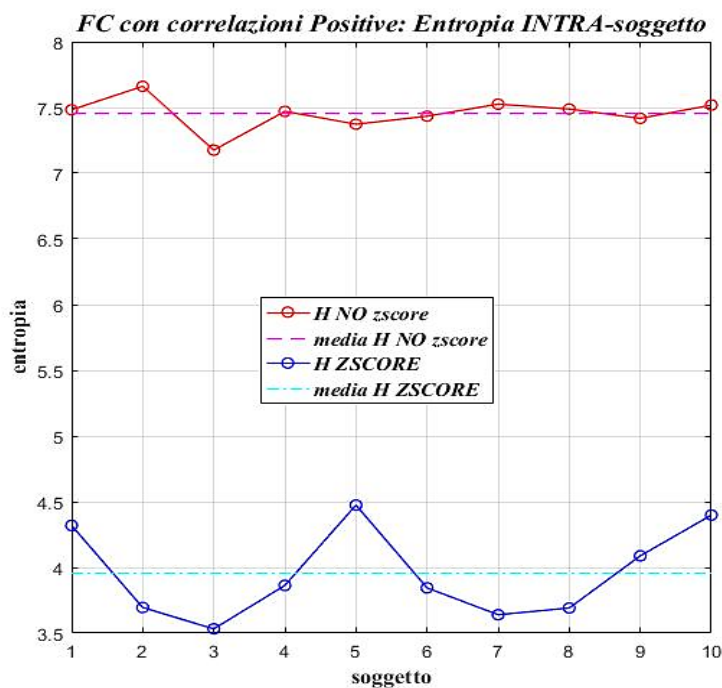


Figura 4.12 H: Entropia Intra-Soggetto per valori di correlazione positiva per distribuzioni normalizzate e non

4.1.4. Connettività Funzionale al variare della soglia

Il terzo caso esaminato si riferisce alle matrici FC a cui è stata applicata una soglia crescente compresa nel range $[-0.5745; 0.01; 0.92]$, con passo di un centesimo, con l'obiettivo di quantificare l'effetto della sogliatura sugli indici di network.

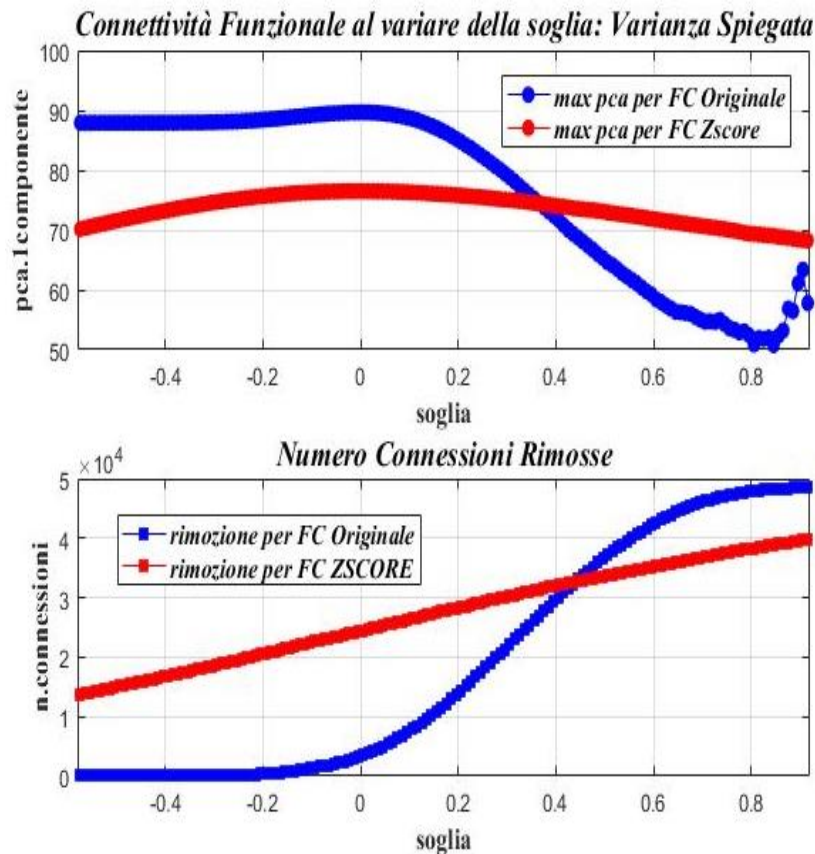


Figura 4.13 Varianza Spiegata e Variazione del numero di connessioni rimosse per le matrici di connettività funzionale valutate al variare della soglia, con distribuzioni normalizzate e non

Nel pannello superiore della **Figura 4.13**, sono rappresentati gli andamenti della varianza spiegata dalla prima componente della PCA, sia per le distribuzione dei valori di connettività funzionale normalizzate e non. In corrispondenza della distribuzione con valori di connettività funzionale positivi e negativi si osserva un plateau con valore massimo di circa 90% in corrispondenza della soglia prossima a zero, ossia quando la quasi totalità delle connessioni negative sono rimosse. In seguito, è evidente il suo decadimento dovuto alla rimozione delle connessioni tra i nodi. Per quanto riguarda le distribuzioni normalizzate, la varianza rimane pressoché costante, registrando il massimo valore pari al 60% in corrispondenza della stessa soglia e mantenendo almeno inizialmente una differenza di circa

30% tra i due andamenti. Nel pannello inferiore, invece, sono mostrati i trend relativi al numero di connessioni rimosse nel network soggettivo per distribuzioni con valori positivi e negativi di correlazione e normalizzati. Per il caso normalizzato si osserva una crescita lineare e omogenea all'aumentare della soglia, per il secondo caso, invece, si assiste ad un mantenimento delle connessioni per un valore di soglia prossimo a 0 e una successiva crescita all'aumentare della soglia fino al completo annullamento della rete cerebrale stessa.

Per quanto riguarda l'esito del test della somma dei ranghi, in maniera molto simile ai casi precedenti, l'ipotesi nulla è accettata quando i campioni sono normalizzati solo prima di effettuare le verifiche statistiche con un $p\text{-value} > 0.05$, in corrispondenza ad un valore di soglia maggiore di 0.4. Per il secondo caso, invece, l'ipotesi nulla è rifiutata con $p\text{-value} < 0.01$.

Seguendo l'andamento delle curve di entropia intra-soggetto media e di entropia di popolazione (*Figura 4.14*), in entrambi i casi si riconosce lo stesso trend che evidenzia un aumento progressivo dell'indice fino al raggiungimento di un picco, in prossimità della completa eliminazione dei pesi negativi.

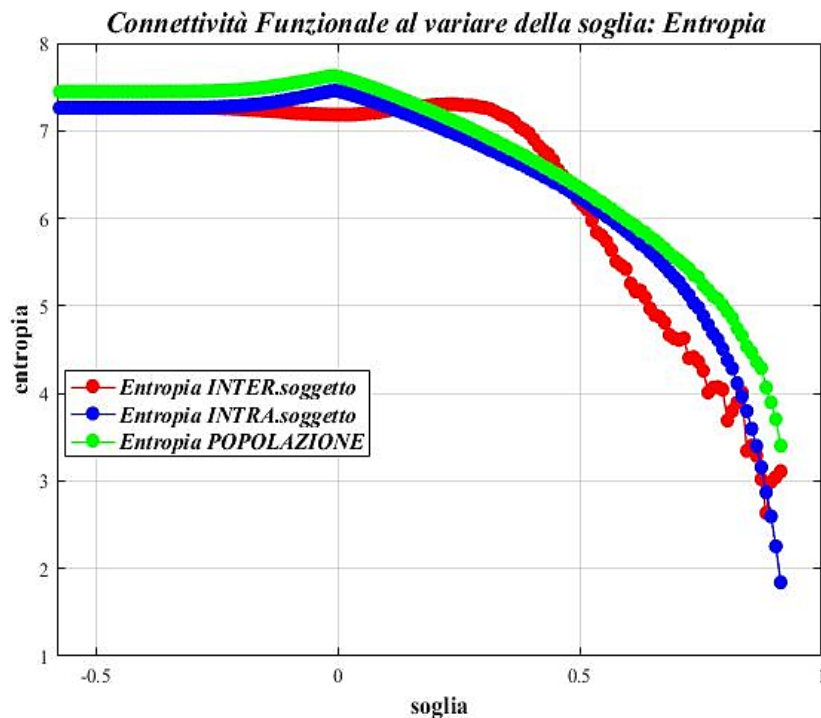


Figura 4.14 Entropia Intra-Soggetto, Inter-Soggetto e di Popolazione per matrici di connettività funzionale originali al variare della soglia assoluta

Successivamente, l'indice di entropia diminuisce in corrispondenza di valori alti per la soglia che prevedono anche la rimozione dei pesi positivi. Per l'indice di entropia inter-soggetto,

la curva assume un andamento più regolare e continuo al variare della soglia, mostrando un picco in prossimità del valore di soglia pari a 0.3.

Infine, gli indici di media, mediana e varianza, sono stati tutti misurati sul campione FC di popolazione. Essi hanno individuato dei dati di correlazione poco sparsi rispetto al *valore vero*, in quanto l'entità della varianza è molto piccola e tende a decrescere all'aumentare della soglia. La media e la mediana, invece, rilevano un continuo aumento all'aumentare della soglia, poichè rimangono esclusivamente le connessioni più forti. L'indice di media, quindi, è significativo e credibile.

4.1.5. Considerazioni Complessive sui Test Preliminari

In sintesi per tutte le analisi svolte e per tutti i casi considerati, le distribuzioni dei valori di correlazione mostrano una significativa differenza di due soggetti sui rimanenti analizzati, ciò ha portato a domandarsi se potessero essere identificati come outlier o no, quindi eliminabili dal resto dell'informazione. Inoltre, le distribuzioni, se standardizzate, risultano equalizzate e non rendono evidente la variabilità intra-soggetto, altrimenti preponderante. Tale evidenza rende particolarmente difficile assumere ipotesi di lavoro omogenee per tutti i soggetti in esame.

La prima componente principale racchiude la maggior parte di informazione quando i valori di connettività funzionale non sono alterati in alcun modo.

I test statistici hanno dato esito positivo per l'utilizzo della prima componente principale per la rappresentazione del campione medio della popolazione dei soggetti, in particolar modo in assenza di connessioni negative.

L'indice di entropia, infine, diminuisce in presenza solo dei pesi positivi con il raggiungimento di uno stato informativo più ricco e più omogeneo, permettendo di pronosticare un valore di soglia oltre al quale l'informazione non è fisiologicamente credibile.

In base ai risultati derivanti dai test preliminari si sono assunte le seguenti scelte progettuali quali: comprendere il ruolo delle connessioni negative attraverso le metriche di grafo prima di dettare una scelta definitiva, non alterare il segno della connessione, in quanto preponderante dal punto di vista fisiologico e per il significato di connettività tra i nodi della network di ciascun soggetto, dal momento che la variabilità intra-soggetto predomina e non permette di raggiungere la convergenza dei risultati, in tutti i possibili casi di studio. Tuttavia

la quantificazione della variabilità dell'informazione delle matrici FC ha evidenziato la difficoltà di determinare una possibile soglia da applicare ai dati basata sull'apporto informativo delle matrici FC, motivo per cui si è passati all'analisi delle metriche di grafo.

4.2. *Metriche di grafo*

Già dai risultati dei test preliminari, l'importanza del segno del valore della correlazione, la distinzione tra connessioni positive e negative e l'impossibilità di poterne attribuire un'adeguata spiegazione si sono rilevati fattori non trascurabili e degni di approfondimento. Infatti, in prima istanza è stato indagato l'impatto della compresenza delle connessioni negative e positive sul calcolo delle metriche di grafo. Esaminando la quantità complessiva delle connessioni negative e positive, prese singolarmente, si è notato come il contenuto negativo costituisse una minima parte dei link delle network studiate. Dal momento che costituiscono il 7% dei valori di correlazioni all'interno delle matrici FC, i pesi negativi sono risultati difficilmente valutabili attraverso le metriche di grafo. In supporto a tale osservazione, sono state ispezionate le distribuzioni delle metriche di grafo più utilizzate in letteratura, quali: *betwenness centrality*, *clustering coefficient* e *local efficiency*. Nella

Tabella 4.3 vengono riportati gli indici di skewness medio, gli indici di curtosi medio e i coefficienti di variazione medio per le singole metriche relative ai grafi dei singoli soggetti.

<i>Indice</i>	<i>Betwenness Centrality</i>	<i>Clustering Coefficient</i>	<i>Local Efficiency</i>
<i>Skewness</i>	2.82	-0.87	-0.88
<i>Coefficiente di variazione</i>	0.28	0.5381	0.4866
<i>Curtosis</i>	12.42	3.52	3.55

Tabella 4.3 *Indici di variabilità di grafo di matrici di connettività positive*

Dai dati si evince che le distribuzioni dei valori di tutte le metriche per matrici di connettività positive sono asimmetriche. I valori di curtosi rivelano curve di distribuzioni leptocurtiche per tutte e tre le metriche. Infine, il coefficiente di variazione mostra una maggiore credibilità del valor medio della *skewness* per la *betweenness centrality*, rispetto alle altre due misure. Per matrici di connettività negative non è stata riportata alcuna tabella in quanto non è stato possibile quantificare le caratteristiche delle distribuzioni delle metriche. La natura sparsa delle matrici di connettività per valori di correlazioni negativi, infatti, ha influenzato molto la quantificazione degli indici.

Inoltre, lo studio della soglia per intervalli simmetrico, asimmetrico e con un valore fisso hanno sottolineato la difficoltà nella scelta di un range adeguato per l'incapacità di eliminare solamente gli elementi di rumore e non l'informazione utile.

4.2.1. Entropia di metrica di grafo vs Entropia Connettività Funzionale

Noti gli andamenti dell'entropia intra-soggetto, intra-soggetto media, inter-soggetto, di popolazione e di matrice relativo ai dati di connettività funzionale per valori positivi di correlazione, ricavati dai test precedenti, si è cercato di caratterizzare le metriche confrontando l'entropia di ciascuna misura con l'entropia delle mappe FC, al variare della soglia. Al fine di caratterizzare non solo le metriche più note in letteratura ma anche altre più avanzate, l'analisi è stata estesa ad altre misure, quali: *global efficiency e strength*.

L'analisi del trend entropico a livello intra-soggetto per tutte le metriche analizzate: *betweenness centrality, clustering coefficient, local efficiency, global efficiency e strength* ha marcato una forte variabilità intra-soggetto al variare della soglia. Nella **Tabella 4.4** sono riportati per i 10 soggetti e per ogni metrica il valore medio dell'entropia della metrica del singolo soggetto al variare della soglia e la deviazione standard. I valori medi risultano non consistenti in quanto i valori di deviazione standard sono elevati rispetto al *valore vero*. Le misure, quindi, hanno bassa precisione dal momento che esse sono molto distanti dal valor medio. Inoltre, si assiste ad un'elevata variabilità intra-soggetto per piccole variazioni di soglia, dimostrata da valori alti di coefficiente di variazione, non riportati in questo elaborato.

La *strength*, invece, descrive un comportamento speculare al trend entropico, non rivela alcun cambiamento all'aumentare della soglia fino ad un valore circa pari a 0.4 per poi

aumentare e annullarsi in corrispondenza dell'eccessivo impoverimento della network. In coincidenza del punto di svolta dell'andamento della strength si registra la rimozione del 60,5 % delle connessioni cerebrali.

A livello inter-soggetto, invece, le curve degli indici di entropia denotano un trend simile tra gruppi di metriche. Per *betwenness centrality* e *strength* si può notare dalla **Figura 4.15** un'insensibilità a soglia fino ad un valore pari a 0.4, associata a valori elevati di entropia relativi alle mappe FC. Superato il punto critico si assiste all'incremento della misura in linea alla continua eliminazione degli archi del grafo e depauperamento informativo della rete cerebrale.

<i>SOG</i>	<i>H_GE</i>	<i>H_LE</i>	<i>H_BC</i>	<i>H_CC</i>	<i>H_STR</i>
<i>Sog1</i>	0.5418±1.6204	0.3630±1.1700	0.0754±0.2336	0.4825±1.4926	0.0659±0.3467
<i>Sog2</i>	0.5271±1.5730	0.4565±1.4210	0.0625±0.2038	0.5067±1.5782	0.0763±0.3382
<i>Sog3</i>	0.4159±1.2970	0.3671±1.2535	0.0529±0.1944	0.4029±1.3635	0.0974±0.4742
<i>Sog4</i>	0.4949±1.5266	0.3898±1.2721	0.0583±0.1999	0.4541±1.4651	0.0749±0.3719
<i>Sog5</i>	0.5062±1.5395	0.3465±1.1494	0.0664±0.2146	0.4427±1.4111	0.0649±0.3238
<i>Sog6</i>	0.4585±1.4081	0.3746±1.2466	0.0509±0.1851	0.4478±1.4493	0.0739±0.3857
<i>Sog7</i>	0.4853±1.4711	0.4279±1.3455	0.0598±0.2031	0.4848±1.5114	0.0759±0.3736
<i>Sog8</i>	0.4762±1.4321	0.3960±1.2783	0.0556±0.1948	0.4677±1.4847	0.0781±0.3672
<i>Sog9</i>	0.4743±1.4593	0.3632±1.1979	0.0517±0.1862	0.4290±1.3943	0.0900±0.4354
<i>Sog10</i>	0.5307±1.5931	0.3976±1.2402	0.0748±0.2349	0.4771±1.4670	0.0626±0.3017

Tabella 4.4 Indici di confronto media deviazione standard per i 10 soggetti al variare della soglia: *Sog*: soggetto; *H*: entropia; *GE*: global efficiency; *LE*: local efficiency; *BC*: *betwenness centrality*; *CC*: clustering coefficient; *STR*: strength

Per *global efficiency*, *local efficiency* e *clustering coefficient* lo stato di disordine dell'informazione si mantiene molto elevato fino ad una soglia pari a 0.6 per poi decrescere rapidamente in linea con l'eliminazione dei collegamenti inter-cerebrali (**Figura 4.15**).

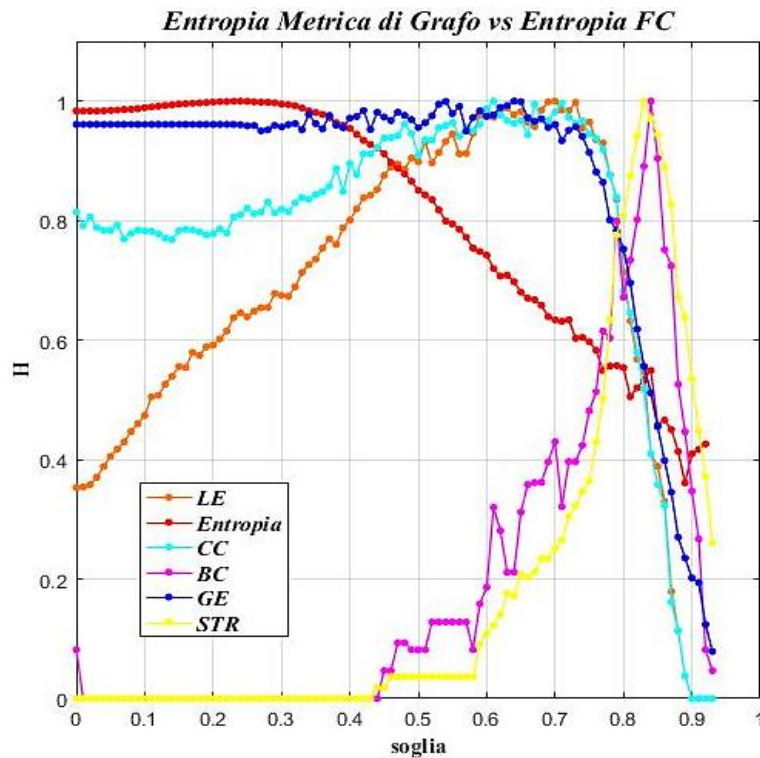


Figura 4.15 Entropia di metrica di grafo vs Entropia FC Inter-soggetto GREYNA. LE: Local Efficiency; H: entropia; CC: Clustering Coefficient; BC: Betweenness Centrality; GE: Global Efficiency; STR: strength

Analoga situazione ai risultati ottenuti a livello intra-soggetto si osserva per l'entropia di popolazione e di matrice. Per questo motivo la fase successiva di analisi è stata ristretta solo al livello intra-soggetto e inter-soggetto.

Come specificato nel paragrafo 3.3.4 l'implementazione delle metriche è stata eseguita attraverso l'utilizzo di due software, quali GREYNA e BCT. Il software di recente sviluppo, GREYNA, presenta una maggiore precisione di calcolo dei valori della metrica e una sostanziale differenza nel calcolo della *global efficiency* dal momento che enfatizza l'entità della distanza tra coppie di nodi del grafo. Tuttavia i range di misura e i trend in tutte le metriche sono simili per la maggior parte dei casi, se non addirittura identici. I valori, quindi, che si riportano fanno riferimento al software GREYNA.

Il confronto tra le entropie, purtroppo, non ha permesso di stabilire un metodo efficace per la caratterizzazione delle metriche, in quanto estremamente variabile a livello intra-soggetto e non permette di individuare un valore di soglia comune per tutte le metriche. La

metodologia risulta non ripetibile né robusta. Tuttavia, attraverso questo tipo di analisi tutte le metriche sono risultate poco sensibili e robuste alla soglia per valori più bassi o per valori più alti.

Infine, la modellizzazione dell'entropia tramite linearizzazione della curva non ha aggiunto risultati accettabili per l'individuazione della soglia. Il valore del punto critico della soglia, infatti, relativo ad ogni curva entropica di ciascuna metrica varia ad ogni livello di indagine. Per la *strength* si registrano come punti critici: 0.4 a livello intra-soggetto e 0.6 a livello inter-soggetto. Analoga situazione si riscontra per la *betweenness centrality*, a livello inter-soggetto, però, la soglia aumenta di un decimo. Per *clustering coefficient*, *global efficiency* e *local efficiency* il punto di ginocchio si registra a 0.6 a livello intra-soggetto e 0.75 per inter-soggetto. Si riportano solo le distinzioni a livello intra-soggetto e inter-soggetto in quanto gli andamenti delle curve entropiche a livello di popolazione e di matrice ripropongono il caso intra-soggetto. Di conseguenza si è affrontato il problema attraverso i *survival plot*.

4.2.2. Survival Plot

Tramite i *survival plot* si sono caratterizzate in modo più puntuale le metriche di grafo. A questo stadio dell'indagine, l'andamento dell'entropia della connettività funzionale è stato confrontato direttamente con il trend delle metriche. L'insieme delle misure di grafo è stato arricchito con altre metriche che hanno permesso di caratterizzare a livello globale le network cerebrali, quali: *modularity* e *char path*.

I casi di studio sono stati in tutto 6 al variare della soglia. Per i casi PAN e PPN, in cui i valori FC sono sottoposti alla procedura di normalizzazione, si assiste ad un decorso uniforme delle metriche. Tutte le misure analizzate quali: *modularity*, *char path*, *global efficiency*, *strength*, *clustering coefficient*, *local efficiency* e *betweneess centrality* presentano un andamento lineare e omogeneo tra tutti i soggetti. Tale uniformità si verifica anche nelle distribuzioni statistiche al variare della soglia, durante l'ispezione visiva delle immagini rappresentanti le matrici dei valori delle metriche ricavate, sia nell'analisi intra-soggetto sia in quella inter-soggetto e di popolazione; scompare del tutto la variabilità intra-soggetto, i valori della *global efficiency* dopo la presenza di un plateau iniziale decrescono all'aumentare della soglia, la *modularity* tende a saturazione al crescere della soglia e il *char path* durante la sua crescita si arresta per una soglia pari a 0.4. Per la *strength* la curva è

simile alla *global efficiency*, il *clustering coefficient* assume un comportamento simile al *char path*, analogo trend si osserva per la *local efficiency*. La *betweenness centrality*, invece, mostra una decrescita esponenziale. In corrispondenza di tutte le metriche questa tipologia di dinamica evidenzia robustezza e insensibilità al cambiamento della soglia. Inoltre, la normalizzazione delle distribuzioni rende le valutazioni sulle metriche inattendibili a livello fisiologico. In particolare, l'applicazione dello z-score rende ancora più critica l'analisi di grafo. Tramite il calcolo della metrica *node degree* al variare della soglia imposta, infatti, si è potuta osservare un'alta variabilità della quantità dei nodi rimossi all'interno del grafo di ciascun soggetto. Questa osservazione non può coesistere con l'uniformità delle distribuzioni. Pertanto, le analisi di grafo successive sono state svolte non normalizzando le distribuzioni ma cercando di aumentare il carattere di gaussianità delle stesse.

Se si considerano i valori FC sia positivi sia negativi o con l'applicazione della trasformata z-Fisher, per i quali i casi di studio si identificano in PAF, PPF, PON e POP, non si apprezzano sostanziali differenze. La trasformata z-Fisher ha la proprietà di aumentare la gaussianità delle distribuzioni ottenute con l'analisi di correlazione di Pearson, rendendo le distribuzioni asimmetriche e leptocurtiche. L'asimmetria, in particolare, si verifica nello spostamento della mediana verso valori più elevati o inferiori nei boxplot delle singole misure. Si è notato che lo studio dei grafi è risultato più agevole con matrici z-Fisher trasformate.

Di conseguenza, si è scelto di affrontare le verifiche sui 15 soggetti restanti utilizzando soltanto la trasformata z-Fisher per i dati delle matrici FC. Di seguito sono riportati i grafici dei survival plot intra-soggetto (**Figura 4.16**) ed inter-soggetto (**Figura 4.17**) relativi al solo caso di valori FC con trasformazione z-Fisher per grafi pesati e non.

Dai grafici si può notare come la *global efficiency* mantenga un andamento costante fino ad un valore di soglia pari a 0.3. Dopo tale punto critico i valori delle metriche decrescono all'aumentare della soglia. Sebbene l'andamento della curva è simile in tutti i soggetti, i valori della *global efficiency* mettono in risalto un'accentuata variabilità intra-soggetto. Per alcuni soggetti si registrano valori molto bassi, inferiori a 0.4, per altri, invece, quantità elevate, maggiori a 0.5.

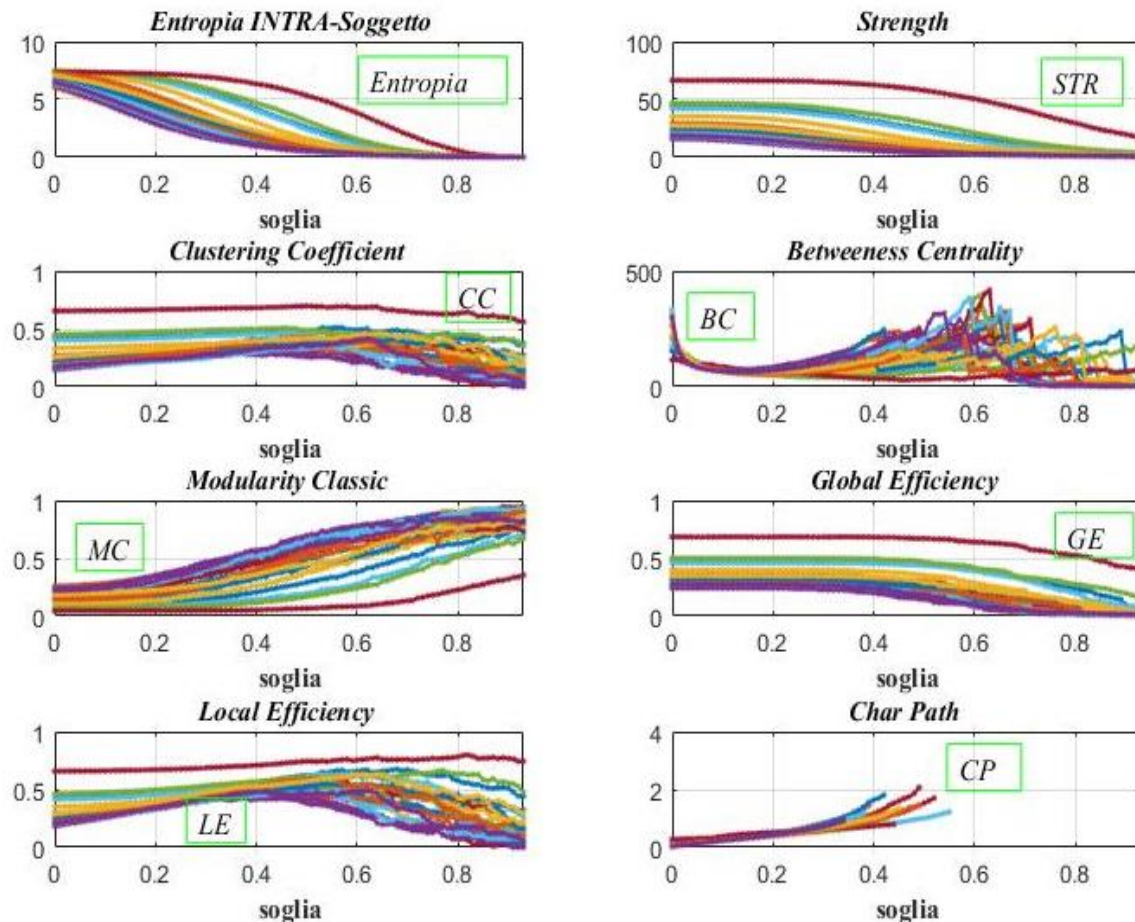


Figura 4.16 Survival Plot Intra-Soggetto per il caso PPF e 25 soggetti; LE:local efficiency, CP:char path, GE:global efficiency, MC: modularity classic, BC:betweenness centrality, STR:strength

La *modularity* giunge a saturazione al variare della soglia sia quando si considerano moduli piccoli, sia medi e grandi. I moduli sono definiti piccoli se ciascun raggruppamento di nodi contiene in media una percentuale di connessioni pari al 24%, medi se contengono in media il 48 % delle connessioni, grandi se contengono il 97% degli archi dei grafi. Il suo trend è speculare a quello entropico. A livello intra-soggetto la metrica possiede una bassa sensibilità a soglia pari a 0.2 (**Figura 4.16**), caratteristica che la rende robusta a piccole variazioni di soglia. Le immagini relative alle matrici della metrica, infatti, contraddistinguono una spiccata uniformità tra i soggetti. Inoltre, con la modularità si è in grado di identificare un possibile limite superiore nella scelta della soglia da applicare alle matrici FC. Tale limite è stabilito intorno a 0.3 se si notano gli andamenti della *modularity* nei *survival plot* sia a livello intra-soggetto sia inter-soggetto. Al passaggio del punto critico la variazione intra-soggetto aumenta e rende la metrica più sensibile alla continua rimozione di ulteriori archi nei grafi. Di maggiore interesse e rilevanza è la modularità riferita a moduli

di piccola grandezza o nella tipologia classica, in quanto si ricorda che la segmentazione delle ROI, effettuata in base all'atlante, è caratterizzata principalmente da regioni di tali dimensioni, non superiori. Difatti, la modularità per regioni grandi presenta variazioni puntuali dell'indice, fattore che indica sensibilità a piccolissime variazioni di soglia. La modularità, quindi, ha una migliore capacità di lavoro per range di misura intermedi e di svolgere indagini inter-soggetto a livello macroscopico. Per quanto riguarda l'implementazione della modularità, sono stati usati due algoritmi: *Danon* e *Newman*. Nessuna differenza significativa è stata registrata tra i due. Si può consigliare, dunque, l'utilizzo dell'algoritmo di ottimizzazione Newman in quanto detiene una più alta efficienza computazionale. Nei survival plot riportati i dati fanno riferimento all'algoritmo di Newman.

La *char path* come la modularità aumenta al crescere della soglia ed evidenzia una minore sensibilità per soglie iniziali, tuttavia, svela un brusco arresto per valori di soglia vicini a 0.3 dovuto alla definizione stessa della metrica, come è evidente dalla **Figura 4.16**. Tale interruzione evidenzia che non è possibile applicare una soglia maggiore del limite superiore, dal momento che la *char path* è un indice dipendente dalla soglia. Le immagini delle matrici della metrica di grafo appaiono binarie proprio in corrispondenza della soglia appena indicata, poiché dopo 0.3 l'indice è affetto da cammini molto lunghi e poco efficienti e questa situazione si ripropone per tutti i soggetti.

Un andamento simile alla *global efficiency* si riscontra anche per *strength* e *clustering coefficient*. A livello intra-soggetto essi marcano una forte variabilità, distinguendo individui con valori estremamente alti da altri con valori estremamente bassi. Tuttavia la *strength* ha una minore resistenza alla soglia rispetto alla *global efficiency*, in quanto i suoi valori decadono prima. Il *clustering coefficient*, invece, mostra una resistenza superiore non solo alla *strength* ma anche alla *global efficiency*, dato il suo andamento pressoché uniforme per tutti i valori di soglia, con un lieve picco per la soglia pari a 0.5 e decremento in corrispondenza dell'eccessiva rimozione degli archi del grafo, come anche la *local efficiency*. Le immagini evidenziano uniformità di valori per la maggior parte dei soggetti tranne per coloro che costituiscono gli estremi dell'intervallo della misura in questione.

La *betweenness centrality* denota un comportamento simile a *modularity* e a *char path*. Essa decresce all'aumentare della soglia con un trend speculare a quello entropico e presenta

una minore sensibilità intra-soggetto per soglie di piccola entità, caratteristica che la rende più robusta all'applicazione della soglia.

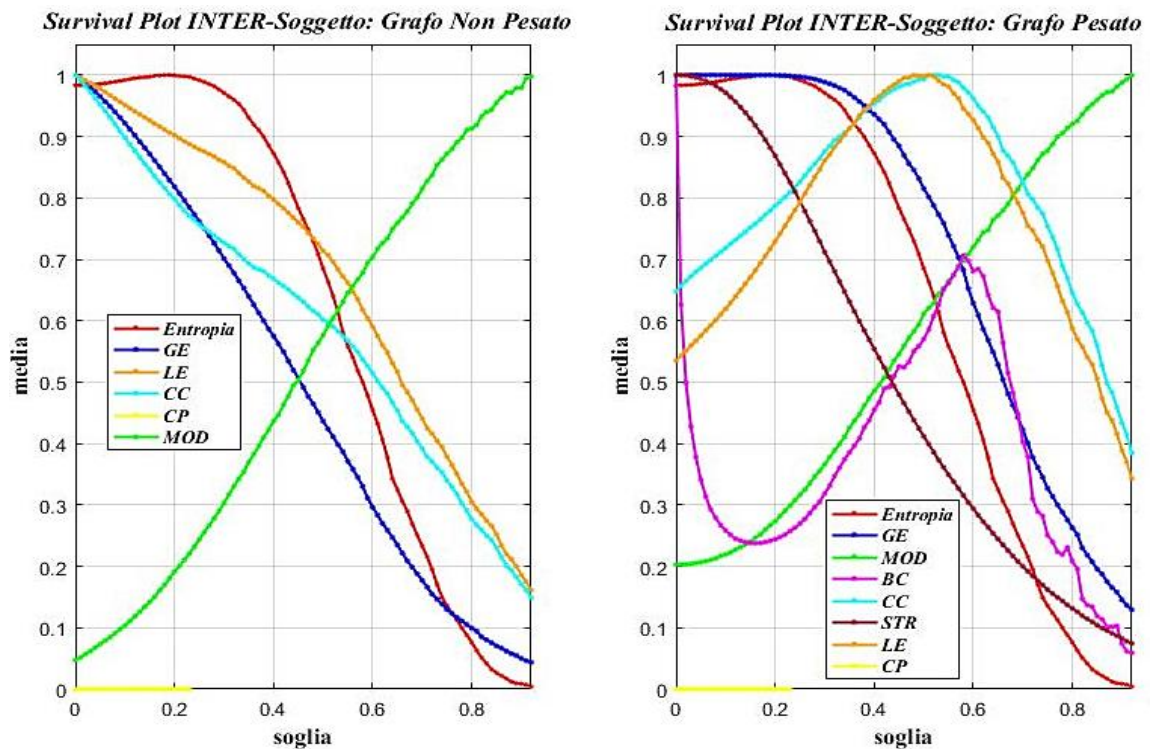


Figura 4.15 Survival Plot Inter-soggetto caso PPF e PAF

A questo livello dello studio è stata esaminata anche la differenza tra grafi pesati e non pesati. La prima evidenza è che non tutte le metriche sono in grado di lavorare con grafi non pesati. *Strength* e *Betweenness centrality* non sono definite per grafi non pesati, di conseguenza non appaiono rappresentate in **Figura 4.17**. Mentre le altre metriche implementate possono essere applicate sia a grafi pesati sia ai non pesati. La differenza nelle due tipologie di grafi risiede nella maggiore sensibilità, nella robustezza delle metriche e nei range di misura, per questo motivo sono di seguito riportati i grafici della variabilità e sensibilità rispetto alla soglia delle metriche, nel caso di FC con applicazione della trasformate z- Fisher e grafi pesati e non, **Figura 4.17**.

Nel caso di grafi pesati la *global efficiency* assume valori più alti, maggiore robustezza alla soglia e minore sensibilità alla stessa rispetto al caso non pesato in cui le considerazioni sono complementari. Il *clustering coefficient* mostra maggiore robustezza, minore sensibilità e un range di valori simile per grafi non pesati, stesso andamento per la *local efficiency*. Nel caso pesato, invece, si assiste ad una decrescita pressochè lineare all'aumentare della soglia che ne determina un comportamento completamente opposto al precedente. La *modularity* evidenzia lo stesso trend per entrambe le tipologie di grafo in termini di variabilità, il range

di valori è più alto e la sensibilità è minore nel caso di grafo pesato. Infine, la *char path* non presenta sostanziali cambiamenti tra grafi pesati e non, sia per robustezza sia per sensibilità e range di misura.

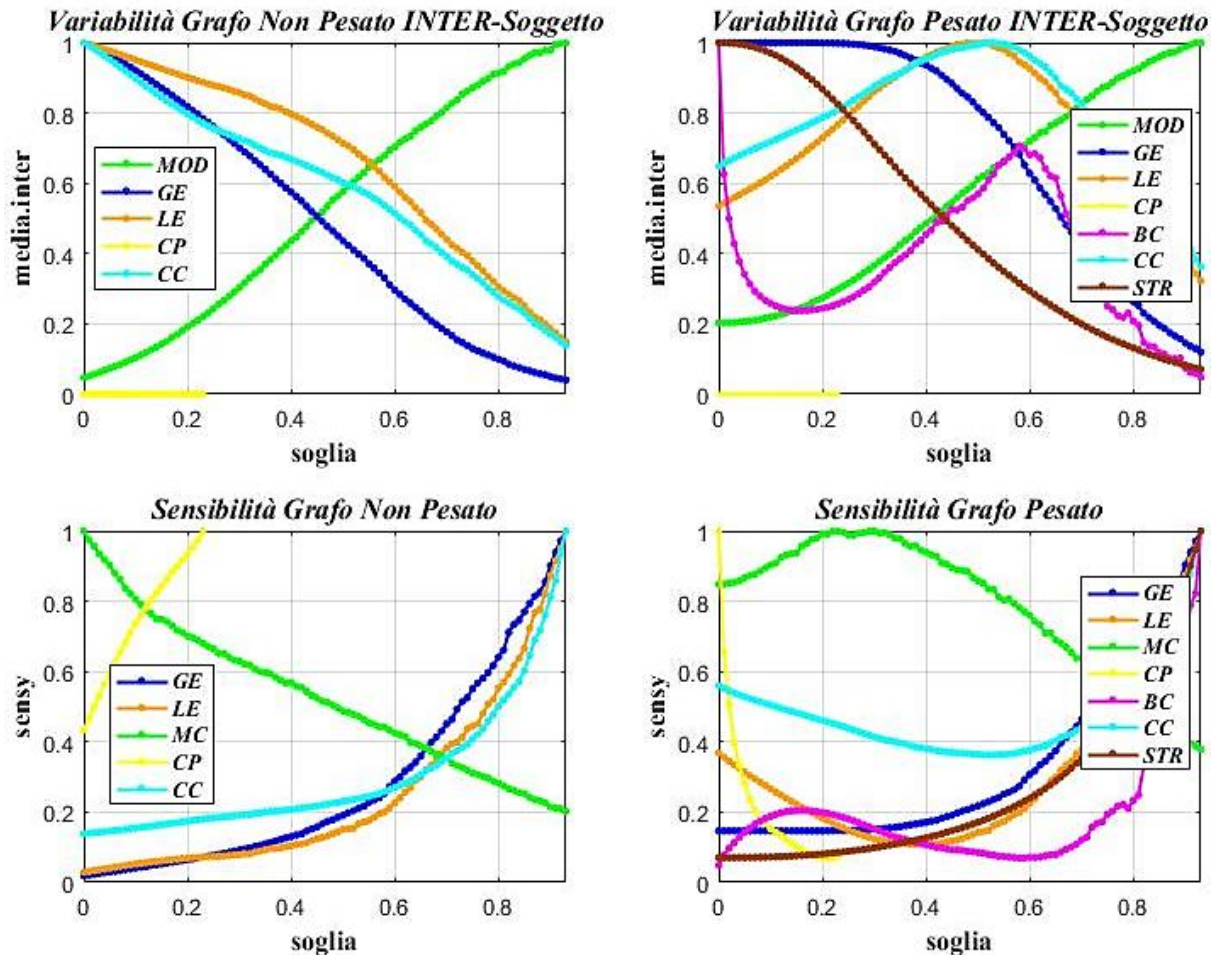


Figura 4.17 Variabilità e sensibilità a soglia metriche di grafo inter-soggetto

Infine, l'analisi ha riguardato anche i 15 soggetti rimanenti. I casi di studio sono stati ristretti soltanto a PAF e PPF, dal momento che la trasformata z-Fisher ha reso più agevoli le valutazioni delle metriche. I risultati ottenuti per i 10 soggetti si sono ripetuti anche per il gruppo completo di individui. Le uniche differenze rispetto al caso precedente si notano nella diminuzione del punto critico segnalato pari a 0.2 per la *char path* e una maggiore sensibilità a soglia delle metriche nel caso di grafi non pesati. I grafici riportati fanno riferimento al gruppo completo di soggetti del data-set.

4.2.3. *Analisi topologica dei grafi*

L'analisi svolta sui grafi mediante le metriche di centralità, integrazione e di segregazione ha evidenziato una notevole complessità nella ricerca del valore di soglia. Nessuna metrica considerata singolarmente si è rivelata un marcatore adeguato per la scelta della soglia e delle caratteristiche intrinseche delle network dei singoli soggetti. Motivo per cui si è cercato di semplificare l'analisi scegliendo un indice di più pratico utilizzo nella caratterizzazione delle network dei singoli soggetti. A questo punto dell'indagine, infatti, è stata quantificata l'entropia topologica. In questo modo è stato possibile capire il contenuto informativo a livello topologico di ciascuna rete cerebrale e quali ROI contenessero la maggior parte di esso. Dapprima per il gruppo ristretto di 10 soggetti, a livello intra-soggetto si è notata una rapida decrescita esponenziale dell'indice. Il disordine informativo a livello strutturale è molto alto se contiene tutte le connessioni positive ma si azzerava per una soglia pari a 0.3. Tale prospetto è ripetibile per tutti i soggetti, in quanto a livello intra-soggetto il coefficiente di variazione intra-soggetto varia da 0.01 a 0.4 molto lentamente fino al punto critico. Esso, quindi, si mantiene molto basso, inferiore a 0.4 per soglie iniziali, successivamente si registra una rapida crescita. La variabilità intra-soggetto tende a crescere in corrispondenza del limite superiore 0.3. In contemporanea il diametro tra le ROI aumenta e assume valore infinito in corrispondenza dello stesso punto critico. In questo modo conferma questo valore come limite superiore per un'eventuale soglia dei dati e marca che le aree maggiormente informative sono le più lontane tra loro. Tali considerazioni sono verificate anche per i restanti 15 soggetti, con la differenza che il limite superiore diminuisce di un decimo (*Figura 4.18*). Tale metrica appare poco sensibile al variare della soglia per valori piccoli, robusta al suo cambiamento e fortemente ripetibile.

Dalla *Figura 4.18* è possibile notare come le considerazioni descritte finora sono valide non solo per i grafi pesati ma anche per i non pesati.

Inoltre, se consideriamo i valori di media e deviazione standard nel caso di grafi pesati per i 10 soggetti essi sono 3886 ± 6621 con coefficiente di variazione pari a 1.9539 per i 25 soggetti, 3509 ± 6793 e coefficiente di variazione pari a 1.936. Per grafi non pesati, sia per i 25 soggetti sia per 10 soggetti non si osserva alcuna variazione dei valori riportati per i grafi pesati. Questi dati dimostrano che i risultati ottenuti dai 10 soggetti sono ripetibili per i 25 soggetti ma il valore medio non è consistente in quanto i dati sono molto sparsi e mancano di precisione, dimostrati da un coefficiente di variazione elevato.

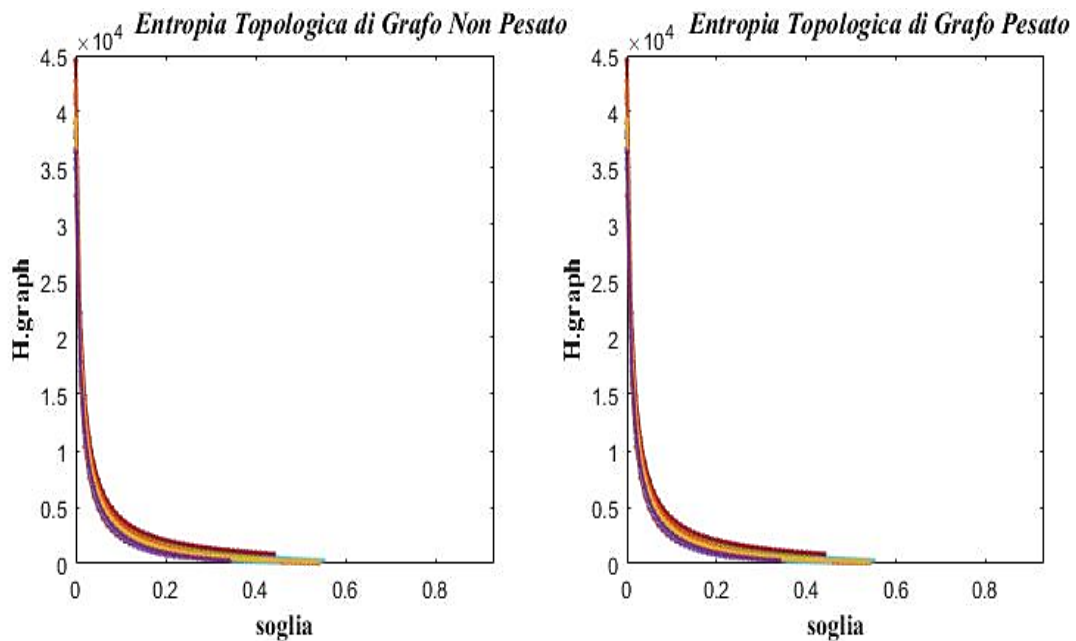


Figura 4.18 Entropia Topologica Caso PPF e PAF

Inoltre, è stata effettuata l'analisi topologica tramite la costruzione dei grafi tridimensionali relativi ai soggetti al fine di comprendere visivamente l'impatto della soglia sui grafi. In questo elaborato sono riportati come esempio i grafi relativi ad un solo soggetto e per una sola metrica, la betweenness centrality, per due livelli di soglia. La rappresentazione è stata riportata per una visione sagittale del lobo destro della superficie cerebrale ed è stata ottenuta tramite il software BrainNetViewer. I nodi rappresentano le ROI e gli acronimi indicano il nome delle singole aree volumetriche, dichiarate in appendice

Tabella 6 e Tabella 7. Nella **Figura 4.19** e **Figura 4.20** la dimensione dei nodi varia con la quantità della metrica, più è alto il valore più il nodo è grande. Analoga considerazione vale per gli archi. Maggiore è il peso della connessione tra i nodi più spesso è il link. Per il valore di soglia pari a 0.3, punto critico per la totalità delle metriche, il grafo è depauperato del 44,5% delle connessioni totali, per una soglia pari a 0.7 manca il 94% delle connessioni. E' possibile osservare visivamente una rete meno densa e meno fitta nel secondo caso rispetto al grafo corrispondente al punto critico. Questo conferma la mancanza di credibilità fisiologica a livello cerebrale dei range delle metriche, al superamento della soglia 0.3.

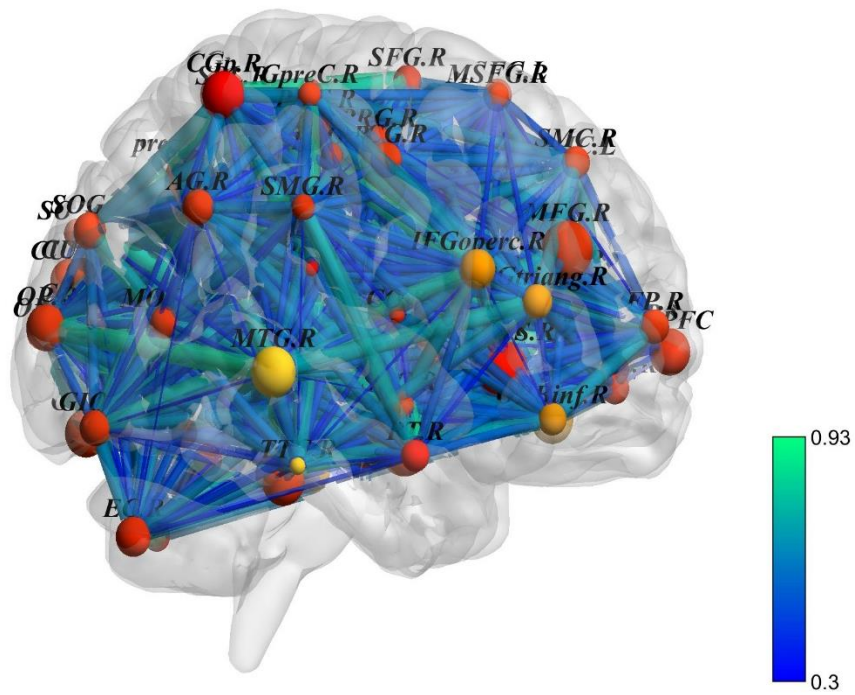


Figura 4.19 Grafo soggetto 1 per la rappresentazione della betweenness centrality con soglia applicata ai valori di correlazione pari a 0.3

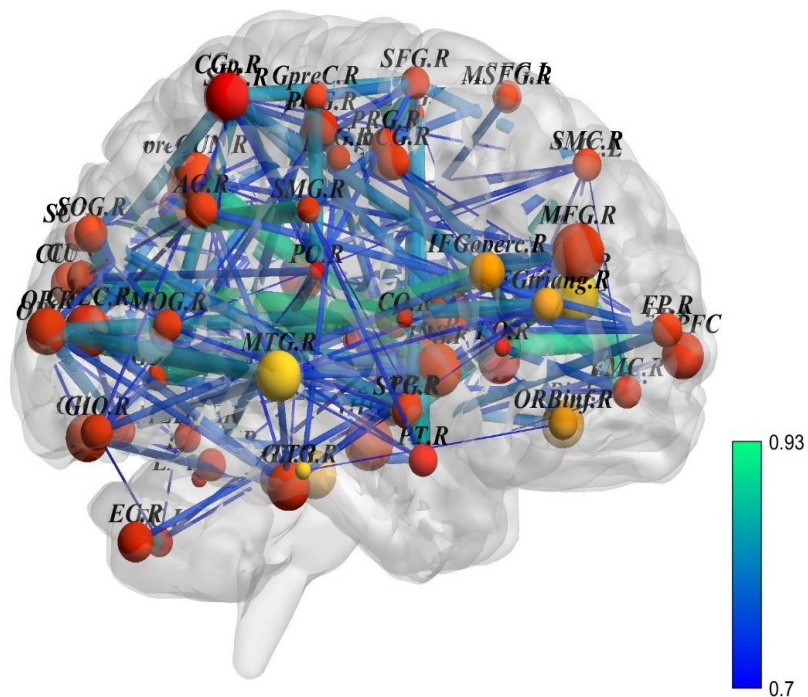


Figura 4.20 Grafo soggetto 1 per la rappresentazione della betweenness centrality con soglia applicata ai valori di correlazione pari a 0.7

4.2.4 Analisi robustezza e sensibilità complessiva

Per riassumere le informazioni in termini di variabilità e sensibilità a soglia per ciascuna metrica e per grafi pesati e non, in relazione a mappe FC con valori di correlazioni positive e z-Fisher trasformate, si riporta di seguito la **Tabella 4.5**. Con gli acronimi G^{UW} e G^W si fa riferimento rispettivamente a grafo non pesato e grafo pesato per i quali sono state implementate le metriche. Il range di valori di variabilità e di sensibilità si identificano in: *bassa* per l'intervallo $[0, 0.3]$, *alta* per range $]0.3, 0.92]$.

Per la *char path*, invece, si indica con *altissima* l'intervallo $]0.92, \text{Inf}]$. In primo luogo, è necessario notare come non tutte le metriche siano in grado di lavorare con grafi non pesati. Questo si evidenzia per le misure: *strength* e *betweenness centrality*. Nella tabella, infatti, tale condizione è indicata con $N G^{UW}$, ossia Not UnWeighted Graph. Per quanto riguarda, quindi, la stabilità delle metriche di grafo è necessario studiare le combinazioni *valore della soglia-tipo di grafo*.

Per grafi pesati e soglie inferiori a 0.3 tutte le metriche mostrano una bassa sensibilità all'applicazione della soglia tranne per la *modularity* in cui si registra un'alta sensibilità alla variazione della soglia. Inoltre, tra le misure poco variabili si distingue la *modularity*. Le misure con sensibilità bassa costituiscono le metriche più robuste e stabili a piccole variazioni della soglia. Tale osservazione non appartiene alla *modularity*. La *char path*, invece, registra un arresto brusco, dovuta alla definizione stessa della metrica. Si rileva, dunque, una misura dipendente dalla soglia perché fin da valori bassi appare estremamente sensibile a minime variazioni. Per questo motivo la *char path* non può essere considerata una metrica affidabile per l'analisi globale del grafo.

La *modularity* per soglie inferiori a 0.3 fornisce una bassa densità di archi tra le ROI intra-modulo, ma alta inter-modulo e una suddivisione in moduli poco efficace. Dal momento che i valori sono positivi, la modularità indica che il numero di archi presenti nelle reti neurali dei soggetti sono maggiori di quelli attesi [34]. In associazione ai valori bassi di modularità, però, la *clustering coefficient* permette di comprendere l'assetto geometrico dei moduli, i quali includono tante triangolarità al loro interno.

La *char path*, invece, indica la presenza di pochi cammini brevi e più efficienti tra le ROI. La *local efficiency* mostra un'alta variabilità e bassa sensibilità, indicando che le connessioni tra i moduli sono forti sebbene siano distanti. A livello globale l'efficienza delle reti neurali è preservata dai cammini lunghi, come si evince da valori superiori a 0.3 per la *global efficiency*.

	<i>Soglia ≤ 0.3</i>		<i>0.3 < Soglia ≤ 0.92</i>	
G^{UW} G^W	<i>Variabilità</i>	<i>Sensibilità</i>	<i>Variabilità</i>	<i>Sensibilità</i>
<i>Modularity</i> <i>Classic</i>	BASSA BASSA	ALTA ALTA	ALTA ALTA	BASSA BASSA
<i>Global</i> <i>Efficiency</i>	ALTA ALTA	BASSA BASSA	ALTA BASSA	ALTA ALTA
<i>Local</i> <i>Efficiency</i>	ALTA ALTA	BASSA BASSA	BASSA ALTA	ALTA ALTA
<i>Char Path</i>	ALTIS.MA ALTIS.MA	ALTIS.MA ALTIS.MA	ALTIS.MA ALTIS.MA	ALTIS.MA ALTIS.MA
<i>Clustering</i> <i>Coefficient</i>	ALTA ALTA	BASSA BASSA	BASSA ALTA	ALTA ALTA
<i>Betweenness</i> <i>Centrality</i>	$N G^{UW}$ ALTA	$N G^{UW}$ BASSA	$N G^{UW}$ ALTA	$N G^{UW}$ BASSA
<i>Strength</i>	$N G^{UW}$ ALTA	$N G^{UW}$ BASSA	$N G^{UW}$ BASSA	$N G^{UW}$ ALTA
<i>Entropia</i> <i>Topologica</i>	ALTA ALTA	BASSA BASSA	BASSA BASSA	ALTA ALTA

Tabella 4.5 Caratterizzazione metriche di grafo sulla base della variabilità e sensibilità a soglia; G^{UW} : UnWeighed Graph; G^W : Weighed Graph; $N G^{UW}$: Not UnWeighed Graph; Bassa: $[0, 0.3]$, Alta $]0.3, 0.92]$; Altissima $]0.92, Inf[$.

Per la *betweenness centrality*, il valor medio elevato è indicatore di una grande quantità di nodi e archi centrali all'interno delle reti in grado di gestire lo scambio di informazione. Questa caratteristica si verifica anche nei valori alti della *strength*, i link che persistono sono i più informativi e caratterizzanti delle network e nell'*entropia topologica* che comunica un disordine di informazione maggiore ma che tende a scemare fino al valore di soglia pari a 0.3, in associazione di una ricchezza informativa non eliminabile.

Per soglie alte, invece, la *betweenness centrality* mantiene una bassa sensibilità e un'alta variabilità, la *char path* continua ad essere fortemente sensibile, la *clustering coefficient* diventa maggiormente variabile mantenendo una sensibilità alta e la *local efficiency* aumenta la sua sensibilità continuando ad essere altamente variabile. Per tutte le altre metriche la situazione è speculare a quella per soglie basse. Il comportamento speculare è giustificato, d'altronde, dal continuo depauperamento delle connessioni cerebrali che supportano la presenza di un'alta sensibilità alla soglia.

Per i grafi non pesati le valutazioni sono simili, soltanto per la *global efficiency*, la *local efficiency* e la *clustering coefficient* si registrano dei cambiamenti in coincidenza delle soglie superiori a 0.3. La *global efficiency* marca un' spiccata variabilità, a differenza della *local efficiency* per cui la variabilità diminuisce. La *clustering coefficient*, invece, assume un trend speculare per soglie alte rispetto a quello per soglie basse, giustificando la disgregazione di moduli triangolari alle alte soglie.

Capitolo 5

Discussione

In questo capitolo sono discussi i risultati in modo da comprendere gli impatti delle varie scelte progettuali nei riguardi sia delle mappe FC di partenza sia della caratterizzazione delle metriche di grafo.

5.1 Variabilità intra-soggetto e inter-soggetto

La variabilità intra-soggetto ha evidenziato una forte presenza sia durante i test preliminari, sia nel primo tentativo di caratterizzazione e nell'analisi iniziale dello studio delle metriche di grafo. Il suo contributo è stato intuito grazie alla quantificazione delle caratteristiche relative alle distribuzioni delle matrici FC dei singoli individui, per poi essere verificata dal calcolo dell'entropia di ciascuna metrica di grafo a livello intra-soggetto. In maniera definitiva è stata accertata con l'implementazione della metrica *Node Degree* al variare della soglia. La variabilità intra-soggetto ha reso le valutazioni più complicate dal momento che l'obiettivo prefissato è stato di ricavare misure ripetibili per ciascun soggetto e tra più individui. Le cause da imputare alla variabilità sono molteplici. Possono influire artefatti di movimento durante l'esame diagnostico, difficoltà di riproduzione della tipologia di condizione cerebrale, ossia di *resting-state*, da parte del singolo soggetto durante la procedura ed altri ancora. In ogni caso sono difficilmente riconoscibili ed estrapolabili dai dati. Tuttavia non è possibile trascurare questa variabilità e semplificare il problema uniformando i dati dei soggetti. Infatti, il problema si è riscontrato per tutte le metriche di grafo e per tutti gli indici statistici calcolati.

La variabilità inter-soggetto, invece, ha permesso di monitorare i trend delle misure di grafo al variare della soglia e di poter ottenere informazioni riguardanti eventuali limiti per i valori di soglia. Tuttavia, tale variabilità non è credibile. Durante la caratterizzazione delle

distribuzioni individuali, sia per mappe FC non sogliate sia al variare della soglia, si sono distinti più gruppi di soggetti con range di misura simili. Nella caratterizzazione delle metriche di grafo mediante indice di entropia, i profili delle curve dell'indice entropico a livello di popolazione e di matrice hanno ripresentato il trend della curva a livello intra-soggetto media. L'unico profilo differente è stato riportato dall'indice entropico a livello inter-soggetto. In ogni caso dai grafici a livello intra-soggetto ed inter-soggetto si sono potute caratterizzare le metriche in termini di variabilità e di sensibilità a soglia, indicando anche l'impatto della scala numerica applicata alle mappe di connettività funzionale.

5.2 Impatto della scala numerica

Nel percorso sperimentale si è potuto osservare l'influenza di una scala numerica differente applicata ai valori di correlazione delle matrici di connettività funzionale dei soggetti analizzati. In primo luogo, la scala normale è stata scelta per poter agevolare la valutazione sulla caratterizzazione delle metriche di grafo. La normalizzazione, per definizione, rende uniformi e gaussiane le distribuzioni dei valori di correlazione, riportandole a media nulla e varianza unitaria. La gaussianità e l'uniformità dei valori di correlazione, nello specifico, sono fattori determinanti per la ripetibilità delle metriche di grafo. Tuttavia la scala normale ha influenzato non solo l'analisi della variabilità dell'informazione relativa alla connettività funzionale ma anche l'analisi di grafo. Durante la quantificazione della variabilità derivante dalle matrici di connettività, la condizione di uniformità è stata evidente non solo a livello intra-soggetto ma anche inter-soggetto. Tale omogeneità ha determinato l'annullamento della variabilità intra-soggetto e ha influenzato il contenuto informativo della prima componente di PCA. Ben il 30% dell'informazione è risultata non spiegata dalla prima componente sia per matrici di connettività non sogliate sia al variare della soglia. Per l'analisi di grafo, l'uniformità delle misure a tutti i livelli di indagine ha complicato le valutazioni delle metriche determinando risultati inattendibili e non affidabili a livelli fisiologico. Benché queste condizioni fossero auspicabili perché ben note e più semplici da gestire, la spiccata variabilità intra-soggetto ha ridotto la possibilità di semplificare il problema tramite la scala normale.

In secondo luogo, invece, per l'analisi di grafo le matrici di connettività sono state trasformate con z-Fisher. La trasformazione ha accentuato il carattere di gaussianità delle distribuzioni dei valori di correlazione rendendole leptocurtiche. Ha permesso di evidenziare

gli outlier delle distribuzioni, favorendo l'individuazione dell'informazione utile e ha reso più agevole la caratterizzazione delle metriche di grafo. L'applicazione, inoltre, di questa scala numerica non ha apportato sostanziali modifiche rispetto allo stato inalterato delle mappe FC sia a livello intra-soggetto sia inter-soggetto, in quanto non si sono osservati cambiamenti significativi per gli andamenti delle metriche sia come profilo sia come range di valori. Di conseguenza, tali osservazioni hanno marcato migliori condizioni di lavoro tramite la scala di Fisher a tutti i livelli di analisi, monitorando in maniera più puntuale i trend delle metriche e non alterando l'informazione in ingresso. La trasformata z-Fisher, quindi, si è rivelata più adeguata all'analisi dei grafi rispetto alla scala normale. In questo modo si conferma l'utilità della trasformata di Fisher per l'analisi di grafo, come mostrato in letteratura[9].

5.3 Impatto della soglia per grafi pesati e non

Per tutte le metriche è stato importante comprendere il loro cambiamento al variare della soglia applicata. In tutti gli stadi del lavoro è stata utilizzata una soglia assoluta con un passo di campionamento di un centesimo, al fine di indagare la variabilità topologica delle misure di grafo. Ogni metrica ha mostrato un comportamento diverso sia in base al valore della soglia sia in base alla tipologia di grafo, pesato o non, in quanto per definizione misurano aspetti differenti del grafo e hanno diverso significato. Inoltre, la modellizzazione del trend entropico a livello inter-soggetto, non riportato per esteso in questo elaborato, ha potuto individuare un punto di ginocchio pari a 0.3 oltre al quale le misure non sono in grado di descrivere le network. Tale valore si è rilevato effettivamente un punto critico.

Tra le metriche di centralità è stata analizzata la *betweenness centrality*. Per definizione il suo ruolo è di identificare i nodi centrali della rete che gestiscono i flussi informativi più consistenti. Dalle analisi tale metrica è risultata poco sensibile per soglie inferiori al punto critico e robusta. Le reti cerebrali, dunque, sono in grado di mantenere la gestione dei flussi informativi più onerosi.

Per le misure di integrazione sono state studiate la *global efficiency* e la *char path*. Solo la *global efficiency* si è rivelata una metrica affidabile per la valutazione delle network in termini di efficienza. La *char path* è risultata estremamente sensibile all'applicazione della soglia fin da valori perfino inferiori al punto critico. Tuttavia, queste due misure rilevano la

robustezza delle mappe FC per soglie basse in quanto sono in grado di conservare un'elevata efficienza.

Per le misure di segregazione, sono state monitorate la *clustering coefficient*, la *modularity* e la *local efficiency*. Sia la *clustering coefficient* sia la *local efficiency* si sono dimostrate metriche robuste a variazioni di soglia inferiori a 0.3, tranne la *modularity* che registra un'alta sensibilità per soglie più elevate. Per la valutazione, dunque, della tendenza alla formazioni di moduli e l'influenza sulla connettività della rete si consigliano le metriche: *clustering coefficient* e *local efficiency* per un'indagine più conservativa, la *modularity* per indagini meno conservative.

Come misura di elasticità, si è quantificata la *strength*. Per soglie basse le reti cerebrali appaiono resistenti ad attacchi esterni e mantengono le connessioni più forti e informative, data la bassa sensibilità della misura. In generale, quindi, le reti cerebrali dei soggetti per soglie basse appaiono efficienti, gestiscono una grande dose di informazione, scambiano dati a livello inter-modulare in maniera più efficace rispetto al livello intra-modulare e i legami tra le ROI sono molto resistenti e forti. Tutto questo accade fino ad un valore di soglia inferiore a 0.3, per cui le reti neurali sono depauperate da una quantità non eccessiva di connessioni cerebrali. Per soglie alte, invece, le connessioni maggiormente informative sono rimosse e le reti si destabilizzano. In definitiva, le metriche sopra citate mostrano una maggiore capacità di lavoro per intervalli di soglie basse in quanto poco sensibili e maggiormente robuste e, come già descritto nei risultati, si può identificare un limite superiore della soglia che assume il valore 0.3. Inoltre, per i grafi non pesati, la *strength* e la *betweenness centrality* non sono definite per essere implementate per matrici di adiacenza. Per tutte le altre metriche le considerazioni descritte per il caso pesato valgono anche per il non pesato.

In relazione a quanto è riportato in letteratura, si riconfermano le metriche di *betweenness centrality*, *clustering coefficient* e *local efficiency* come robuste ma per valori di soglia inferiori a 0.3. In aggiunta allo stato dell'arte, quindi, si sono identificati i limiti di robustezza e sensibilità delle metriche più usate e di altre ancora poco note. Inoltre, i range di misura delle metriche sono in linea con quelli riportati in letteratura e risultano credibili a livello fisiologico e topologico. All'aumentare della soglia la costruzione tridimensionale dei grafici, infatti, marca visivamente la rimozione di un'alta percentuale di connessioni, pari al 94% per la soglia 0.7, sintomo di instabilità delle network e, di conseguenza, delle instabilità delle metriche relative all'analisi dei grafi.

5.4 *Impatto dei valori negativi di correlazione*

Da letteratura tuttora non è chiaro il significato fisiologico delle correlazioni negative per le mappe FC né le modalità di analisi adeguate. In questo elaborato si è osservato che la loro esigua quantità, pari al 7%, ha reso complesso il loro studio. Durante la quantificazione della variabilità dell'informazione derivante dalle matrici FC, in loro assenza si sono registrati: un contenuto informativo maggiore per la prima componente principale, esito positivo per i test statistici e minore entropia a livello intra-soggetto. Per l'analisi dei grafi, invece, se isolati attraverso il principio della sovrapposizione degli effetti, non è stato possibile quantificare il valore delle metriche né caratterizzare le distribuzioni delle stesse a causa della natura sparsa delle matrici di connettività. In definitiva, si è riscontrata l'inattendibilità dei valori negativi di correlazione per l'analisi di grafo, nessun chiaro significato fisiologico e mappe FC più robuste in loro assenza. Si consiglia, dunque, di effettuare l'analisi basate su mappe FC per valori di correlazione positivi in quanto si è potuto concludere che i pesi negativi fossero trascurati per notevoli difficoltà nella comprensione del loro contributo nelle reti cerebrali.

5.5 *Impatto dei software*

A livello implementativo sono stati usati due software per l'analisi del grafo, come più volte ribadito nell'elaborato. Sebbene BCT sia il toolbox più utilizzato in letteratura, GRETNA ha permesso di indagare meglio la presenza delle correlazioni negative. Il software, infatti, contiene molte metriche già sviluppate per l'analisi di mappe FC con anticorrelazioni. Tuttavia, per lo scopo di questo studio e per le metriche analizzate, non si sono rilevate significative differenze di implementazione, sia in termini di definizione della metrica sia in termini di range di misura e di conseguenza del profilo delle metriche. L'unica sostanziale differenza è stata riportata nella definizione della *global efficiency* che non è semplicemente la media delle *global efficiency*, ma è l'inverso della stessa, fattore che marca il valore della *global efficiency* durante l'analisi del grafo. Tuttavia, i due software sono validi entrambi ma si consiglia GRETNA perché contiene una più ampia scelta di metriche.

Capitolo 6

Conclusione

Il lavoro di tesi ha avuto lo scopo di caratterizzare le metriche di grafo quali: *betweenness centrality*, *modularity*, *global efficiency*, *local efficiency*, *char path*, *strength*, *clustering coefficient* ed *entropia topologica* in soggetti di controllo, sani. Il percorso scelto ha riguardato un'analisi preliminare dell'informazione contenuta all'interno delle mappe FC dei singoli soggetti, un tentativo iniziale di caratterizzazione delle metriche tramite la valutazione dell'indice di entropia delle matrici FC contro quello delle misure di grafo e lo sviluppo del metodo di caratterizzazione attraverso i *survival plot*. Le metriche sono state valutate in termini di variabilità, sensibilità e range di misura al variare della soglia. Si è esaminato, in definitiva, l'impatto della scala numerica applicata alle mappe FC, l'impatto dell'utilizzo della soglia, le influenze della componente di variabilità intra-soggetto e l'impatto della presenza delle correlazioni negative all'interno delle mappe FC.

In prima istanza, quindi, è stato indagato il contenuto informativo delle mappe di connettività funzionale dei singoli soggetti, con varie ipotesi di lavoro, con lo scopo di comprendere se le metriche di grafo fossero in grado di apportare una più chiara valutazione delle network cerebrali. Si è svolta l'analisi di popolazione al fine di ricavare una metodologia per la descrizione univoca degli individui e si è studiato il significato fisiologico delle connessioni cerebrali con peso negativo, per identificare il loro ruolo all'interno delle reti neurali. Da questi test preliminari è risultato che il contenuto informativo delle mappe FC non fosse uguale ai vari livelli di analisi, che la prima componente della PCA fosse in grado di rappresentare il campione medio di popolazione e che per i link con peso negativo non fossero valutabili né attendibili gli andamenti delle misure implementate.

In seguito si è testato e modellato il metodo di caratterizzazione delle metriche di grafo, monitorando i trend dell'indice entropico delle singole metriche contro quello delle mappe

FC. Nello specifico, i grafici sono stati valutati a livello intra-soggetto, inter-soggetto, di popolazione e di ogni singola matrice in modo da avere una visione puntuale del comportamento delle singole misure, al variare della soglia. Purtroppo questo tentativo non ha portato a risultati soddisfacenti e non ha individuato un metodo ripetibile e affidabile per la ricerca della tipologia di soglia da applicare.

Infine, è stata esaminata la metodologia dei *survival plot* per la caratterizzazione delle misure di grafo sia a livello intra-soggetto sia inter-soggetto, verificando l'influenza della variabilità intra-soggetto e la robustezza delle singole misure al variare della soglia. Il confronto diretto tra il trend entropico delle mappe FC e delle misure di grafo si è dimostrato un ottimo metodo di valutazione per la catalogazione delle metriche, in quanto si è potuto identificare un limite superiore dei valori di soglia oltre al quale si è notata la completa destabilizzazione delle reti neurali e per la scelta della scala numerica più adeguata alle mappe FC di partenza, in grado di non alterare l'informazione in ingresso.

Inoltre, le mappe FC sono state valutate a livello topologico. E' stata implementata, dunque, la misura di entropia topologica e sono stati costruiti i grafici tridimensionali delle reti neurali per ogni singolo soggetto al variare della soglia. L'ispezione visiva dei grafi ha dato prova della corretta interpretazione dei risultati ottenuti con le singole metriche di grafo

In aggiunta, ad ogni stadio del lavoro di tesi sono stati testati due tipologie di software per l'implementazione delle misure, uno di più recente sviluppo, GRETNA e quello maggiormente utilizzato in letteratura, BCT, per i quali non si è riscontrata una sostanziale differenza. Dunque, sono validi entrambi i software, ma si consiglia il GRETNA per indagini più approfondite.

In particolare tutti gli stadi dell'indagine sono stati affrontati su un gruppo ristretto di soggetti per poi essere riconfermati da un numero superiore di individui. I risultati per le metriche di grafo ottenuti per i 10 soggetti si sono riproposti per i 25 e hanno permesso di effettuare la caratterizzazione delle metriche finali ritenendo più adeguate tali scelte progettuali: la trasformata z-Fisher come scala numerica, valori positivi di correlazione, distinzione per grafi non orientati pesati e non e punto critico di soglia pari a 0.3.

In conclusione, la caratterizzazione delle metriche di grafo tramite *survival plot* ha introdotto incoraggianti risultati per l'individuazione di un metodo utile per sogliare le mappe FC, comprendere l'impatto della scala numerica sulle stesse, l'impatto topologico della soglia e l'influenza dell'implementazione tramite software diversi.

In relazione ai dati analizzati in questo elaborato, dunque, la matrice di connettività ottima è costituita da valori di correlazione z-Fisher trasformati, solo positivi e superiori al valore

di soglia 0.3, **Figura 6.21** si riportano le immagini della matrice di connettività media e della deviazione standard.

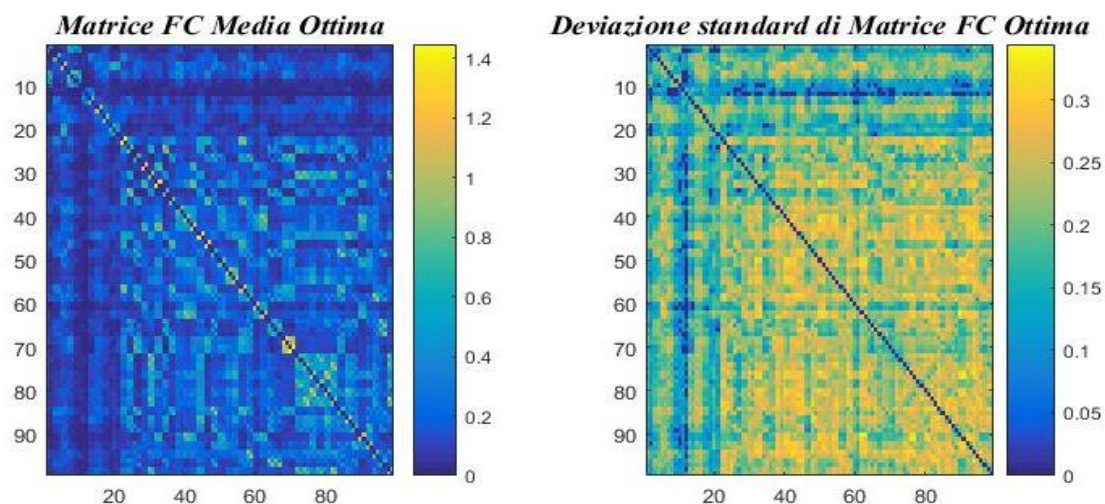


Figura 6.21 Immagini della matrice media ottima e deviazione standard

6.6 *Sviluppi futuri*

I risultati, finora ottenuti, si sono rilevati importanti per una futura distinzione tra soggetti sani e patologici attraverso un immediato riconoscimento del trend delle metriche e per uno studio delle patologie neurodegenerative a vari livelli di gravità. In futuro la ricerca si può sicuramente estendere a molti aspetti che rimangono ancora oscuri e ignoti.

Ad esempio, all'interno dell'elaborato è stata dimostrata una forte presenza della componente di variabilità intra-soggetto all'interno delle mappe FC. Questa caratteristica ha reso la scelta dei valori della soglia alquanto difficoltosa. Questo potrà essere ulteriormente approfondito al fine di identificare con facilità un intervallo di soglia da applicare e di comprendere in quali condizioni applicarlo. Inoltre, la condizione di variabilità intra-soggetto ha come diretta conseguenza l'incapacità di attribuire una soglia comune per tutti i soggetti, come si è notato durante lo svolgimento delle analisi. Risulta necessario, quindi, identificare metodi affidabili per l'attribuzione in modo da creare un modello per il calcolo della soglia per ogni soggetto sulla base dei dati ricavati dalle metriche di grafo.

Inoltre, una maggiore ricerca risiede nel confrontare i trend relativi ai soggetti sani con gli andamenti ricavati da un pattern di soggetti patologici, accomunati da una singola tipologia

di malattia neurodegenerativa e di caratterizzarne i cambiamenti più significativi rispetto ai soggetti di controllo. In questo modo si potrebbero individuare le caratteristiche chiave per l'identificazione di varie malattie, quali Parkinson, Alzheimer e predirne l'insorgenza in soggetti predisposti.

Ulteriori analisi possono riguardare i software implementativi, in quanto esistono varie interpretazioni delle metriche di grafo da applicare sulle reti cerebrali e non contengono tutte le misure di recente sviluppo. Un altro ambito di ricerca futura, infatti, può interessare la creazione e la caratterizzazione di ulteriori metriche di grafo da applicare alle reti neurali, come è avvenuto già con l'entropia topologica durante il lavoro di tesi.

L'obiettivo futuro di questa analisi è di rendere più immediato il riconoscimento di patologie neurodegenerative da parte del medico attraverso la valutazione dei trend delle metriche di grafo, di snellire l'analisi ingegneristica per ricavare le informazioni necessarie ai casi di studio e, di conseguenza, ricavare un modello ad hoc per il calcolo delle soglie che sia in grado di riconoscere le varie tipologie di malattie neurologiche e che possa essere facilmente aggiornato e arricchito in coincidenza agli upgrade degli studi di ricerca su nuove patologie e su nuove metriche di grafo più robuste, ripetibili e affidabili.

Appendice

Elenco degli acronimi

<i>ACRONIMO</i>	<i>SIGNIFICATO</i>
<i>fMRI</i>	Risonanza magnetica funzionale
<i>BOLD</i>	Blood Oxygenation Level Dependent
<i>ROI</i>	Region Of Interest
<i>AAL</i>	Automated Anatomical Labeling
<i>SPM</i>	Statistical Parametric Mapping
<i>FSL</i>	FMRIB Software Library
<i>DMN</i>	Default Mode Network
<i>FC</i>	Connettività Funzionale
<i>PCA</i>	Principal Component Analysis
<i>PON</i>	Positivo Originale Non pesato
<i>POP</i>	Positivo Originale Pesato
<i>PAF</i>	Positivo Adiacenza Fisher
<i>PAN</i>	Positivo Adiacenza Normale
<i>PPF</i>	Positivo Pesato Fisher
<i>PPN</i>	Positivo Pesato Normale
<i>BCT</i>	Brain Connectivity Toolbox
<i>GRETNA</i>	GRaph thEoretical Network Analysis
G^w	Weigthed Graph
G^{uw}	UnWeigthed Graph
$N G^{uw}$	Not UnWeigthed Graph
<i>LE</i>	Local Efficiency
<i>GE</i>	Global Efficiency

<i>STR</i>	Strength
<i>BC</i>	Betweenness Centrality
<i>CC</i>	Clustering Coefficient
<i>CP</i>	Char Path
<i>H</i>	Entropia
<i>MC</i>	Classic Modularity

Elenco delle figure

Figura 1 Resting-state network: N1_network_visivo_1; N2_network_visivo_2; N3_network_visivo_3; DMN_default_mode_network; N4_network_cerebellare; N5_network_sensomotorio; N6_network_uditivo; N7_network_controllo_esecutivo; N8.1_network_frontoparietale_sinistro; N8.2_network_frontoparietale_destro (modificata[3]).....	8
Figura 2.2 Diagramma di flusso relativo alla costruzione della network cerebrale: 1.Acquisizione immagini di R-fMRI, 2.Parcellizzazione anatomica per analisi basata su ROI, 3.Estrazione delle serie temporali di connettività funzionale tra le regioni cerebrali, 4.Generazione delle matrici di connettività, 5.Analisi del grafo tramite metriche descrittive [2]	10
Figura 3.3 Entropia Inter-Soggetto al variare della soglia: punto di ginocchio evidenziato nel riquadro significativo per la caratterizzazione delle metriche di grafo	25
Figura 3.4 Survival Plot: la curva di entropia e la curva della metrica di grafo sono confrontate nello stesso grafico	26
Figura 4.5 Distribuzioni individuali relativi ai 10 soggetti dei valori di connettività funzionale sia positivi sia negativi e normalizzati.....	32
Figura 4.6 Immagini della connettività funzionale media intra-soggetto e della deviazione standard relativi ai 10 soggetti.....	33
Figura 4.7 Variabilità individuale del numero di connessioni nulle al variare della soglia con passo di campionamento di un decimo	33
Figura 4.8 Profilo di skewness medio al variare della soglia delle distribuzioni per valori di correlazione sia positivi sia negativi.....	34
Figura 4.9 Varianza Spiegata per Matrici di Connettività senza l'applicazione di una soglia, nel caso di dati normalizzati e non	35
Figura4.10 Entropia Intra-Soggetto per matrici di connettività funzionale originali per distribuzioni normalizzate e non	36

Figura 4.11 Varianza spiegata per mappe di connettività funzionale con correlazioni positive per distribuzioni normalizzate e non.....	37
Figura 4.12 H: Entropia Intra-Soggetto per valori di correlazione positiva per distribuzioni normalizzate e non.....	38
Figura 4.13 Varianza Spiegata e Variazione del numero di connessioni rimosse per le matrici di connettività funzionale valutate al variare della soglia, con distribuzioni normalizzate e non.....	39
Figura 4.14 Entropia Intra-Soggetto, Inter-Soggetto e di Popolazione per matrici di connettività funzionale originali al variare della soglia assoluta.....	40
Figura 4.15 Entropia di metrica di grafo vs Entropia FC Inter-soggetto GRETNA. LE: Local Efficiency; H: entropia; CC: Clustering Coefficient; BC: Betweenness Centrality; GE: Global Efficiency; STR: strength	45
Figura 4.16 Survival Plot Intra-Soggetto per il caso PPF e 25 soggetti; LE:local efficiency, CP:char path, GE:global efficiency, MC: modularity classic, BC:betweenness centrality, STR:strength.....	48
Figura 4.17 Variabilità e sensibilità a soglia metriche di grafo inter-soggetto	51
Figura 4.18 Entropia Topologica Caso PPF e PAF	53
Figura 4.19 Grafo soggetto 1 per la rappresentazione della betweenness centrality con soglia applicata ai valori di correlazione pari a 0.3.....	54
Figura 4.20 Grafo soggetto 1 per la rappresentazione della betweenness centrality con soglia applicata ai valori di correlazione pari a 0.7	54
Figura 6.21 Immagini della matrice media ottima e deviazione standard.....	67

Elenco delle tabelle

Tabella 3.1 Soggetti di Controllo.....	20
Tabella 3.2 Sommario dei software per l'imaging connettomico: confronto tra GRETNA e BCT [1].....	30
Tabella 4.3 Indici di variabilità di grafo di matrici di connettività positive	42
Tabella 4.4 Indici di confronto media deviazione standard per i 10 soggetti al variare della soglia: Sog: soggetto; H: entropia; GE: global efficiency; LE: local efficiency; BC: betweenness centrality; CC: clustering coefficient; STR: strength	44
Tabella 4.5 Caratterizzazione metriche di grafo sulla base della variabilità e sensibilità a soglia; GUW: UnWeighed Graph; GW:Weighed Graph; N GUW:Not UnWeighed Graph; Bassa: [0, 0.3], Alta]0.3, 0.92]; Altissima]0.92, Inf[.....	56
Tabella 6 Labeling delle ROI con Atlante Cerebrale	75
Tabella 7 Labeling Cerveletto con Atlante Cerebrale	76

ATLANTE Resting-state Networks

<i>Nome Area Cerebrale</i>	<i>Label ROI Destra</i>	<i>Label ROI Sinistra</i>
Amygdala	31	32
Nucleo Caudato	36	37
Cervelletto Esterno	38	39
Ippocampo	47	48
Ventricolo Laterale Inferiore	49	50
Pallidum	55	56
Putamen	57	58
Thal Pr	59	60
Diencefalo Ventrale	61	62
ACgG Giro Cingolato Anteriore	100	101
AIns Insula Anteriore	102	103
AnG Giro Angolare	106	107
Calc Corteccia Calcarina	108	109
CO Operculum Centrale	112	113
Cun Cuneo	114	115
FO Operculum Frontale	118	119
FRP Polo Frontale	120	121
FuG Giro Fusiforme	122	123
IOG Giro occipital inferiore	128	129
LiG Giro Linguale	134	135
MCgG Giro Cingolato medio	138	139
MFC Corteccia mediale frontale	140	141
MFG Giro mediale frontale	142	143
MOG Giro occipital mediale	144	145
MPoG Giro postcentrale segment mediale	148	149
MPrG Giro precentrale mediale	150	151
MSFG Giro Frontale Superiore Mediale	152	153

MTG	Giro temporale medio	154	155
OCP	Polo occipitale	156	157
OFuG	Giro Occipitale Fusiforme	160	161
OpIFG	Parte Opercolare del giro frontale inferiore	162	163
OrIFG	Parte Orbitale del giro frontale inferiore	164	165
PCgG	Giro cingolato posteriore	166	167
PCu	Precuneo	168	169
PIns	Insula posteriore	172	173
PO	Operculum parietale	174	175
PoG	Giro Postcentrale	176	177
PP	Planum Polare	180	181
PrG	Giro Precentrale	182	183
PT	Planum Temporale	184	185
SFG	Giro Superiore Frontale	190	191
SMC	Corteccia Motoria Supplementare	192	193
SMG	Giro Sopramarginale	194	195
SOG	Giro superiore Occipitale	196	197
SPL	Lobulo Superiore Parietale	198	199
STG	Giro Superiore Temporale'	200	201
TrIFG	Parte Triangolare del giro inferior frontale	204	205
TTG	Giro Temporale Trasverso	206	207

Tabella 6 Labeling delle ROI con Atlante Cerebrale

<i>Nome Area Cerebrale</i>	<i>Label ROI</i>
Cerebellar Vermis Lob I-V	71
Cerebellar Vermis Lob VI- VII	72
Cerebellar Vermis Lob VIII- X	73

Tabella 7 Labeling Cerveletto con Atlante Cerebrale

Bibliografia

1. Wang, J., et al., *GRETNA: a graph theoretical network analysis toolbox for imaging connectomics*. Front Hum Neurosci, 2015. **9**: p. 386.
2. Gleiser, P.M. and V.I. Spoomaker, *Modelling hierarchical structure in functional brain networks*. Philos Trans A Math Phys Eng Sci, 2010. **368**(1933): p. 5633-44.
3. Bernadet Klaassens, J.v.G., Jeroen Van der Grond, *Functionale Brain Networks in Aging and Alzheimer's Disease*. 2016.
4. van den Heuvel, M.P. and H.E. Hulshoff Pol, *Exploring the brain network: a review on resting-state fMRI functional connectivity*. Eur Neuropsychopharmacol, 2010. **20**(8): p. 519-34.
5. Stam, C.J. and J.C. Reijneveld, *Graph theoretical analysis of complex networks in the brain*. Nonlinear Biomedical Physics, 2007. **1**: p. 3-3.
6. Buckner, R.L. and J.L. Vincent, *Unrest at rest: Default activity and spontaneous network correlations*. NeuroImage, 2007. **37**(4): p. 1091-1096.
7. van Eimeren, T., et al., *Dysfunction of the default mode network in Parkinson disease: a functional magnetic resonance imaging study*. Arch Neurol, 2009. **66**(7): p. 877-83.
8. Boccaletti, S., et al., *Complex networks: Structure and dynamics*. Physics Reports, 2006. **424**(4-5): p. 175-308.
9. Garrison, K.A., et al., *The (in)stability of functional brain network measures across thresholds*. Neuroimage, 2015. **118**: p. 651-61.
10. Rubinov, M. and O. Sporns, *Complex network measures of brain connectivity: uses and interpretations*. Neuroimage, 2010. **52**(3): p. 1059-69.
11. Poldrack, R.A., *Region of interest analysis for fMRI*. Soc Cogn Affect Neurosci, 2007. **2**(1): p. 67-70.
12. Wang, J., X. Zuo, and Y. He, *Graph-based network analysis of resting-state functional MRI*. Front Syst Neurosci, 2010. **4**: p. 16.
13. Damoiseaux, J.S., et al., *Consistent resting-state networks across healthy subjects*. Proc Natl Acad Sci U S A, 2006. **103**(37): p. 13848-53.
14. Fox, M.D. and M.E. Raichle, *Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging*. Nat Rev Neurosci, 2007. **8**(9): p. 700-711.

15. Gusnard, D.A. and M.E. Raichle, *Searching for a baseline: Functional imaging and the resting human brain*. Nat Rev Neurosci, 2001. **2**(10): p. 685-694.
16. Greicius, M.D., et al., *Functional connectivity in the resting brain: A network analysis of the default mode hypothesis*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(1): p. 253-258.
17. Bullmore, E. and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems*. Nat Rev Neurosci, 2009. **10**(3): p. 186-98.
18. Butts, C.T., *Revisiting the foundations of network analysis*. Science, 2009. **325**(5939): p. 414-6.
19. Honey, C.J., et al., *Predicting human resting-state functional connectivity from structural connectivity*. Proc Natl Acad Sci U S A, 2009. **106**(6): p. 2035-40.
20. Latora, V. and M. Marchiori, *Efficient behavior of small-world networks*. Phys Rev Lett, 2001. **87**(19): p. 198701.
21. Buscema, M., *Algoritmo di Pruning di un grafo*.
22. Akhondi-Asl, A., et al., *Whole Brain Group Network Analysis Using Network Bias and Variance Parameters*. Proc IEEE Int Symp Biomed Imaging, 2012. **2012**: p. 1511-1514.
23. Simpson, S.L., F.D. Bowman, and P.J. Laurienti, *Analyzing complex functional brain networks: Fusing statistics and network science to understand the brain*dagger*. Stat Surv, 2013. **7**: p. 1-36.
24. Pamplona, G.S., et al. *Analyzing the association between functional connectivity of the brain and intellectual performance*. Front Hum Neurosci 2015 [cited 9; 61]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25713528>.
25. Rubinov, M. and O. Sporns, *Weight-conserving characterization of complex functional brain networks*. Neuroimage, 2011. **56**(4): p. 2068-79.
26. Chang, C. and G.H. Glover, *Effects of model-based physiological noise correction on default mode network anti-correlations and correlations*. Neuroimage, 2009. **47**(4): p. 1448-59.
27. Chen, G., et al., *Negative functional connectivity and its dependence on the shortest path length of positive network in the resting-state human brain*. Brain Connect, 2011. **1**(3): p. 195-206.
28. Goelman, G., N. Gordon, and O. Bonne, *Maximizing negative correlations in resting-state functional connectivity MRI by time-lag*. PLoS One, 2014. **9**(11): p. e111554.
29. Wang, H. and P.A. Yushkevich, *Multi-atlas segmentation with joint label fusion and corrective learning-an open source implementation*. Front Neuroinform, 2013. **7**: p. 27.

30. Sladky, R., et al., *Slice-timing effects and their correction in functional MRI*. Neuroimage, 2011. **58**(2): p. 588-94.
31. Holland, D., J.M. Kuperman, and A.M. Dale, *Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo Planar Imaging*. Neuroimage, 2010. **50**(1): p. 175-83.
32. Power, J.D., et al., *Methods to detect, characterize, and remove motion artifact in resting state fMRI*. Neuroimage, 2014. **84**: p. 320-41.
33. Crinion, J., et al., *Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses*. Neuroimage, 2007. **37**(3): p. 866-75.
34. Newman, M.E., *Modularity and community structure in networks*. Proc Natl Acad Sci U S A, 2006. **103**(23): p. 8577-82.