



Istituto Virtuale Nazionale delle Malattie Cerebrovascolari

**PROTOCOLLO DI STUDIO**

**“Registro dati nazionale dell’Istituto Virtuale delle Malattie Cerebrovascolari”**

**(“National Database of the Virtual Institute of Cerebrovascular Diseases”)**

Protocollo BIG-MENTI V.1.0 del 11 Maggio 2023

## PROTOCOLLO DI STUDIO

**“Registro dati nazionale dell’Istituto Virtuale delle Malattie Cerebrovascolari”**

**(“National Database of the Virtual Institute of Cerebrovascular Diseases”)**

**Acronimo: BIG-MENTI**

**Sperimentatore principale (PI):** Prof. Paolo Calabresi, U.O.C. Neurologia (Direttore), Dipartimento di Scienze dell’Invecchiamento, Neurologiche, Ortopediche e della Testa-Collo, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Francesco Vito 1 – 00168 Roma. E-mail [paolo.calabresi@policlinicogemelli.it](mailto:paolo.calabresi@policlinicogemelli.it)

**Sub-Investigators:** Dott.ssa Irene Scala, Dott. Giovanni Frisullo, Dott. Mauro Monforte, U.O.C. Neurologia, Dipartimento di Scienze dell’Invecchiamento, Neurologiche, Ortopediche e della Testa-Collo, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Francesco Vito 1 – 00168 Roma

**Promotore:** Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Agostino Gemelli 8, 00168, Roma (Italia)

**Finanziamento:** No Profit cofinanziato

**Cofinanziatore:** Ministero della Salute, Viale Giorgio Ribotta 5, 00144, Roma (Italia)

**Data Collection:** Data Collection Research Core Facility, Gemelli Science and Technology Park – GSTeP, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Francesco Vito 1 – 00168 Roma (RM); e-mail: [datacollection@gemelligenerator.it](mailto:datacollection@gemelligenerator.it)

**Versione:** 1.0

**Data:** 11 Maggio 2023

## Background e razionale

L'Organizzazione Mondiale della Sanità (OMS) nel 1978 ha definito le malattie cerebrovascolari come “tutti quei disordini in cui c'è un'area del cervello transitoriamente o permanentemente colpita da ischemia o sanguinamento, o nei quali uno o più vasi sanguigni cerebrali sono primariamente coinvolti in un processo patologico, o una combinazione delle due condizioni di cui sopra” e in cui “il processo patologico include anomalie della parete vasale, stenosi vasale od occlusione vasale da parte di un trombo o di un embolo, alterata permeabilità vasale al plasma ed ad altre cellule ematiche” [1].

Secondo uno special report del 1990 del National Institute of Neurological Disorders and Stroke (NINDS) le malattie cerebrovascolari sono state classificate in diverse categorie dal punto di vista clinico e patologico e tale classificazione comprende una vasta gamma di malattie, tra cui lo stroke ischemico ed emorragico, la trombosi venosa cerebrale, le malformazioni artero-venose cerebrali, gli aneurismi dei vasi cerebrali, le stenosi del circolo carotideo e vertebro-basilare e la patologia dei piccoli vasi intracranici [2].

Nel loro insieme le patologie cerebrovascolari costituiscono un'importante causa di disabilità e mortalità a livello globale. Considerando nello specifico lo stroke, la malattia cerebrovascolare di gran lunga più frequente, è stato stimato che nel 2017 questo costituisca in Europa la seconda causa di mortalità e di disabilità [3], con un'incidenza di circa 135.2/100.000 individui nel 2019 [4]. Di conseguenza, tale patologia comporta importanti costi socio-sanitari che, nel 2015, si stimavano essere di circa 45 miliardi di euro annui a livello europeo<sup>3</sup>. Secondo il Global Burden of Disease la prevalenza dei differenti sottotipi

di stroke è così suddivisa: circa il 28% è di tipo emorragico, il 62% circa di tipo ischemico e poco meno del 10% è rappresentato dall'emorragia subaracnoidea [5]. In particolare, andando ad indagare l'incidenza dello stesso in Italia, secondo i dati riportati dal ministero della Salute nel 2019 sono stati registrati 86.360 ricoveri per ictus, mentre i dati Istat indicano che nel 2018 le malattie cerebrovascolari nel loro complesso costituivano la seconda causa di morte, dopo le malattie ischemiche del cuore, con 55.434 decessi (l'8,8% di tutti i decessi), di cui 22.062 maschi (7,3%) e 33.372 femmine (10,1%) [6]. In tale contesto, tuttavia, va considerata l'ampia variabilità interregionale dell'incidenza dell'ictus nel nostro paese, che comprende regioni a bassa ed alta prevalenza [7]. Infine, anche patologie cerebrovascolari con un'incidenza minore, quali le malformazioni artero-venose cerebrali e la trombosi venosa cerebrale, assumono una netta importanza quando valutate nell'ottica delle possibili complicanze [8], e dell'età giovanile d'insorgenza [9].

Considerando l'elevata incidenza di tali patologie appare quindi ovvio come, per poter dimostrare nessi causali e per ricercare fattori prognostici con un elevato livello di evidenza, sia necessaria una raccolta sistematica di dati su larga scala che coinvolgano differenti strutture sparse in tutto l'ambito nazionale. Tale raccolta sistematica consentirebbe sia, come appena menzionato, di ottenere dati validi su patologie ad elevata prevalenza, quali lo stroke, sia di ottenere evidenze sufficienti per l'inquadramento diagnostico e prognostico di quelle malattie cerebrovascolari ad incidenza minore.

Inoltre, considerando l'elevata mole di dati, appare evidente come sia necessario un sistema di raccolta di informazioni rapido ed efficiente che riesca a garantire un risultato ottimale senza gravare eccessivamente sull'attività del personale sanitario.

In tale contesto nel 2017 il Ministero della Salute ha istituito la rete IRCCS delle Neuroscienze e della Neuroriabilitazione (RIN), il più grande network di ricerca italiana d'ambito, fondata per stimolare la collaborazione tra gli Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS), favorire la diffusione delle informazioni sulla attività clinico-scientifica e coordinare azioni di rilevanza internazionale volte ad aumentare il rilievo e la competitività del settore. All'interno di questa cornice si sviluppano in maniera trasversale numerosi Istituti Virtuali, specializzati per ambiti di patologia, che hanno lo scopo di fornire alla Rete una propria forte identità, armonizzare le attività degli IRCCS, razionalizzare gli investimenti e le risorse, costruire vaste coorti ed interagire con i network internazionali.

A questo scopo è stata inviata a tutti i membri della Rete nel mese di ottobre 2020 una survey per indicare, nell'ambito specifico delle Malattie Cerebrovascolari:

- Le Risorse umane, infrastrutturali e tecnologiche disponibili;
- Il volume e la tipologia di attività clinica specifica svolta nell'anno 2019.

Alla survey iniziale hanno risposto 19 centri (4 specificatamente dedicati alla riabilitazione neuromotoria, 2 a indirizzo pediatrico e 1 a specifico indirizzo di ricerca farmacologica), manifestando l'interesse alla partecipazione del costituendo Istituto Virtuale Nazionale

(IVN) delle Malattie Cerebrovascolari. In seguito, ulteriori due centri si sono aggiunti, arrivando all'attuale composizione di 14 centri al Nord, 5 al Centro e 2 al Sud.

Complessivamente, l'IVN delle Malattie Cerebrovascolari (IVNMC) coinvolge circa 10 Stroke Unit, 1400 operatori sanitari, 200 posti letto per acuti e 790 posti letto per riabilitazione, con un volume di attività di 7000 ischemie cerebrali, 2000 emorragie cerebrali e circa 3000 trattamenti e procedure per anno. Durante il kick-off meeting si è insediato il consiglio direttivo composto dai Referenti dei singoli IRCCS ed è stato nominato il Prof. Paolo Calabresi come coordinatore e la Prof. Michela Matteoli come vice-coordinatore.

Peculiarità di questo IVN è l'elevato volume di pazienti affetti da patologie così frequenti, la grande specializzazione clinica, l'alto livello tecnologico dell'assistenza in tutte le fasi della malattia (dall'acuzie fino alla riabilitazione), la dinamicità e tempestività di intervento richiesta e necessaria. I singoli componenti di questo IVN posseggono un grande e diversificato expertise sia nell'attività assistenziale (gestione dei trattamenti della fase acuta, degenza in terapia sub-intensiva, riabilitazione neuromotoria, peculiarità delle malattie cerebrovascolari nella popolazione pediatrica) che di ricerca clinica e pre-clinica (modelli sperimentali, biomarcatori). Inoltre, considerando l'elevato carico sociosanitario dell'insieme di tali patologie, in gran parte legato alla disabilità permanente che comportano, l'IVN delle Malattie Cerebrovascolari avrà la peculiarità di poter seguire tutto il percorso del Paziente dalle fasi acute (ospedaliere) a quelle di riabilitazione (in centri di neuroriabilitazione) sino al rientro a domicilio.

Attraverso tale studio si mira alla creazione di un registro nazionale per queste patologie.

## OBIETTIVI DELLO STUDIO

### *Obiettivo primario*

Obiettivo primario dello studio è l'identificazione e definizione dei vari sottotipi di ciascuna patologia cerebrovascolare, rientrante nei seguenti codici dell'International Classification of Diseases, 9th revision (ICD-9): ICD-9 430-438, associati a prognosi sfavorevole.

### *Obiettivi secondari*

- Identificazione delle variabili cliniche di ciascuna patologia cerebrovascolare correlate con la prognosi della stessa;
- Identificazione di biomarcatori di ciascuna patologia cerebrovascolare correlati con la prognosi della stessa;
- Identificazione delle variabili radiologiche di ciascuna patologia cerebrovascolare correlate con la prognosi della stessa;
- Identificazione delle variabili neurofisiologiche di ciascuna patologia cerebrovascolare correlate con la prognosi della stessa;
- Identificazione di score prognostici per ciascuna malattia cerebrovascolare, basati sulla combinazione di variabili cliniche, radiologiche, neurofisiologiche e di biomarcatori;

- Identificazione delle variabili cliniche, radiologiche e neurofisiologiche e dei biomarcatori di ciascuna patologia cerebrovascolare che aiutino il clinico nella diagnosi differenziale precoce;
- Identificazioni delle variabili cliniche, radiologiche e neurofisiologiche e dei biomarcatori di ciascuna patologia cerebrovascolare in grado di predire lo sviluppo di complicanze spontanee o legate alle procedure terapeutiche;
- Identificazioni delle variabili cliniche, radiologiche e neurofisiologiche e dei biomarcatori di ciascuna patologia cerebrovascolare in grado di predire il buon esito delle terapie farmacologiche o procedurali;
- Identificazioni delle variabili cliniche, radiologiche e neurofisiologiche e dei biomarcatori di ciascuna patologia cerebrovascolare utile ai fini della personalizzazione del trattamento riabilitativo ed in grado di predire il buon esito funzionale a seguito degli stessi trattamenti riabilitativi.

## MATERIALI E METODI

### *Disegno dello studio*

Studio multicentrico nazionale, osservazionale, ambispettico su larga scala. Lo studio è co-finanziato dal ministero della Salute nell'ambito della Rete Italiana Neuroscienze (rete RIN).

### *Popolazione*

I partecipanti verranno reclutati in ogni singolo centro dal principal investigator (PI) o da uno dei sub-investigators (sub-I) locali in occasione di una fase qualsiasi della gestione delle malattie cerebrovascolari: il ricovero in fase acuta, la riabilitazione neuromotoria successiva in regime residenziale o di Day Hospital, o la visita ambulatoriale di follow-up a tre mesi dall'evento acuto presso la struttura per acuti o di neuroriabilitazione o eventuali visite di follow-up successive, effettuate entro tre anni dall'esordio della patologia cerebrovascolare. I pazienti verranno reclutati retrospettivamente dal 1° gennaio 2016 e si conta un arruolamento di circa 400 pazienti annui per centro, per un numero totale di pazienti di circa 285.600 pazienti.

### *Centri partecipanti*

21 centri su rete nazionale. La distribuzione geografica dei centri partecipanti è così composta: 14 al Nord, 5 al Centro e 2 al Sud. La lista dei centri partecipanti allo studio con i rispettivi PI è disponibile nell'**allegato 1**.

## *Durata dello studio*

La durata prevista dello studio è di 360 mesi a partire dalla data di approvazione da parte del comitato etico della Fondazione Policlinico Universitario Agostino Gemelli IRCCS, che sarà il Central Ethics Committee (CEC) del progetto stesso. In seguito all'approvazione da parte del CEC, la medesima documentazione verrà sottomessa a tutti gli altri comitati etici dei centri partecipanti al progetto. I primi 312 mesi dello studio saranno deputati all'arruolamento dei pazienti, i successivi 36 all'eventuale follow-up dei pazienti ed gli ultimi 12 all'analisi statistica (6 mesi) e all'eventuale stesura di manoscritti per la pubblicazione (6 mesi).

### *Criteri di inclusione*

- Diagnosi di patologia cerebrovascolare definita secondo lo special report del NINDS del 1990<sup>2</sup> come “tutti i disturbi in cui è presente un'area del cervello transitoriamente o permanentemente affetti da ischemia o sanguinamento e/o in cui uno o più vasi sanguigni di il cervello sono principalmente alterati da un processo patologico”, rientrante nei seguenti codici dell' International Classification of Diseases, 9th revision (ICD-9): ICD-9 430-438;
- Consenso informato fornito dal paziente o dal suo rappresentante legale nel caso di incapacità di intendere e di volere o dai genitori/tutori nel caso di soggetti di minore età.

### *Criteri di esclusione*

- Il paziente, il suo rappresentante legale nel caso di incapacità di intendere e di volere o i suoi genitori/tutori nel caso di soggetti di minore età esplicitano in maniera inconfutabile la volontà di uscire dallo studio;
- Insorgenza della malattia cerebrovascolare prima del 1 gennaio 2016.

### *Variabili e procedure*

I partecipanti verranno reclutati in ogni singolo centro dal *principal investigator* (PI) o da uno dei *sub-investigators* (sub-I) locali in occasione di una fase qualsiasi della gestione delle malattie cerebrovascolari: il ricovero in fase acuta, la riabilitazione neuromotoria successiva, o la visita ambulatoriale di follow-up a tre mesi dall'evento acuto. Ad ogni paziente verrà assegnato un codice alfanumerico univoco, che sarà caratteristico per ogni centro partecipante. La lista dei codici alfanumerici verrà custodita indipendentemente da ogni centro. Ai partecipanti dello studio non verrà richiesto alcun pagamento in denaro o di altra natura per i costi medici o di altro genere che incorreranno durante il periodo della sperimentazione. I pazienti inclusi verranno sottoposti a tutte le procedure effettuate di routine ai pazienti con ictus ischemico ed emorragico, includendo le visite in regime di ricovero ed ambulatoriale, esami ematici, ecocolordoppler dei vasi epiaortici, TC o risonanza magnetica dell'encefalo e qualunque procedura diagnostica/terapeutica prevista in maniera routinaria nei pazienti con patologie cerebrovascolare secondo la buona pratica clinica e basandosi sulle conoscenze attuali allo stato dell'arte: per tale ragione il costo di queste procedure mediche è a carico del sistema sanitario nazionale.

Tutti i dati clinici, i biomarcatori, i dati radiologici, neurofisiologici e di neuroriabilitazione verranno raccolti dai centri partecipanti durante ogni fase del processo diagnostico-terapeutico, secondo le esigenze, l'expertise e le buone norme di pratica clinica che si confanno allo specifico setting assistenziale (ricovero in fase acuta, fase di

neuroriabilitazione in regime di ricovero, *Day Hospital* o teleneuroriabilitazione, e visita di controllo ambulatoriale a 3 mesi dall'evento acuto ed eventuali successive visite di follow-up a lungo termine). Di seguito viene riportata la lista di dati che verranno presi in considerazione per la compilazione delle Case Report Form elettroniche (*electronic case report form – e-CRF*):

- Dati clinici, laboratoristici e strumentali – Fare riferimento ad **allegato 2**
- Dati di neuroriabilitazione – Fare riferimento ad **allegato 3**

I dati verranno collezionati in forma di e-CRF che verranno completate dai clinici durante la valutazione al baseline di ciascun differente setting assistenziale (ricovero in fase acuta, fase di neuroriabilitazione in regime di ricovero, *Day Hospital* o teleneuroriabilitazione, e visita di controllo ambulatoriale a 3 mesi dall'evento acuto).

## ENDPOINTS

### *Endpoint primario*

Endpoint primario dello studio è la prevalenza dei diversi sottotipi di ciascuna patologia cerebrovascolare, rientrante nei codici ICD-9 430-438, associati a prognosi sfavorevole.

### *Endpoints secondari*

- Valutazione di potenziali predittori di prognosi tra le variabili cliniche, i biomarcatori, le variabili radiologiche e neurofisiologiche di ciascuna patologia cerebrovascolare;
- Ipotesi di score prognostici per ciascuna malattia cerebrovascolare, basati sulla combinazione di variabili cliniche, radiologiche, neurofisiologiche e di biomarcatori emersi come predittori di patologia (v. endpoint precedente);
- Valutazione di potenziali predittori tra variabili cliniche, radiologiche e neurofisiologiche e dei biomarcatori di ciascuna patologia cerebrovascolare per lo sviluppo di complicanze spontanee o legate alle procedure terapeutiche;
- Valutazione di potenziali predittori di outcome terapeutico/procedurale favorevole tra variabili cliniche, radiologiche e neurofisiologiche e dei biomarcatori di ciascuna patologia cerebrovascolare.

## PIANO DI ANALISI STATISTICA

### *Dimensionamento campionario*

Trattandosi di uno studio di popolazione sull'incidenza dell'ictus in Italia, i.e. su larga scala, il cui endpoint primario è puramente descrittivo, non si rende necessario alcun dimensionamento formale del campione.

### *Analisi statistica*

Il campione verrà descritto nelle sue caratteristiche cliniche e demografiche mediante gli appropriati indici di statistica descrittiva. In dettaglio, i dati qualitativi saranno espressi come frequenza assoluta e relativa percentuale, mentre le variabili quantitative come media e deviazione standard (DS) o mediana e range interquartile (IQR), a seconda dei casi. Per verificare la distribuzione gaussiana delle variabili quantitative verrà applicato il test di Shapiro-Wilk. In caso di valori mancanti, questi verranno trattati mediante il pacchetto R *imputeR*, utilizzando l'imputazione multipla con metodi di Regressione Lasso centrati sulla media per i dati quantitativi, mentre alberi di classificazione per l'imputazione attraverso la funzione "rpartC", centrati sulla moda, ovvero la più rappresentata oggetto classe, sarà applicato su dati qualitativi [10].

Le differenze nei parametri clinici, biomarcatori, variabili radiologiche e neurofisiologiche tra le diverse patologie cerebrovascolari saranno valutate mediante test del Chi-quadrato o

test esatto di Fisher-Freeman-Halton per le variabili qualitative, mentre l'ANOVA ad una via o il test di Kruskal-Wallis saranno applicati sui dati quantitativi. I confronti a coppie, in quest'ultimo caso, verranno valutati mediante test t di Student o test U di Mann-Whitney, come appropriato. Le differenze nei parametri clinici, biomarcatori, variabili radiologiche e neurofisiologiche tra le diverse patologie cerebrovascolari saranno ulteriormente rappresentate da "violin plots", utilizzando i pacchetti R "ggpubr", "ggstatsplot", "ggplot2", "ggprism" e "ggsignif" [11-15]. Sarà utilizzata la correzione di Benjamini-Hochberg per confronti multipli.

La prognosi sarà definita in termini di sopravvivenza globale (OS) definita come il tempo che intercorre tra la diagnosi e la morte/ultimo follow-up del paziente. Le differenze in termini di prognosi tra le diverse patologie cerebrovascolari saranno valutate mediante l'analisi di Kaplan Meier utilizzando il pacchetto R "survival" [16,17]. La curva di Kaplan Meier verrà ulteriormente illustrata utilizzando i pacchetti R "ggplot2" e "survminer" [13,18].

Al fine di valutare potenziali predittori di prognosi saranno fittati modelli di regressione di Cox uni- e multivariabile a rischi proporzionali e riportati gli Hazard Ratio (HR) e gli intervalli di confidenza (CI) al 95%. La proporzionalità degli hazard sarà valutata mediante un'ispezione visiva degli hazard e mediante i residui di Schoenfeld. In caso di proporzionalità dubbia, verranno fittati modelli di regressione di Cox pesati [19,20].

Verrà inoltre valutata la possibilità di creare score prognostici per ciascuna malattia cerebrovascolare attraverso la creazione di due coorti, derivazione (per sviluppare il

modello di punteggio) e validazione (per convalidare e valutare le capacità diagnostiche del punteggio), in modo casuale. Sulla coorte di derivazione verranno eseguiti modelli di regressione di Cox univariata e multivariata per identificare fattori prognostici indipendenti da includere nel sistema di punteggio. In dettaglio, saranno calcolati Hazard Ratios (HR) e intervalli di confidenza al 95%. Le variabili incluse nel modello multivariabile saranno selezionate sulla base dei risultati dell'analisi univariata ( $p < 0,05$  o suggestivi) e dell'opinione degli esperti [21]. Le prestazioni del modello verranno valutate sulla base di diversi metodi, come la correlazione del rango Dxy di Somers, l'indice C, il valore Nagelkerke R<sup>2</sup>, l'intercetta e la pendenza della calibrazione e il punteggio di Brier [21]. Infine, il test di bontà di adattamento di Hosmer-Lemeshow consentirà la calibrazione nella coorte di derivazione [22]. I grafici di calibrazione forniranno inoltre una rappresentazione grafica dell'associazione tra i risultati previsti e quelli osservati mediante il livellamento del grafico a dispersione ponderato localmente. La convalida interna del modello verrà eseguita sulla base di una procedura bootstrap [22]. Successivamente sarà assegnato un punteggio ad ogni predittore incluso nel sistema di punteggio basato sull'HR di ciascuna variabile nella coorte di derivazione [22]. Verranno fissati valori di cut-off appropriati per un approccio rule-in e rule-out per aiutare nel processo decisionale. I predittori quantitativi indipendenti saranno stati ulteriormente trasformati in variabili qualitative ordinali o nominali, attraverso metodi di selezione appropriati basati sull'analisi delle curve ROC, mediante i pacchetti R "*pROC*" [23], ed "*OptimalCutpoints*" applicando il metodo di selezione SpEqualSe, che restituisce la

massima accuratezza [24]. Verranno quindi calcolate le capacità diagnostiche (cioè sensibilità, specificità, rapporto di verosimiglianza positivo e rapporto di verosimiglianza negativo) di ciascun punteggio. Per la convalida esterna, il punteggio sviluppato verrà applicato alla coorte di validazione, descrivendo le prestazioni di discriminazione e calibrazione, come sopra. La performance complessiva verrà descritta in termini di sensibilità, specificità, accuratezza (o valore predittivo positivo, PPV), punteggio F1, accuratezza, tasso di falsi positivi (FPR), tasso di false scoperte (FDR) e tasso di falsi negativi (FNR).

Per identificare predittori di ciascuna patologia cerebrovascolare per lo sviluppo di complicanze spontanee o legate alle procedure terapeutiche verranno invece utilizzati modelli di regressione logistica uni- e multi-variabile mediante il pacchetto R “*rms*”, che verrà altresì utilizzato per identificare potenziali predittori del buon esito delle terapie farmacologiche/procedurali [25].

La significatività statistica sarà fissata per valori di  $p < 0,05$  saranno considerati statisticamente significativi. Saranno riportati anche valori di  $p$  suggestivi ( $0,05 \leq p < 0,10$ ). Tutte le analisi saranno condotte con il software R, versione 4.2.0 (CRAN®, R Core 2022, Vienna, Austria) [26]. Verrà prevista un’analisi statistica al termine del progetto di ricerca, e multiple analisi ad interim che si svolgeranno ogni due anni in seguito all’approvazione dello studio da parte del CEC.

## GESTIONE DEI DATI E DELLA QUALITA'

### *Raccolta e gestione dei dati e dei registri*

I dati clinici verranno raccolti nel sistema di acquisizione elettronica dei dati (electronic data capture - EDC), REDCap [27,28]. Tale sistema si trova su di un server che sarà gestito dalla facility di data collection del "Gemelli Science and Technology Park" (GSTeP).

I compiti specifici di GSTeP sono:

1. Regolamentare le attività regolatorie dell'infrastruttura di ricerca;
2. Sviluppo delle e-CRF per la raccolta dei dati;
3. Training da remoto o in presenza del personale per la raccolta dei dati nelle e-CRF;
4. Controllo di qualità di primo livello dei dati raccolti;
5. Redazione di rapporti periodici sull'arruolamento dei pazienti e la raccolta dei dati;
6. Aggiornamento periodico del sistema EDC.

Lo staff coinvolto nel progetto riuscirà ad accedere al sistema EDC da un qualsiasi computer presente nei centri partecipanti per visualizzare le informazioni riguardo i soggetti arruolati dal proprio centro, completare le e-CRF e caricare i risultati riguardo le varie indagini mediche direttamente nel database stesso. Il database elettronico verrà istituito solo in seguito l'approvazione del progetto da parte del comitato etico del centro coordinatore.

**Sicurezza dei dati:** tutti i dati raccolti in relazione al progetto di ricerca devono obbligatoriamente essere sottoposti ad un monitoraggio continuo per garantire la

riservatezza dei dati sensibili. Tale aspetto assume particolare rilevanza nei periodi durante i quali i dati sono in transito tra un sistema e l'altro. Tutti i dispositivi ed i sistemi di raccolta e conservazione dei dati devono essere protetti da una password con un alto livello di sicurezza. L'accesso ad i dati identificabili di un partecipante allo studio sarà limitato esclusivamente al personale del centro che ha reclutato quel paziente.

### **Gestione delle informazioni personali**

Gli elementi identificativi ed i dati dovranno essere raccolti in file separati, protetti da password o criptati ed ogni file dovrebbe essere salvato separatamente dagli altri.

## CONSIDERAZIONI ETICHE

### *Privacy e sicurezza dei partecipanti allo studio*

La tutela delle informazioni personali dei soggetti partecipanti allo studio avverrà ai sensi del decreto legge n° 196 del 30/06/2003 del codice in materia di protezione dei dati personali.

In linea con la normativa internazionale riguardo la protezione dei dati, verranno adottate le seguenti misure:

- a) l'accesso al database contenente i dati verrà ristretto alle sole persone autorizzate;
- b) i questionari e tutti i dati cartacei verranno immagazzinati sottochiave;
- c) tutti i dati identificativi verranno criptati all'interno del database ed i soggetti saranno identificati solo con un codice alfanumerico.

In particolare, per poter identificare in maniera univoca un paziente all'interno della rete, verrà definito un codice identificativo (ID) del paziente, derivabile a partire dai dati del soggetto stesso. Il codice identificativo non sarà scritto in chiaro all'interno del REDCap, per evitare che si possa risalire all'identificazione del paziente.

Per la generazione di questo ID si è optato per utilizzare la concatenazione tra codice fiscale, data di nascita (in formato GGMMAAAA), e sesso (in formato "F"/"M") del paziente. La concatenazione di questi 25 caratteri viene poi trasformata tramite algoritmo SHA-256 in un codice alfanumerico che rappresenterà d'ora in avanti il codice identificativo del paziente.

SHA-256 è una funzione di crittografia e, la cosa più importante per l'IVNMC, è che non esiste un modo diretto per decodificarlo.

Tutti gli altri dati presenti nel CRF saranno scritti all'interno del database REDCap in chiaro.

A ciascun soggetto sarà inoltre garantito l'anonimato durante le presentazioni dei dati dello studio nel corso di eventi scientifici o nella pubblicazione degli stessi in giornali scientifici.

### *Terminazione precoce dello studio*

Il coordinatore del progetto o qualunque altra autorità competente possono stabilire di terminare prematuramente il progetto nel caso di problematiche di tipo finanziario.

### *Emendamenti/modifiche del protocollo*

Emendamenti e modifiche significative del protocollo dello studio dovranno essere approvate dal comitato etico di ciascun centro partecipante.

Il coordinatore dello studio dovrà sottomettere al comitato etico ogni variazione proposta e dovrà provvedere informazioni sufficienti riguardo le ragioni che hanno condotto alla proposta di tali variazioni. Emendamenti sostanziali verranno implementati solo in seguito alla loro approvazione da parte del comitato etico.

## POLITICA DI PUBBLICAZIONE E DISSEMINAZIONE DEI DATI

### *Pubblicazione dei risultati*

Le pubblicazioni scientifiche derivanti dal presente studio riguarderanno i PIs dei centri partecipanti all'IVNMC ed il personale di ogni centro che interverrà nella fase di raccolta ed analisi dei dati; verranno seguite le direttive per le attività di collaborazione scientifica contenute nel Regolamento RIN n.1 del 5 luglio 2018 (**allegato 4**).

## GOVERNANCE DEL PROGETTO

### *Coordinazione del progetto*

Il progetto è coordinato dal Prof. Paolo Calabresi, con sede presso la Fondazione Policlinico Universitario A. Gemelli IRCCS in Roma (RM), Largo Agostino Gemelli n° 8, 00168, Roma (RM). Il ruolo del coordinatore dello studio, essendo questo la costituzione di un registro nazionale di dati, consiste nel verificare che un supporto adeguato alla corretta esecuzione dello studio sia fornito agli investigatori dei diversi centri partecipanti. La progettazione del disegno dello studio, la raccolta, la gestione, l'analisi e l'interpretazione dei dati, la stesura di eventuali lavori scientifici, presentazioni ad eventi scientifici sono di responsabilità di tutti gli investigatori dell'IVNMC partecipanti al progetto.

## *Sperimentatori principali*

Ogni centro partecipante è guidato da un PI, che è responsabile di tutte le decisioni mediche di pertinenza del sito stesso. I PIs possono eventualmente delegare le decisioni mediche ai *Sub-investigators* (SUB-I) afferenti allo stesso dipartimento o ad altri dipartimenti della stessa istituzione o ad istituzioni affiliate al centro principale dello studio.

## *Condotta etica dello studio*

Lo studio verrà effettuato in accordanza con il protocollo e con i principi della Dichiarazione di Helsinki, le norme di *Good Clinical Practice* (GCP) istituite dalla *International Conference on Harmonisation* (ICH), dalla legge italiana e dai requisiti della normativa italiana.

## *Dichiarazione di interessi*

Tutti i partecipanti al progetto non riportano alcun tipo di conflitto di interessi.

### *Informazione dei partecipanti e consenso informato*

Gli investigatori dello studio spiegheranno in maniera estensiva e comprensibile a ciascun partecipante la natura dello studio, i suoi scopi, le procedure necessarie allo svolgimento dello studio stesso, la durata prevista dello studio, i potenziali rischi e benefici ed ogni disagio che la partecipazione possa comportare. Ogni partecipante sarà informato del fatto che la partecipazione allo studio è volontaria e che egli/ella potrà ritirarsi dallo studio appena lo desidera e che il suo ritiro dallo studio non avrà ripercussioni sulla sua successiva assistenza medica e sui trattamenti che gli/le verranno proposti. I partecipanti verranno altresì informati del fatto che le loro cartelle cliniche potrebbero essere esaminate da altri individui autorizzati oltre che dai loro medici curanti.

A tutti i soggetti candidati alla partecipazione allo studio verranno forniti un foglio informativo ed un modulo di consenso informato descrittivi in maniera estensiva il progetto, inclusivi di tutte le informazioni necessarie affinché il soggetto possa effettuare una decisione ponderata ed informata riguardo la volontà di partecipare allo studio. Inoltre, al paziente verrà chiaramente spiegato che il suo rifiuto di partecipare allo studio non comporterà alcuna conseguenza e che continuerà ad essere sottoposto alle migliori cure disponibili secondo le attuali norme di buona pratica clinica. Il consenso scritto alla partecipazione allo studio sarà ottenuto prima che venga raccolto alcun dato dalle cartelle cliniche del soggetto per la costituzione del registro. Il soggetto leggerà e valuterà le informazioni contenute all'interno della documentazione prima di firmare e datare il

modulo di consenso informato e gli verrà fornita una copia del documento firmato e datato.

Il consenso informato sarà inoltre firmato e datato dallo sperimentatore che proporrà al paziente la partecipazione allo studio e tale modulo sarà conservato tra la documentazione dello studio. Il modulo di consenso informato ed il foglio informativo per i partecipanti allo studio saranno sottomessi al CEC e, successivamente, ai comitati etici locali di ciascun centro per essere approvati. Nel caso di soggetti minori di età (<18 anni) o di soggetti incapaci di intendere e volere, il modulo di consenso informato verrà firmato e datato su base volontaria rispettivamente dai genitori/tutori del partecipante o dal suo rappresentante legale, previa adeguata informazione degli stessi secondo le procedure di cui sopra. Agli stessi verrà inoltre fornita una copia del modulo di consenso informato datato e firmato da loro stessi e dal medico proponente la partecipazione allo studio del loro figlio/tutelato/rappresentato.

Ai soggetti minori (età <18 anni), sebbene inabili a firmare un consenso informato legalmente valido, verranno comunque dettagliatamente illustrati il disegno dello studio, le sue finalità, le modalità di svolgimento dello stesso e tutte le informazioni inerenti la raccolta, la custodia e l'utilizzo dei dati personali in una modalità che sia congrua al loro livello di età e di maturità e si terrà in conto la loro volontà di partecipazione o meno allo studio.

Infine, per la raccolta dati retrospettiva, sarà possibile includere tutti i pazienti che abbiano firmato il modulo "INFORMATIVA E CONSENSO PER IL TRATTAMENTO DEI DATI PERSONALI" presso i singoli IRCCS, laddove tale modulo sia stato già predisposto



precedentemente dai singoli centri e laddove venga approvato dal CEC e, successivamente, dai singoli comitati etici locali per lo studio in oggetto. Nel caso in cui il centro non abbia predisposto un modulo apposito, possono essere inclusi retrospettivamente i pazienti che sono ancora all'interno del follow-up (consenso specifico) o che hanno terminato il consenso ma che firmino comunque il consenso informato (consenso specifico).

## BIBLIOGRAFIA

- [1] World Health Organization (WHO), W. H. Cerebrovascular disorders : a clinical and research classification. World Health Organization, WHO offset publication, n 431978, ISBN 9241700432.
- [2] Special report from the National Institute of Neurological Disorders and Stroke. Classification of cerebrovascular diseases III. Stroke 21, 637-676, doi:10.1161/01.str.21.4.637 (1990).
- [3] Wilkins E, W. L., Wickramasinghe K, Bhatnagar P, Leal J, Luengo-Fernandez R, Burns R, Rayner M, Townsend N. ( European Heart Network, Brussels, 2017).
- [4] Timmis, A. et al. European Society of Cardiology: cardiovascular disease statistics 2021. Eur Heart J 43, 716-799, doi:10.1093/eurheartj/ehab892 (2022).
- [5] Collaborators, G. B. D. S. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 20, 795-820, doi:10.1016/S1474-4422(21)00252-0 (2021).
- [6] Ministero della salute, [https://www.salute.gov.it/portale/news/p3\\_2\\_1\\_1\\_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=5682#:~:text=A%20livello%20globale%20si%20stima,1.240%2C3%20per%20100.000,2021;Data ultimo accesso: 21 maggio 2023](https://www.salute.gov.it/portale/news/p3_2_1_1_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=5682#:~:text=A%20livello%20globale%20si%20stima,1.240%2C3%20per%20100.000,2021;Data%20ultimo%20accesso:21%20maggio%202023)).
- [7] Sacco, S., Stracci, F., Cerone, D., Ricci, S. & Carolei, A. Epidemiology of stroke in Italy. Int J Stroke 6, 219-227, doi:10.1111/j.1747-4949.2011.00594.x (2011).

- [8] Ozpinar, A., Mendez, G. & Abla, A. A. Epidemiology, genetics, pathophysiology, and prognostic classifications of cerebral arteriovenous malformations. *Handb Clin Neurol* 143, 5-13, doi:10.1016/B978-0-444-63640-9.00001-1 (2017).
- [9] Ferro, J. M. & Aguiar de Sousa, D. Cerebral Venous Thrombosis: an Update. *Curr Neurol Neurosci Rep* 19, 74, doi:10.1007/s11910-019-0988-x (2019).
- [10] Feng L, Moritz S, Nowak G, Welsh AH, O'Neill TJ (2020). *\_imputeR: A General Multivariate Imputation Framework\_*. R package version 2.2, <<https://CRAN.R-project.org/package=imputeR>>.
- [11] Kassambara A (2020). *\_ggpubr: 'ggplot2' Based Publication Ready Plots\_*. R package version 0.4.0, <<https://CRAN.R-project.org/package=ggpubr>>.
- [12] Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software*, 6(61), 3167, doi:10.21105/joss.03167
- [13] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
- [14] Dawson C (2021). *\_ggprism: A 'ggplot2' Extension Inspired by 'GraphPad Prism'\_*. R package version 1.0.3, <<https://CRAN.R-project.org/package=ggprism>>.
- [15] Ahlmann-Eltze, C., & Patil, I. (2021). *ggsignif: R Package for Displaying Significance Brackets for 'ggplot2'*. *PsyArxiv*. doi:10.31234/osf.io/7awm6
- [16] Therneau T (2022). *\_A Package for Survival Analysis in R\_*. R package version 3.3-1, <<https://CRAN.R-project.org/package=survival>>.

- [17] Terry M. Therneau, Patricia M. Grambsch (2000). *\_Modeling Survival Data: Extending the Cox Model\_*. Springer, New York. ISBN 0-387-98784-3.
- [18] Kassambara A, Kosinski M, Biecek P (2021). *\_survminer: Drawing Survival Curves using 'ggplot2'\_*. R package version 0.4.9, <<https://CRAN.R-project.org/package=survminer>>.
- [19] Schemper M. Cox Analysis of Survival Data with Non-Proportional Hazard Functions. *Journal of the Royal Statistical Society. Series D (The Statistician)* 1192; 41(4): 455-65. doi: 10.2307/2349009.
- [20] Schemper M, Wakounig S, Heinze G. The estimation of average hazard ratios by weighted Cox regression. *Stat Med.* 2009;28(19):2473-2489. doi:10.1002/sim.3623
- [21] Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1):W1-W73. doi:10.7326/M14-0698
- [22] Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, And Updating*. Springer; 2009.
- [23] Xavier Robin, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez and Markus Müller (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, p. 77. DOI: 10.1186/1471-2105-12-77 <http://www.biomedcentral.com/1471-2105/12/77/>

- [24] Monica Lopez-Raton, Maria Xose Rodriguez-Alvarez, Carmen Cadarso Suarez, Francisco Gude Sampedro (2014). OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests. *Journal of Statistical Software*, 61(8), 1-36. DOI 10.18637/jss.v061.i08.
- [25] Harrell Jr FE (2022). *\_rms: Regression Modeling Strategies\_*. R package version 6.3-0, <<https://CRAN.R-project.org/package=rms>>.
- [26] R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [27] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381. doi:10.1016/j.jbi.2008.08.010
- [28] Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform.* 2019;95:103208. doi:10.1016/j.jbi.2019.103208