

VALUTAZIONE D'IMPATTO SULLA PROTEZIONE DEI DATI

DPIA (Data Protection Impact Assessment)

“REte di Genomica Integrata per Nuove Applicazioni in medicina di precisione – REGINA”

Descrizione del trattamento	Attività di ricerca clinica, nonché le attività amministrative connesse all'attività medesima (attività di gestione degli aspetti autorizzativi, contrattualistici, normativo-regolatori, giuridico-contabili relativi ai progetti di ricerca clinica e sperimentazioni).
Team elaborazione DPIA	Il presente documento è stato elaborato da: Marco Masala, Nadina Foggetti, Vito Flavio Liciulli, Alessandro Orro
Titolare del trattamento	Consiglio Nazionale delle Ricerche
RPD/DPO	Il Responsabile protezione dati per il CNR è l'ing. Roberto Puccinelli, nominato con provvedimento del direttore generale n. 91/2023 raggiungibile ai seguenti contatti: email: rpd@cnr.it oppure dpo@cnr.it ; PEC: rpd@pec.cnr.it
Data ultimo aggiornamento	
Data di validazione	
Validatore	Dr. Antonio Cerasa, Direttore Dipartimento di Scienze Biomediche

1 Sommario

2	DEFINIZIONI E ABBREVIAZIONI	3
3	FINALITÀ DEL DOCUMENTO.....	4
4	DESCRIZIONE DEL TRATTAMENTO	4
4.1	Flusso dei dati	7
5	CATEGORIE DI DATI PERSONALI E BASE GIURIDICA DEL TRATTAMENTO	13
5.1	Le categorie di dati personali	13
5.2	La base giuridica	13
6	TRASPARENZA E DURATA	15
7	ANALISI DEL CONTESTO E AVVIO DELLA VALUTAZIONE.....	16
7.1	Tipologia di trattamento.....	16
7.2	Valutazione preliminare dell'utilizzo dei dati	16
8	IMPOSTAZIONE DELL'ANALISI DI RISCHIO PRELIMINARE.....	21
8.1	Tecnologie utilizzate	21
8.2	Coinvolgimento di altre strutture	22
8.3	Modifiche alle modalità di trattamento dei dati.....	24
8.4	Modifiche alle procedure di trattamento dei dati.....	24
8.5	Esenzioni dalla applicazione delle disposizioni del GDPR (ex art.2, comma 2 GDPR).....	25
9	ESITO DELL'ANALISI PRELIMINARE DEI RISCHI.....	26
9.1	Identificazione preliminare dei rischi	26
9.2	Decisione su come procedere.....	26
9.3	Congruità con altre leggi, codici o regolamenti afferenti alla protezione dei dati	26
10	CONTENUTI ANALITICI DELLA DPIA	26
10.1	Descrizione analitica delle operazioni di trattamento, con indicazione delle finalità perseguite dal Titolare del Trattamento	26
10.2	Valutazione della necessità e proporzionalità delle operazioni di trattamento, in relazione alle finalità	26
10.3	Misure esistenti o pianificate	27
10.4	Rischi.....	29
10.5	Panoramica dei rischi	32
10.6	Indicare le salvaguardie previste se i dati personali sono trasferiti o comunicati a paesi terzi ...	33
10.7	Eventuale coinvolgimento del DP	33
11	REVISIONE E AGGIORNAMENTO, CON RIESAME DI CONGRUITÀ CON LE ESIGENZE DI PROTEZIONE DEI DATI - ART 35 GDPR.....	33
11.1	Data entro la quale deve essere condotto il riesame di congruità	34
12	APPENDICE A - Lista di controllo della congruità del trattamento previsto con le esigenze di protezione dei dati	35

2 DEFINIZIONI E ABBREVIAZIONI

DPIA	Valutazione d'impatto (Data Protection Impact Assessment) - Valutazione ai sensi dell'art. 35 del GDPR che il Titolare del trattamento dei dati è chiamato a svolgere in via preliminare ogni volta che si appresta ad eseguire un trattamento che, per la natura, lo scopo o l'ambito di applicazione presenti un elevato rischio specifico per i diritti e le libertà dell'interessato
GDPR	REGOLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati). E' il principale testo normativo europeo in materia di protezione dei dati personali
Anonimizzazione	È una operazione di trattamento che attraverso la de-identificazione trasforma in maniera irreversibile i dati personali in dati anonimi in modo tale che non si possano più attribuire a un interessato specifico.
Pseudonimizzazione	Trattamento dei dati personali in modo tale che i dati personali non possano più essere attribuiti a un interessato specifico senza l'utilizzo di informazioni aggiuntive, a condizione che tali informazioni aggiuntive siano conservate separatamente e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile
PI/Sperimentatore	Principal Investigator, responsabile dell'esecuzione dello studio. Se lo studio è svolto da un gruppo di persone nello stesso centro, lo sperimentatore responsabile del gruppo è definito Sperimentatore principale
Delegation Log	Documento in cui il PI elenca le persone idoneamente qualificate cui ha delegato compiti significativi relativi ad una sperimentazione/studio
N.A.	Non applicabile
eCRF	Case Report Form, schede informatizzate di raccolta dati utilizzate nell'ambito di studi clinici, survey, gestione di database e progetti di ricerca
Open Data	Dati aperti, accessibili a tutti, che possono essere liberamente utilizzati, riutilizzati e ridistribuiti da chiunque
Sistemi di intelligenza artificiale	Sistemi sia hardware che software che permettono di dotare le macchine di determinate caratteristiche quali, ad esempio, le percezioni visive, spazio-temporali e decisionali
ToU	Terms of Use. Questo documento stabilisce le condizioni e le regole specifiche che devono essere seguite quando si utilizza il servizio.
AUP	Acceptable Use Policy. Definisce le linee guida per l'uso accettabile dei servizi offerti dal provider.

3 FINALITÀ DEL DOCUMENTO

Il trattamento di dati personali deve avvenire nel rispetto della normativa applicabile in materia di protezione dei dati personali e, in particolare, delle disposizioni del Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 (di seguito, il “Regolamento”) e del d.lgs. n. 196 del 30 giugno 2003 (Codice in materia di protezione dei dati personali – di seguito, il “Codice”).

I dati inoltre devono essere raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in modo che non sia incompatibile con tali finalità e comunque, “adeguati, pertinenti e limitati a quanto necessario alle finalità per le quali sono trattati” (principi della limitazione della finalità e di minimizzazione dei dati - art. 5, par. 1, lett. b) e c) del Regolamento)

Con particolare riferimento al trattamento per finalità di ricerca scientifica, si evidenzia che i dati personali devono essere “trattati in modo lecito corretto e trasparente” (principio di “liceità, correttezza e trasparenza” e “in maniera da garantire un’adeguata sicurezza (...), compresa la protezione, mediante misure tecniche e organizzative adeguate, da trattamenti non autorizzati o illeciti (principio di “integrità e riservatezza”)” (art. 5, par. 1, lett. a) e f) del Regolamento).

Il Regolamento prevede poi che il titolare del trattamento valuti i rischi che un trattamento può determinare sui diritti e libertà fondamentali degli interessati e, conseguentemente, metta in atto “misure tecniche e organizzative adeguate per garantire un livello di sicurezza adeguato al rischio”, tenendo conto, tra l’altro, “della natura, dell’oggetto, del contesto e delle finalità del trattamento, come anche del rischio di varia probabilità e gravità per i diritti e le libertà delle persone fisiche” (art. 32 del Regolamento).

Pertanto, lo scopo del presente documento è quello di fornire una descrizione dell’impatto che il trattamento dei dati, relativo all’attività di studio, ricerca e sperimentazione clinica, potrebbe avere sui diritti e sulle libertà dei soggetti interessati, al fine di valutarne la necessità e la proporzionalità.

Tale valutazione è necessaria, inoltre, per consentire la gestione dei rischi derivanti dal trattamento stesso, anche attraverso la programmazione di misure organizzative e tecniche idonee a garantire il rispetto del Regolamento (UE) n. 2016/679 (GDPR).

L’attività di trattamento oggetto del presente DPIA considerati la natura, l’oggetto, il contesto e le finalità, potrebbe presentare un rischio elevato per i diritti e le libertà delle persone fisiche, secondo i criteri di cui all’art. 35, paragrafo 3, del GDPR e declinati nel provv. 467/2018 del Garante italiano dell’11 ottobre 2018. Nel seguito del documento si darà, quindi, evidenza delle misure tecniche e organizzative messe in atto per mitigare l’impatto del trattamento e ridurre così il rischio residuale per i diritti e le libertà degli interessati.

4 DESCRIZIONE DEL TRATTAMENTO

La presente valutazione ha ad oggetto il trattamento dei dati personali nell’attività di ricerca denominata “REte di Genomica Integrata per Nuove Applicazioni in medicina di precisione – REGINA”.

Il progetto REGINA è teso alla “creazione di un programma di precisione per la mappatura del genoma umano su scala nazionale”. Esso si propone di generare profili genomici ad alta risoluzione tramite tecniche di whole genome sequencing o genotipizzazione tramite SNP array, di pazienti affetti da:

- malattie cardiovascolari complesse
- cardiomiopatie e sindromi aritmogene
- disturbi neurocognitivi quali Mild Cognitive Impairment e Subjective Cognitive Decline
- demenze ad esordio precoce e demenze con elevata familiarità (Alzheimer, demenza frontotemporale)

- Multiple System Atrophy nelle manifestazioni cerebellari e parkinsoniane
- malattie genetiche rare caratterizzate da patologia degenerativa neurologiche o neuromuscolari ad eziologia genetica ignota.

Il progetto REGINA si propone di creare un'infrastruttura, pienamente integrata nell'ecosistema infrastrutturale nazionale ed europeo, in grado di supportare vari ambiti disciplinari che vanno dall'analisi genomica alla medicina di precisione. L'infrastruttura verrà utilizzata per generare profili genomici ad alta risoluzione tramite tecniche di whole genome sequencing e genotipizzazione mediante SNP array a media densità in pazienti afferenti alle tre macroaree di intervento (patologie cardiologiche, neurologiche, genetiche), al fine di generare profili di rischio genetico e identificare cause genetiche e molecolari che conducano a una migliore prevenzione, diagnosi e cura delle patologie oggetto di studio.

Il progetto, con il codice T3-AN-03, a valere sulla programmazione nazionale della Ricerca per il periodo 2015-2020, è stato ammesso al contributo con il decreto direttoriale del 20 giugno 2022 prot. n. 2351, che ha approvato la graduatoria di merito dei progetti presentati nell'ambito della Traiettorie 3 "Medicina rigenerativa, predittiva e personalizzata", Azione 3.1 del Piano Operativo Salute.

I partner di progetto sono:

- Consiglio Nazionale delle Ricerche (capofila)
- Istituto delle Scienze Neurologiche dell'Azienda USL di Bologna
- Istituto Neurologico Carlo Besta
- Ordine Ospedaliero di San Giovanni di Dio – Fatebenefratelli
- Centro Cardiologico SpA "Fondazione Monzino"
- Policlinico San Donato SpA
- Istituto Clinico Humanitas – Humanitas Mirasole SpA

L'azione progettuale sarà declinata in tre macroaree di intervento:

1. Macroarea patologie cardiologiche

Le attività previste in tale macroarea mirano al raggiungimento degli obiettivi specifici O2 e O3 e consistono nell'utilizzare il profilo genomico per:

- identificare individui con rischio elevato di sviluppare malattie cardiovascolari complesse (prevenzione primaria) o recidiva e progressione severa della malattia cardiovascolare (prevenzione secondaria) (O2);
- ottimizzare la diagnosi e alla stratificazione fenotipica di pazienti affetti da cardiomiopatie e sindromi aritmogene (O3).

2. Macroarea patologie neurologiche

Le attività previste in tale macroarea mirano al raggiungimento degli obiettivi specifici O4-O6, e consistono nell'utilizzare l'analisi genomica per identificare determinanti genetici associati a:

- progressione verso disturbi neurocognitivi maggiori in pazienti affetti da *Mild Cognitive Impairment* (MCI) e *Subjective Cognitive Decline* (SCD) (O4);
- demenze ad esordio precoce e demenze con elevata familiarità (O5); *Multiple System Atrophy* (MSA) e sua manifestazione in forme cerebellari (MSA-C) o parkinsoniane (MSA-P) (O6)

3. Macroarea patologie genetiche rare

Le attività previste in tale macroarea mirano al raggiungimento dell'obiettivo specifico 7 (O7) e consiste nell'utilizzo dell'analisi genomica per la caratterizzazione di casi irrisolti di patologie genetiche a fenotipo degenerativo neurologico o neuromuscolare. Il raggiungimento di questi obiettivi apporterà importanti elementi per un approccio personalizzato alla diagnosi e alla terapia di queste patologie complesse.

Uno schema generale della struttura dello studio REGINA è illustrato nella Tabella 1.

O2 (cardio comuni)	O3 (cardio rare)	O4 (neuro comuni)	O5 (neuro comuni)	O6 (neuro rare)	O7 (genetiche rare)
Coorti cliniche 1250 EPIFANIA 535 ProSalute 1704 POGEAC 2100 Prevital 150 PLACCA Coorti di popolazione 7000 ProgeNIA 2250 CILENTO (200)	Coorti cliniche 632 Brugada 201 cardiomiopatia dilatativa (DCM) 70 cardiomiopatia ipertrofica (HCM) 41 cardiomiopatia aritmogena (ARVC) 50 sindrome QT lungo (LQTS)	Coorti cliniche 500 mild cognitive impairment (MCI)	Coorti cliniche 1156 Alzheimer, demenza frontotemporale (AD, FTD) 700 controlli (comuni a O6)	Coorti cliniche 318 Multiple System Atrophy (MSA) 700 controlli (comuni a O5)	Coorti cliniche 50 Sindrome di Cockayne, Tricotodistrofia (CS/TTD) 90 Distrofia Muscolare (DM) 100 Distrofia fascioscapuloumerale (FSHD) 50 Distrofia Muscolare dei Cingoli (DMC) 50 Atassia 50 Paraparesi Spastica Ereditaria (PSE)
Partecipanti: Humanitas, Monzino IRGB, IGB*	San Donato* IBIOM	Besta, FBF, ISNB IBBC, IN	Besta, FBF, ISNB IBBC, IFT	Besta, ISNB	IGM, IBBC, IRIB
Approccio sperimentale:					
Genotipizzazione (*sequenziamento per IGB e San Donato)			Sequenziamento		
Calcolo del PRS e correlazione con esordio, progressione o caratteristiche cliniche della malattia			Annotazione delle varianti; analisi della segregazione	Annotazione delle varianti; analisi del carico mutazionale; associazione con varianti comuni	Annotazione delle varianti
Finalità: Predizione dell'esordio e della progressione della patologia	Classificazione e predizione della progressione e della variabilità delle manifestazioni cliniche della malattia	Predizione della progressione della patologia	Identificazione di varianti genetiche di suscettibilità	Identificazione di varianti genetiche di suscettibilità; predizione gravità della malattia	Identificazione di varianti genetiche di suscettibilità
Numerosità al 31/07/23; in rosso: da fare; in nero: già fatti con altri fondi; in blu: parzialmente fatti con altri fondi					

Tabella 1. Schema riassuntivo del progetto REGINA

L'attività di ricerca nel settore della bioinformatica sarà effettuata utilizzando i servizi forniti dall'infrastruttura del nodo italiano di ELIXIR.

L'Italia ha partecipato alla fase preparatoria di ELIXIR tramite il CNR (Dipartimento di Scienze della Vita).

Nel 2012, il MIUR ha firmato il Memorandum of Understanding (MoU) esprimendo l'interesse ad aderire all'infrastruttura di ricerca (RI). La Joint Research Unit (JRU) ELIXIR-IT è stata istituita nel 2013, coinvolgendo tutte le istituzioni accademiche e tecnologiche che durante la fase preparatoria avevano risposto a una "call for expression of interest".

Il modello legale e di governance scelto per ELIXIR sfrutta lo status giuridico dell'EMBL (progetto speciale EMBL).

L'Italia ha firmato l'ELIXIR Consortium Agreement (ECA) alla fine del 2015 ed è quindi un membro a pieno titolo di ELIXIR.

ELIXIR è stato identificato come Landmark Project nella Roadmap ESFRI 2016. Il nodo italiano, configurato come Joint Research Unit (JRU) denominata Infrastruttura Italiana di Bioinformatica (ELIXIR-IT, precedentemente ELIXIR-ITA), è coordinato dal Consiglio Nazionale delle Ricerche (CNR) e comprende oltre 30 partner, tra cui università, istituti di ricerca e fornitori pubblici di servizi Cloud e High Performance Computing (HPC). ELIXIR-IT si occupa di coordinare i servizi di bioinformatica nazionali e integrarli nell'infrastruttura ELIXIR.

Il Nodo italiano di ELIXIR ha l'ambizione di costituire una Infrastruttura Italiana di Bioinformatica (IIB) distribuita su più centri e si propone di supportare i ricercatori italiani che operano nel campo della Bioinformatica, favorendo lo scambio e lo sviluppo di competenze, di mettere a sistema le varie risorse bioinformatiche già riconosciute in ambito internazionale e pubblicamente disponibili, e di contribuire alla loro integrazione in seno all'infrastruttura europea.

Le attività di ELIXIR-IT sono suddivise in aree tecnologiche, denominate piattaforme. Coordinano la fornitura di servizi computazionali di alta qualità per le scienze della vita e guidano l'integrazione dei servizi nazionali

all'interno dell'infrastruttura ELIXIR-IT. ELIXIR-IT comprende sei piattaforme operative (Compute, Data, Interoperability, Tools, Omics e Training).

ELIXIR è un'organizzazione intergovernativa per i dati e le risorse delle scienze della vita.

ELIXIR è organizzato secondo un modello di Hub e Node, in cui l'Hub funge da coordinatore dell'infrastruttura, che è composta da servizi gestiti dai Node. L'Hub e i Node svolgono funzioni diverse e, di conseguenza, trattano differenti tipologie di dati personali.

- L'**ELIXIR Hub** si occupa dell'amministrazione e coordina l'infrastruttura di ELIXIR.
- Un **Node ELIXIR** è un insieme di istituti di ricerca all'interno di un paese membro di ELIXIR. I Node gestiscono le risorse e i servizi scientifici che fanno parte dell'infrastruttura di ELIXIR.

L'ELIXIR Hub non ha accesso ai dati scientifici elaborati dai Node per progetti interni e finanziati dall'UE, i dati genetici o sanitari raccolti nei progetti sono gestiti solo dai Nodi. Nei Node ELIXIR situati all'interno dell'UE, i dati sono protetti dalla normativa europea sulla protezione dei dati (GDPR) (cfr. <https://elixir-europe.org/about-us/data-protection#legal-model>).

4.1 Flusso dei dati

4.1.1 Raccolta dei dati e campioni biologici

Nell'ambito del progetto REGINA, i dati e i campioni biologici impiegati non sono oggetto di raccolta iniziale. Essi derivano, infatti, da coorti di studio preesistenti, le quali includono soggetti che hanno precedentemente espresso il loro consenso informato per l'analisi genetica e per il riutilizzo dei campioni a fini di ricerca. La raccolta di questi dati è stata condotta in stretta aderenza agli standard etici e legali vigenti per la protezione dei dati personali, assicurando il pieno rispetto delle normative applicabili.

Elenco delle Coorti di Studio e Dettagli sui Numeri:

1. **ProgeNIA/SardiNIA** (CNR-IRGB):
 - **Partecipanti:** Circa 7000
 - **Ubicazione:** Lanusei - Sardegna
2. **Cilento** (CNR-IGB)
 - **Partecipanti:** Circa 2300
 - **Ubicazione:** Gioi (SA); Napoli - Campania
3. **EPIFANIA** (IRCCS - Monzino):
 - **Partecipanti:** Circa 1250
 - **Ubicazione:** Milano - Lombardia
4. **ProSALUTE** (IRCCS - Monzino):
 - **Partecipanti:** Circa 535
 - **Ubicazione:** Milano - Lombardia
5. **POGEAC** (IRCCS - Monzino):
 - **Partecipanti:** Circa 1704
 - **Ubicazione:** Milano - Lombardia

6. **PREVITAL** (IRCCS Humanitas):
 - **Partecipanti:** Circa 2100
 - **Ubicazione:** Milano - Lombardia.
7. **SMAC** (IRCCS Humanitas):
 - **Partecipanti:** Circa 1000
 - **Ubicazione:** Milano - Lombardia.
8. **PLACCA** (IRCCS Humanitas):
 - **Partecipanti:** Circa 200
 - **Ubicazione:** Milano - Lombardia.
9. **Brugada PSD** (IRCCS Policlinico San Donato):
 - **Partecipanti:** Circa 1000
 - **Ubicazione:** San Donato Milanese (MI) - Lombardia.
10. **PSD cardiomyopathies** (IRCCS Policlinico San Donato)
 - **Partecipanti:** Circa 400
 - **Ubicazione:** San Donato Milanese (MI) - Lombardia.
11. **Cardiomiopatie e Sindromi Aritmogene** (CNR-IBIOM)
 - **Partecipanti:** Circa 1000
 - **Ubicazione:** Bari - Puglia.
12. **CS/TTD** (CNR-IGM)
 - **Partecipanti:** Circa 60
 - **Ubicazione:** Pavia - Lombardia; Bologna - Emilia-Romagna.
13. **Distrofia Muscolare** (CNR-IGM)
 - **Partecipanti:** Circa 100
 - **Ubicazione:** Pavia - Lombardia; Bologna - Emilia-Romagna.
14. **FSHD** (CNR-IBBC)
 - **Partecipanti:** Circa 100
 - **Ubicazione:** Monterotondo Scalo (RM) - Lazio.
15. **DMC – PSE – Atassia** (CNR-IRIB)
 - **Partecipanti:** Circa 150
 - **Ubicazione:** Mangone (RM) - Lazio.

4.1.2 Trasferimento dei dati e dei campioni biologici

Il trasferimento dei dati e dei campioni biologici provenienti dalle coorti di studio preesistenti sarà effettuato nel pieno rispetto delle normative vigenti in materia di protezione dei dati personali e sicurezza del trasferimento.

Pseudonimizzazione dei dati:

I dati personali relativi ai partecipanti delle coorti di studio preesistenti saranno forniti al progetto REGINA in forma già pseudonimizzata dai rispettivi titolari del trattamento originari.

Trasferimento dei campioni biologici:

Il trasferimento dei campioni biologici sarà effettuato in conformità con le normative nazionali e internazionali applicabili, inclusi il Regolamento (UE) 2017/745 (MDR) e il Regolamento (CE) 1394/2007, nonché le linee guida dell'Organizzazione Mondiale della Sanità (OMS) relative al trasporto sicuro di materiali biologici. Per garantire la tracciabilità e l'integrità dei campioni durante il trasporto, saranno selezionate aziende specializzate nel trasferimento di materiali biologici che operano in conformità ai più elevati standard di qualità e sicurezza. In particolare troverà applicazione quanto disposto dal Garante Provvedimento recante le prescrizioni relative al trattamento di categorie particolari di dati, ai sensi dell'art. 21, comma 1 del d.lgs. 10 agosto 2018, n. 101.

Trasferimento dei dati:

I dati pseudonimizzati saranno trasferiti in formato elettronico tramite canali criptati implementati da ELIXIR-IT. L'utilizzo di tale piattaforma garantirà:

- **Crittografia dei dati:** I dati saranno cifrati sia durante il trasferimento che nella fase di archiviazione, conformemente ai requisiti di sicurezza dell'art. 32 del GDPR.
- **Autenticazione e controllo degli accessi:** Solo personale autorizzato potrà accedere ai dati tramite tunnel criptati a chiave asimmetrica e utilizzando meccanismi di autenticazione a due fattori e/o basati su sistemi a chiave asimmetrica.
- **Audit trail:** Sarà mantenuto un registro completo di tutte le operazioni effettuate sui dati, per garantire la trasparenza e la rendicontabilità delle attività.

Il trasferimento sarà pianificato e monitorato per minimizzare ogni rischio, assicurando il rispetto degli standard di protezione richiesti dalle normative europee e internazionali.

4.1.3 Genotipizzazione e sequenziamento dei campioni di DNA

Nel presente studio, il primo obiettivo consiste nell'ottenimento dei genotipi o della sequenza dell'intero genoma a partire dai campioni di DNA precedentemente raccolti. Tali attività saranno affidate, in modalità in service, ad aziende esterne, conformemente alle linee guida del finanziamento.

Le suddette aziende verranno configurate come responsabili del trattamento dei dati personali ai sensi dell'art. 28 del Regolamento (UE) 2016/679 (GDPR). A tal fine, saranno stipulati specifici contratti di nomina che definiranno in modo chiaro ed esaustivo i compiti e le responsabilità delle parti, nonché le misure tecniche e organizzative adeguate a garantire la sicurezza e la protezione dei dati personali trattati.

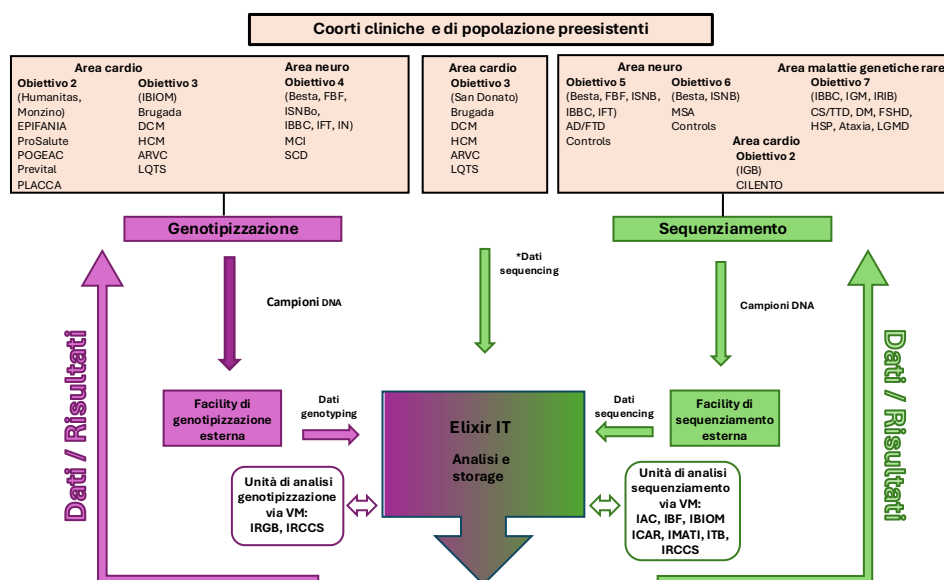
I campioni compresi nelle coorti cliniche di malattie cardiovascolari e di pazienti affetti da mild cognitive impairment (relativi agli obiettivi O2, O3 e O4), ad esclusione delle coorti del IRCCS Policlinico San Donato, saranno **sottoposti a genotipizzazione *genome-wide*** mediante l'utilizzo di SNP arrays appositamente scelti per consentire studi di ricerca clinica, convalida di associazioni, profilazione del rischio, screening preventivo e studi di farmacogenomica e applicazioni di medicina di precisione con largo impiego in aree come il rischio poligenico e la nutrigenomica. A tal proposito, **i campioni di DNA verranno inviati dagli istituti partecipanti all'azienda esterna di competenza.**

IRCCS Policlinico San Donato invierà direttamente ad ELIXIR i dati derivanti dalle attività di sequenziamento *genome-wide* autonomamente eseguite, da cui verranno estratti i dati necessari per le successive analisi statistiche.

I campioni compresi nelle coorti cliniche di malattie neurologiche, sia comuni che rare, incluse le "patologie genetiche rare," saranno **sottoposti a sequenziamento** dell'intero genoma. Tale attività è finalizzata sia alla

profilazione del rischio genetico sia all'identificazione di determinanti genetiche e molecolari associate alle suddette patologie, in linea con gli obiettivi O5, O6 e O7. A tal fine, **i campioni di DNA saranno inviati direttamente dai rispettivi istituti partecipanti alle aziende esterne incaricate**, specificamente selezionate per la realizzazione di questa attività.

Un diagramma riassuntivo del flusso dei campioni e dati nell'ambito del progetto REGINA è illustrato nella Figura 1.



*San Donato invierà direttamente ad ELIXIR i dati derivanti dalle attività di sequenziamento autonomamente eseguite

Figura 1. Schema del flusso dei dati nel progetto REGINA

4.1.4 Analisi dei dati di genotipizzazione

Verranno applicati dei controlli sui dati di genotipizzazione per escludere varianti o campioni di scarsa qualità. Saranno utilizzati o approcci standard già in uso (CNR-IRGB, CNR-IGB) sugli stessi *genotyping array* o *array* simili (Sidore et al Nat Gen 2015, Nutile et al Sci Rep 2019). Successivamente si procederà alla predizione probabilistica di varianti genetiche aggiuntive (circa 39 milioni varianti) attraverso tecniche di inferenza statistica (genotype imputation) e l'utilizzo di un pannello di riferimento di sequenze di genomi europei (Haplotype Reference Consortium) (<https://www.sanger.ac.uk/tool/sanger-imputation-service>). Anche questa tecnica di inferenza è già in uso (Sidore et al Nat Gen 2015), e permetterà di armonizzare tutte le coorti, incluse quelle già caratterizzate con altri arrays, sullo stesso insieme di varianti genetiche. Permetterà inoltre di caratterizzare anche le varianti importanti per il calcolo del rischio poligenico ma non presenti negli array di genotipizzazione.

Per ogni individuo di ciascuna coorte, verrà calcolato il rischio poligenico (*polygenic risk score*, o PRS) per malattie cardiovascolari (CAD, infarto, *heart failure*) e fattori di rischio associati (lipidi, BMI, IMT) come somma pesata, sulla base degli effetti genetici, degli alleli di rischio in varianti genetiche identificate in ampi studi di associazione dell'intero genoma (Wand et al Nature 2021). L'elenco delle stesse varianti genetiche e dei relativi effetti si otterrà utilizzando le informazioni disponibili su GWASCatalog e PGSCatalog (<https://www.ebi.ac.uk/gwas/>, <https://www.pgscatalog.org/>). I profili di rischio così ottenuti verranno correlati con gli outcomes osservati nelle diverse visite di follow-up (esordio, progressione, o caratteristiche cliniche delle diverse patologie) per valutare la loro utilità attraverso il

confronto di un modello che utilizza solo i valori clinici, un modello che utilizza solo il rischio poligenico, e un modello che unisce sia i parametri clinici che il profilo genomico.

4.1.5 Analisi bioinformatiche di sequenze di genomi umani

L'attività di "Analisi bioinformatiche di genomi umani" riguarderà, nel dettaglio, la messa a punto di pipeline per l'analisi dei dati di sequenziamento massivo dell'intero genoma (*Whole Genome Sequencing*, WGS), prevalentemente costituiti da *short read*, a partire dall'allineamento del dato di sequenza grezzo (*raw reads*) sul genoma umano di riferimento (GRCh38), fino alla determinazione ed annotazione di varianti geniche potenzialmente correlate alle patologie oggetto di indagine. L'obiettivo è l'identificazione di varianti a singolo nucleotide (*Single Nucleotide Polymorphisms*, SNP, o *Single Nucleotide Variations*, SNV), inserzioni e delezioni (InDel), variazioni del numero di copie (*Copy Number Variations*, CNV) e varianti strutturali (*Structural Variations*, SV) incluse le ripetizioni in tandem (*Short Tandem Repeats*, STR). L'identificazione di varianti strutturali si potrà eventualmente avvalere anche di tecnologie in grado di generare *long read*.

Per l'analisi delle varianti e l'identificazione delle varianti rare ad alto impatto, le frequenze delle varianti verranno confrontate con i database di riferimento (ExAC_NFE, GnomAD_Exome_NFE). Verrà data la priorità alle varianti ultrarare (ExAC_NFE minore di 0.1% oppure GnomAD_Exome_NFE minore di 0.1%); assenti nel gruppo controllo; con effetti deleteri sulla funzione o struttura della proteina (non tollerate - PolyPhen-2; CADD-score > 20).

Le conseguenze delle varianti non sinonime localizzate in porzioni codificanti del genoma verranno valutate a livello di prodotto genico per verificarne gli effetti sia sulla struttura della proteina (stabilità strutturale e eventuali cambi conformazionali) tramite modelling, analisi delle cavità e dinamica molecolare, che sulle interazioni della proteina mutata con altre proteine, acidi nucleici, ligandi etc. L'effetto delle varianti verrà studiato anche con un approccio di biologia dei sistemi (*systems biology*), che permetterà di valutare l'interazione tra i diversi attori dei sistemi biologici, correlando i processi molecolari (*pathway*) con il fenotipo della malattia neurodegenerativa considerata. Infine, verrà analizzata la rete di interazioni tra geni (*gene network analysis*), tenendo in particolare considerazione la tessuto-specificità, riconducendo le combinazioni di varianti geniche di interesse ai processi biologici coinvolti. L'approccio *pathway-based* permetterà di identificare gruppi di geni (codificanti per funzioni tra loro correlate) disregolati comuni alla stessa malattia. Infine, impiegando tecniche di similarità genetica sia *network/pathway-based* sia non, verrà effettuata la stratificazione dei pazienti per tipo di malattia, ossia verranno individuati gruppi di pazienti contraddistinti da pattern mutazionali e da fenotipi simili, onde permettere una migliore definizione delle caratteristiche genetiche associate al fenotipo.

Ove disponibili i campioni, sarà effettuata l'analisi di segregazione delle varianti selezionate identificate nelle famiglie. Il ruolo dei nuovi geni candidati identificati nelle forme di demenze sporadiche ad esordio tardivo verrà valutato tramite analisi in silico di dati disponibili in banche dati pubbliche (UKBiobank) o generati nell'ambito di progetti/consorzi cui gli Enti partecipano (IFGC consortium <https://ifgcsite.wordpress.com/> PI J Hardy, UCL, London UK EADB consortium, PI JC Lambert, Institut Pasteur de Lille, Lille). Qualora venisse confermato un ruolo degli stessi geni candidati, la validazione in coorti di pazienti con demenza ad esordio tardivo sarà effettuata mediante *gene re-sequencing* o analisi di varianti comuni. A tal fine gli istituti partecipanti dispongono di ampie coorti di casi sporadici ad esordio tardivo (circa n=400 AD e n=400 FTD).

In parallelo a questi approcci, l'attività di "Applicazioni dell'Intelligenza Artificiale all'analisi dei genomi" si occuperà di definire modelli computazionali realizzati ricorrendo alle più recenti metodologie e tecniche di Intelligenza Artificiale, principalmente basati su algoritmi di Machine Learning e Deep Learning, in grado di integrare dati genomici e clinici prodotti nel corso del progetto allo scopo di produrre nuova conoscenza in

ambito clinico. Attraverso tali modelli, per la cui definizione verranno presi in considerazione sia approcci di tipo generativo che di tipo discriminativo, sarà possibile ottenere informazioni utili i) alla stratificazione degli individui arruolati, ii) all'individuazione di bio-marcatori associati all'insorgenza e/o al decorso patologico e iii) alla valutazione del rischio basata su fattori genetici. A tale scopo verranno studiati e realizzati dei classificatori in grado di individuare automaticamente le caratteristiche patologiche discriminanti di un particolare stato patologico e differenziare i soggetti appartenenti al caso di studio in funzione di tali caratteristiche.

In conformità ai principi sanciti dal Regolamento (UE) 2016/679 (GDPR), in particolare dagli articoli 13, paragrafo 2, lettera f), e in coerenza con il Considerando 71, che riconoscono agli interessati il diritto di ottenere informazioni significative sulla logica sottesa ai processi decisionali automatizzati – con specifica attenzione al settore sanitario – verrà posta particolare attenzione alla gestione e all'ottimizzazione dell'interpretabilità dei modelli di Intelligenza Artificiale impiegati nel progetto. Pur non trattandosi di decisioni automatizzate aventi effetti giuridici o significativamente analoghi ai sensi dell'art. 22 GDPR, in un'ottica di accountability e correttezza (art. 5, par. 1, lett. a), si adotteranno accorgimenti per garantire trasparenza e comprensibilità dei modelli tramite tecniche di Explainable AI (XAI), ossia un insieme di processi che consentono all'utente di comprendere l'algoritmo (in termini di logica, più che di processo tecnologico) e quindi accettare consapevolmente i risultati dell'elaborazione. Tali processi permetteranno un uso responsabile dell'utilizzo dell'AI all'interno del progetto, contribuendo, in ottemperanza alla normativa vigente, a creare anche fiducia, equità e trasparenza.

Verrà inoltre osservato il principio di minimizzazione (art. 5, par. 1, lett. c) GDPR), limitando i trattamenti ai dati strettamente necessari

Inoltre, in linea con quanto indicato anche dall'AI Act, durante la definizione dei modelli di intelligenza artificiale per l'analisi dei genomi, verrà posta particolare attenzione sia all'accuratezza dei dati elaborati, che a quella dei modelli proposti. Per quanto riguarda i dati utilizzati per l'apprendimento, verranno presi in considerazione tutti i possibili bias (di interazione, latenti e selettivi) che condizionerebbero l'apprendimento degli algoritmi (come ad esempio dati parziali, inesatti o sbilanciati verso una determinata popolazione), viziando il comportamento del modello e producendo risultati sbilanciati verso un sottogruppo di pazienti (ad esempio più numerosamente rappresentati nel dataset), ignorando o trascurandone altri (ad esempio pazienti con condizioni cliniche sottorappresentate nel dataset). Per quanto riguarda i modelli di AI, questi verranno preventivamente testati allo scopo di includere nel progetto esclusivamente i modelli che garantiranno un adeguato livello non solo di accuratezza nella risposta, ma anche di affidabilità secondo le più attendibili misure di performance presenti in letteratura. Ciò è tanto più importante in quanto anche i modelli di classificazione basati sull'intelligenza artificiale, seppur addestrati con dati genetici e clinici rappresentativi di alta qualità, possono comunque generare output contenenti informazioni imprecise o false.

4.1.6 Gestione dei dati genetici

Il flusso di importazione, immagazzinamento, backup e utilizzo dei dati sarà implementato nelle infrastrutture di calcolo e storage dell'infrastruttura di Ricerca ELIXIR acquisite dal CNR tramite il PON infrastrutturale CNRBioOmics (MIUR: PON Ricerca e Innovazione 2014-2019), in conformità con le direttive ISO rilevanti e la normativa europea e nazionale in materia di protezione dei dati personali, nonché con i provvedimenti emanati dal Garante (misure di garanzia) relativi al trattamento di categorie particolari di dati. Tutte le analisi bioinformatiche verranno effettuate da operatori appartenenti agli istituti coinvolti nel progetto tramite accesso a macchine virtuali (*virtual machine*, VM) inserite in un Secure Processing Environment (SPE), in grado di assicurare i massimi livelli di sicurezza per la gestione di dati sensibili, quali quelli genetici. Memorizzazione, accesso e conservazione dei dati genetici, dei risultati di analisi primarie (es. file VCF delle varianti individuate nel genoma), e dei fenotipi e metadati correlati saranno attuati mediante deposizione nel database EGA (*European Genome-phenome Archive*) nella versione centrale o federata ("Local EGA" nazionale, <https://ega-archive.org>) in fase di realizzazione in Italia. EGA è

sviluppato da EMBL-EBI (UK) ed è riconosciuto come core database dell'infrastruttura ELIXIR e utilizzato da vari nodi europei ELIXIR (è ad esempio utilizzato per la memorizzazione e accesso controllato dei genomi della UK Biobank). Il database utilizza una memorizzazione criptata dei dati genomici e fenotipici ricercabili mediante una interrogazione dei metadati associati. In questo contesto garantisce la proprietà del dato in quanto l'accesso completo e il suo utilizzo viene approvato solo a seguito di un accordo mediato tra il ricercatore richiedente e un *Data Access Committee* (DAC), in capo all'istituzione proprietaria del dato.

5 CATEGORIE DI DATI PERSONALI E BASE GIURIDICA DEL TRATTAMENTO

5.1 Le categorie di dati personali

Per "dato personale" si intende "qualsiasi informazione riguardante una persona fisica identificata o identificabile ("interessato")". Inoltre, "si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale" (art. 4, par. 1, n. 1 del Regolamento).

Le categorie di dati personali oggetto della presente DPIA comprendono dati personali dei pazienti assistiti dai partner di progetto, o dai volontari, partecipanti a progetti di ricerca condotti dai partner di progetto, tra cui:

- Dati anagrafici (sesso, data di nascita e luogo di residenza)
- Dati antropometrici, dati relativi alla salute raccolti durante le visite, i ricoveri, gli accessi al pronto soccorso nonché gli esami e gli accertamenti effettuati presso la Struttura Sanitaria. In particolare, i dati riguardano diagnosi, interventi, terapie somministrate, esiti degli esami di laboratorio, dispositivi e campioni biologici.
- Dati genetici

I soggetti interessati sono:

- pazienti che hanno ricevuto dai partner di progetto del trattamento prestazioni sanitarie nell'ambito della normale pratica clinica; i pazienti possono essere anche minori, persone temporaneamente incoscienti (per es. in coma) e interdetti/inabilitati
- Soggetti volontari che hanno partecipato a precedenti progetti di ricerca condotti dai partner
- Per finalità amministrative, gestionali e/o di monitoraggio dell'attività verranno trattati anche i dati del personale coinvolto nello studio.

5.2 La base giuridica

Il trattamento di dati personali deve avvenire nel rispetto del Regolamento, secondo il quale, in particolare, i dati personali devono essere trattati "in modo lecito, corretto e trasparente nei confronti dell'interessato" (principio di «liceità, correttezza e trasparenza») (art. 5, par. 1, lett. a) del Regolamento).

L'art. 9 del Regolamento sancisce un generale divieto di trattamento delle particolari categorie di dati, tra cui rientrano i dati sulla salute e i dati genetici, a meno che non ricorra una delle specifiche esenzioni a tale divieto (art. 9, par. 2).

Più nello specifico, si evidenzia che in omaggio al principio di liceità del trattamento anche per il perseguimento di scopi di ricerca medica, i titolari devono individuare, tra quelle enucleate dal Regolamento, la base giuridica idonea e, tenendo conto che in tale contesto rilevano dati sulla salute e genetici, l'appropriata deroga al divieto di trattare tali informazioni (artt. 5, par. 1 lett. a), 6 e 9 del Regolamento).

In tale contesto rileva, in particolare, l'esenzione dal divieto al trattamento delle particolari categorie di dati di cui all'art. 9, par. 2, lett. a) del Regolamento, che ne ammette l'uso quando l'interessato abbia prestato il proprio consenso.

Il consenso è stato acquisito al momento della raccolta dei dati che verranno processati ai fini dell'attività di ricerca descritta dai Titolari del trattamento. Il Consenso riguardava la possibilità di utilizzare i dati per attività di ricerca scientifica.

Non è stato possibile acquisire un consenso specifico rispetto al trattamento dei dati ai fini del Progetto Regina, per impossibilità organizzativa.

La mole dei dati e la pluralità di Titolari coinvolti nel progetto rendono particolarmente onerosa la raccolta del consenso specifico da ciascun interessato. Un contatto individuale comporterebbe uno sforzo sproporzionato, considerando sia l'elevato numero di soggetti coinvolti sia la limitata durata temporale del progetto di ricerca. Tale attività rischierebbe di compromettere la realizzazione stessa degli obiettivi scientifici previsti.

Per garantire comunque un'adeguata informazione agli interessati e agevolare l'esercizio dei loro diritti, sono state previste le seguenti misure alternative di pubblicità:

- Pubblicazione di tutti i documenti relativi al trattamento (informativa privacy, DPIA, ecc.) sul sito web istituzionale <https://regina.cnr.it/>, in apposita sezione dedicata e facilmente accessibile
- Diffusione di un avviso sul trattamento attraverso la pubblicazione su un quotidiano a larga diffusione nazionale
- Ove possibile, esposizione di avvisi nelle sedi dei progetti originari dai quali provengono i dati, al fine di raggiungere direttamente i potenziali interessati

La base giuridica individuata per il presente trattamento è quindi l'art. 110 del d.lgs. 196/2003, comma 1, come riformato dall'art. 44, comma 1-bis del decreto-legge 2 marzo 2024, n. 19, convertito con legge 29 aprile 2024, n. 56

Inoltre, si richiamano i seguenti standard e linee guida applicabili al trattamento:

- Clinical Trials Regulation (Regulation (EU) No 536/2014).
- Provvedimento dell'Autorità Garante n. 146 del 5 giugno 2019 recante le "Prescrizioni relative al trattamento di categorie particolari di dati, ai sensi dell'art. 21, comma 1 del d.lgs. 10 agosto 2018, n. 101", con particolare riferimento alle prescrizioni relative al trattamento dei dati genetici
- Convenzione di Oviedo "per la protezione dei Diritti dell'Uomo e della dignità dell'essere umano nei confronti dell'applicazione della biologia e della medicina: Convenzione sui Diritti dell'Uomo e la biomedicina" del 4 aprile 1997.
- Regole Deontologiche per trattamenti a fini statistici o di ricerca scientifica (Provvedimento del Garante del 19 dicembre 2018 n. 515).
- Convenzione di Helsinki (WMA DECLARATION OF HELSINKI – ETHICAL PRINCIPLES FOR MEDICAL RESEARCH INVOLVING HUMAN SUBJECTS).

6 TRASPARENZA E DURATA

Nelle more di approvazione delle nuove Regole deontologiche e ferma la vigenza di quelle di cui all'allegato A5 del Codice, si applicheranno al trattamento in parole le garanzie necessarie per i trattamenti di dati sulla salute per finalità di ricerca medica, biomedica e epidemiologica, riferiti a soggetti deceduti o non contattabili per motivi organizzativi, come disposto dal provvedimento del garante del 9 maggio 2024, pubblicato in GU Serie Generale n. 130 del 5 giugno 2024.

Come accennato in premessa, la presente Valutazione d'impatto, coerentemente con le disposizioni dell'art. 110, comma 1, del Codice, sarà oggetto di pubblicazione sul sito web istituzionale nella sezione privacy dedicata all'attività di studio, ricerca e sperimentazione clinica.

Accanto al presente documento sarà resa pubblica anche l'informativa, dedicata allo studio o sperimentazione, resa ai sensi degli artt. 13 e 14 del GDPR.

Nell'informativa è altresì comunicata agli interessati la durata degli studi e i tempi di trattamento e conservazione dei dati, che non supereranno i dieci anni dalla conclusione del progetto.

In materia di trasparenza l'attività di studio e sperimentazione si conforma alle indicazioni del Consiglio di stato (Sez. VI), il quale nella pronuncia del 13 dicembre 2019 (n. 8472) ha dichiarato che "dal diritto sovranazionale emergono tre principi, da tenere in debita considerazione nell'esame e nell'utilizzo degli strumenti informatici:

- il principio di conoscibilità, per cui ognuno ha diritto a conoscere l'esistenza di processi decisionali automatizzati che lo riguardino ed in questo caso a ricevere informazioni significative sulla logica utilizzata;
- il principio di non esclusività della decisione algoritmica;
- il principio di non discriminazione algoritmica, secondo cui è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, mettendo in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali, secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche.

Inoltre, con riferimento agli studi ed alle sperimentazioni nei quali è previsto l'utilizzo di Intelligenza Artificiale coerentemente con il parere congiunto del Comitato europeo per la protezione dei dati e il Garante europeo, n. 5/2021 (sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale del 21 aprile 2021) è valorizzato un approccio graduato e adattato al rischio concreto connesso alle attività.

Sono concretamente valutati i rischi per i diritti e le libertà fondamentali degli interessati derivanti dall'utilizzo di tali strumenti, alla luce della considerazione per la quale esiste una sostanziale differenza tra le modalità con cui attività quali, generare contenuti, fare previsioni o adottare decisioni in maniera automatica, sono svolte dagli esseri umani.

Non è previsto in alcuno studio l'affidamento in via esclusiva alle macchine del compito di prendere decisioni sulla base di dati senza la sorveglianza (o supervisione) umana, al fine di assicurare il rispetto del diritto di non essere assoggettato a una decisione basata esclusivamente su un trattamento automatizzato.

7 ANALISI DEL CONTESTO E AVVIO DELLA VALUTAZIONE

7.1 Tipologia di trattamento

Attività di ricerca svolta attraverso studio multicentrico che rientra nella seguente tipologia:

1. studio retrospettivo.

7.2 Valutazione preliminare dell'utilizzo dei dati

7.2.1 Come verranno raccolti i dati?

Nel contesto del progetto REGINA, la presente attività di ricerca e di sperimentazione clinica si caratterizza per il suo approccio retrospettivo, dove i dati personali, inclusi dati di particolare sensibilità quali dati genetici e relativi allo stato di salute degli interessati, sono stati raccolti in precedenza attraverso vari progetti di ricerca. Questi dati sono ora analizzati retrospettivamente e provengono dalle seguenti fonti, a seconda delle necessità specifiche dello studio in questione:

a) Direttamente dagli interessati, in progetti precedenti; b) Cartelle cliniche e ambulatoriali preesistenti; c) Database open data; d) Dataset prodotti da enti pubblici o privati, sia nazionali che internazionali; e) Database interni ed esterni; f) Biobanche esistenti; g) Dispositivi elettronici utilizzati in studi passati.

I partner di progetto, operando in qualità di titolari autonomi del trattamento, sono responsabili per la gestione adeguata di queste collezioni di dati, assicurando che tutte le operazioni rispettino le misure di sicurezza prescritte dalle Prescrizioni del Garante, in particolare quelle dettate al punto 4.2, lettere a), b) e d), per la custodia e la sicurezza dei dati genetici e dei campioni biologici.

Inoltre, conformemente ai Material Transfer Agreements (MTA) e Data Transfer Agreements (DTA) stipulati con ciascun partner, si garantisce il rispetto di tutte le normative applicabili. Questi accordi confermano che tutti i consensi necessari sono stati ottenuti in conformità ai requisiti legali ed etici precedentemente, riflettendo l'impegno continuativo dei partner nella protezione dei dati e nel rispetto dell'etica della ricerca

7.2.2 Chi avrà accesso ai dati?

Per le attività di ricerca: PI e soggetti compresi nel delegation log.

Per le attività di manutenzione dei sistemi informativi: servizio ICT ed eventuali fornitori.

Per le attività di vigilanza: Funzione Privacy aziendale e il DPO.

I responsabili del trattamento (fornitori di servizi di genotipizzazione)

Tutti i soggetti sopra elencati sono autorizzati al trattamento dei dati ex art. 29 GDPR.

7.2.3 In che modo i dati verranno eventualmente trasferiti/comunicati?

Coerentemente con le attuali disposizioni aziendali in materia di comunicazioni e trasferimento di documenti contenenti dati sanitari, detta attività avviene attraverso l'utilizzo dello strumento PEC (Posta Elettronica Certificata) allegando i documenti de quo in formato criptato e inviando con una seconda mail o con altro strumento di comunicazione la chiave asimmetrica pubblica dell'utente correlato al trasferimento dei dati.

Inoltre, la comunicazione e il trasferimento potranno avvenire secondo la logica Cloud. Attraverso i processi di autenticazione e autorizzazione, i destinatari dei dati potranno accedere alle piattaforme condivise per l'estrazione del dataset.

I trasferimenti di campioni biologici e dati saranno sottoposti alle seguenti condizioni:

1. La consegna dei campioni sensibili (es. umani) avviene in forma anonima o pseudonimizzata;
2. Il trasporto dei campioni biologici è posto in essere con modalità volte anche a garantirne la qualità, l'integrità, la disponibilità e la tracciabilità.
3. Il trasferimento dei dati genetici in formato elettronico è effettuato mediante canali "sicuri" di comunicazione di tipo tunnel criptato, previa verifica e certificazione sui protocolli di comunicazione sicuri, sull'identità digitale del server che eroga il servizio e sulla postazione client da cui si effettua l'accesso ai dati.
4. I dati genetici e i campioni biologici contenuti in elenchi, registri o banche di dati, sono trattati con tecniche di cifratura o mediante l'utilizzazione di codici identificativi o di altre soluzioni che, considerato il numero dei dati e dei campioni trattati, li rendano temporaneamente inintelligibili anche a chi è autorizzato ad accedervi e permettano di identificare gli interessati solo in caso di necessità, in modo da ridurre al minimo i rischi di conoscenza accidentale e di accesso abusivo o non autorizzato.
5. Laddove gli elenchi, i registri o le banche di dati siano tenuti con strumenti elettronici e contengano anche dati riguardanti la genealogia o lo stato di salute degli interessati, le predette tecniche devono consentire, altresì, il trattamento disgiunto dei dati genetici e sanitari dagli altri dati personali che permettono di identificare direttamente le persone interessate.
6. Restano comunque fermi gli altri obblighi previsti dall'articolo 9 del GDPR e dal provvedimento del Garante per la tutela dei dati personali Provvedimento recante le prescrizioni relative al trattamento di categorie particolari di dati, ai sensi dell'art. 21, comma 1 del d.lgs. 10 agosto 2018, n. 101 Pubblicato sulla Gazzetta Ufficiale Serie Generale n. 176 del 29 luglio 2019), anche per ciò che attiene alla conservazione e al trasporto dei dati all'esterno dei locali protetti e all'accesso controllato a tali locali.
7. L'utente deve essere identificato, mediante Life Science AAI, ovvero altro sistema di autenticazione a due fattori indicato come equivalente.

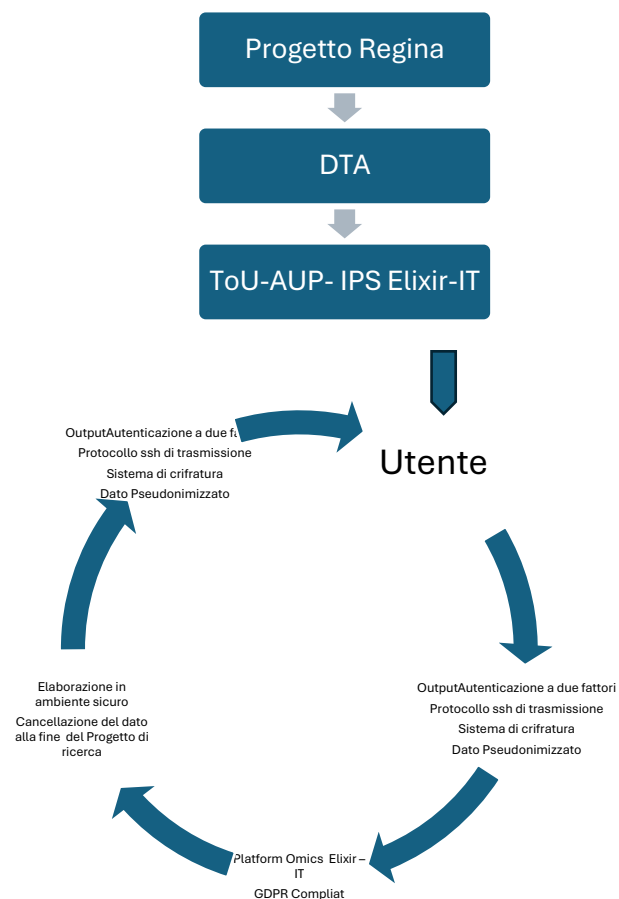
La comunicazione e il trasferimento dei dati per i trattamenti connessi alla presente valutazione d'impatto avverrà previa sottoscrizione di un DTA/MTA (Data Transfer Agreement/Material Transfer Agreement), comunque, in coerenza con quanto previsto dalla normativa vigente e pattuito nel protocollo di studio.

La gestione dei dati genetici avverrà tramite l'infrastruttura ELIXIR-IT e saranno adottate le Policy privacy, i Terms of Use e la Acceptable User Policy specificamente definite per questi servizi.

Una parte delle attività di processamento e di tipizzazione dei campioni verranno affidate mediante procedura di gara.

Il ciclo di comunicazione dei dati è descritto nella Figura 2.

Figura 2. Schema di comunicazione dei dati A- Progetto REGINA

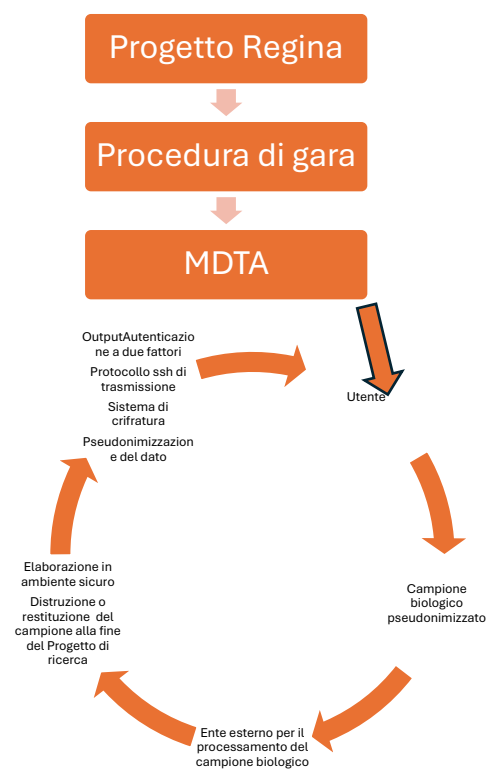


Inoltre, sono previste le seguenti misure:

- I dati e campioni biologici trasferiti dagli IRCCS ai centri responsabili dell'analisi genetica saranno pseudonimizzati mediante un identificativo univoco al fine di proteggere la privacy dei soggetti e assicurare la tracciabilità, e saranno conservati in conformità alle normative vigenti, garantendo sicurezza e riservatezza.
- L'accesso sarà limitato al personale autorizzato, qualificato e istruito su pratiche di sicurezza, conformità normativa e procedure di accesso e gestione dei dati, operante nel rispetto del segreto professionale e scientifico. I dati e i campioni saranno utilizzati esclusivamente per le finalità del progetto e non saranno comunicati a terzi senza il consenso del soggetto.
- Le analisi genomiche e fenotipiche saranno svolte con metodologie validate e standardizzate, seguendo protocolli rigorosi e criteri di qualità elevati, al fine di garantire l'affidabilità e la riproducibilità dei risultati. Le analisi saranno supervisionate da esperti del settore, che ne verificheranno la correttezza e la coerenza.
- I risultati saranno interpretati con cautela e prudenza, tenendo conto dei limiti e delle incertezze delle conoscenze scientifiche attuali, e saranno riportati con chiarezza e trasparenza, evitando ambiguità o equivoci.
- Infine, in ottemperanza a quanto previsto dal Codice novellato all'art. 110 (Ricerca medica, biomedica ed epidemiologica, come modificato da ultimo dalla legge 56/2024) prima del trasferimento di dati e campioni biologici al CNR i 6 IRCCS partner di progetto potrebbero richiedere un parere al CET di propria competenza in ottemperanza alle linee guida e buone prassi adottate da ogni istituzione. Per quanto concerne il CNR tali competenze sono attribuite al CEN come da decreti del Ministero della salute del 1° febbraio 2022 di individuazione dei comitati etici a valenza nazionale e del 30 gennaio 2023 riguardante la definizione dei criteri per la composizione e il funzionamento dei comitati etici. Il progetto è inoltre stato sottomesso alla Commissione per l'Etica e l'Integrità nella Ricerca del CNR che ne ha approvato finalità, metodologie e modalità di svolgimento rendendo parere di Ethical Clearance con valore autorizzativo di cui al prot. n.0239724 del 1° agosto 2023.

Qualora il trasferimento riguardi il campione biologico, la procedura seguirà il flusso descritto nella Figura 3.

Figura 3. Schema di comunicazione dei dati B - Progetto REGINA



7.2.4 Condivisione dei risultati:

I risultati saranno condivisi, tra i titolari autonomi partecipanti al progetto, solo sotto forma di dati aggregati, ovvero dopo essere stati sottoposti ad un processo di combinazione di informazioni provenienti da fonti diverse in un formato sintetico per l'analisi e per scopi statistici, non vi sarà la condivisione dei dati personali neppure nella forma di dati pseudonimizzati.

7.2.5 Come verranno archiviati, aggiornati ed eliminati i dati quando non più necessari?

I dati relativi agli studi saranno aggiornati in funzione delle attività svolte nelle varie fasi previste, archiviati per un tempo non superiore a 10 anni dal termine del progetto, e in seguito eliminati o resi anonimi. La conservazione del campione biologico e del dato di input e di output della piattaforma ELIXIR è a cura dell'utente che chiede il processamento del dato o del campione.

I dati contenuti nelle fonti dati seguono il ciclo di vita proprio definito dalla normativa di settore (massimario di scarto).

8 IMPOSTAZIONE DELL'ANALISI DI RISCHIO PRELIMINARE

Il trattamento dei dati per finalità di ricerca o sperimentazione clinica rientra nei casi per i quali è prevista la necessaria conduzione di una DPIA (Data Protection Impact Assessment), ovvero:

1. Trattamenti sistematici ed estensivi di valutazione di aspetti personali dell'interessato, basati su sistemi automatizzati, inclusa la profilazione, i cui esiti portino a decisioni che possono avere effetti legali diretti ed indiretti sull'interessato (ex articolo 35, paragrafo 3, lett. a) del GDPR);
2. Trattamento, su larga scala, di categorie particolari di dati personali di cui all'articolo 9, paragrafo 1, o di dati afferenti a profili penali e giudiziari come illustrato nell'articolo 10 del GDPR (ex articolo 35, comma 3, lett. b) del GDPR);
3. Altre attività di trattamento che siano inseriti nell'elenco pubblico dell'autorità garante nazionale e che richiedono specificamente lo sviluppo di una DPIA (cfr Allegato n. 1 al Provvedimento Garante Privacy n. 467 dell'11/10/2018);
4. Trattamenti in cui una violazione dei dati può avere un impatto negativo sulla protezione dei dati stessi, nonché la riservatezza e i diritti o i legittimi interessi degli interessati coinvolti;

Trattamenti per i quali la valutazione d'impatto è prevista da disposizioni normative e/o regolamenti

8.1 Tecnologie utilizzate

8.1.1 In questo trattamento verranno utilizzate nuove tecnologie informatiche che potrebbero avere un significativo potenziale di violazione della protezione dei dati personali e riduzione del livello di protezione dei dati, che bisogna garantire agli interessati?

Il trattamento in questione prevede l'utilizzo di tecnologie che possono ridurre o provocare un potenziale rischio per gli interessati:

Verranno utilizzati modelli computazionali di Intelligenza Artificiale (IA) basati su algoritmi di Machine Learning e Deep Learning, al fine di fornire informazioni utili i) alla stratificazione degli individui arruolati, ii) all'individuazione di bio-marcatori associati all'insorgenza e/o al decorso patologico e iii) alla valutazione del rischio basata su fattori genetici.

In conformità ai principi sanciti dal **Regolamento**, che riconoscono agli interessati il diritto di ottenere

informazioni significative sulla logica sottesa ai processi decisionali automatizzati – con specifica attenzione al settore sanitario – si dedicherà particolare cura alla gestione e all’ottimizzazione dell’**interpretabilità** dei modelli di Intelligenza Artificiale impiegati nel progetto e alla **trasparenza** dei risultati da essi prodotti.

I dati raccolti su eCRF, in ossequio al principio di minimizzazione, sono esclusivamente quelli definiti nel dataset dello studio, sono resi in forma pseudonimizzata e inseriti nelle piattaforme messe a disposizione dai Partner di progetto.

Anche quando i dati sono registrati su piattaforme fornite dai promotori degli studi sono rispettati tutti i principi previsti all’interno del GDPR. Le piattaforme offerte dai promotori, infatti, sono compliant con la disciplina in materia di dati personali e sono datate di regole per l’uso dei servizi offerti. (cnf. Par. [7.2.3](#))

8.1.2 Metodi di identificazione degli interessati

In ossequio alle procedure e ai regolamenti aziendali, l’identificazione degli interessati (già effettuata nell’ambito degli studi originari dai quali i dati provengono) è stata realizzata tramite sistemi tradizionali, quali ad esempio:

- la carta d’identità o altro documento di riconoscimento (codice fiscale);
- codificazioni derivanti da numeri nosologici;
- codici di pseudonimizzazione assegnati in fase di rilevazione e registrati in appositi Data Base.

Il trattamento dei dati e dei campioni viene effettuato nel rispetto dei principi etici, a tutela dei diritti e delle libertà fondamentali dei donatori.

- Sono esclusi dall’impiego per le finalità del progetto REGINA i campioni biologici:
- riguardo ai quali i donatori avevano espresso una volontà negativa circa l’utilizzo per ulteriori finalità di ricerca;
- riguardo ai quali i donatori non avevano espresso alcuna volontà in merito a un eventuale utilizzo per ulteriori finalità di ricerca o tale volontà non risulta documentata;
- per i quali, nel consenso informato originario, era prevista la distruzione al termine delle attività autorizzate.

Sulla base del criterio normativo della compatibilità con le finalità originarie della raccolta, sono invece ammessi all’impiego per il progetto REGINA quei campioni biologici per i quali la documentazione trasmessa evidenzia esplicitamente e chiaramente una volontà positiva alla donazione anche ai fini di ulteriori attività di ricerca.

8.2 Coinvolgimento di altre strutture

8.2.1 Questa iniziativa di trattamento coinvolge altre strutture, sia pubbliche, sia private, sia appartenenti a settori non-profit e volontari?

Si tratta di uno studio multicentrico, cui partecipano:

- Consiglio Nazionale delle Ricerche (capofila)
- Istituto delle Scienze Neurologiche dell’Azienda USL di Bologna
- Istituto Neurologico Carlo Besta
- Ordine Ospedaliero di San Giovanni di Dio – Fatebenefratelli
- Centro Cardiologico SpA “Fondazione Monzino”
- Policlinico San Donato SpA

8.2.2 Centri di coordinamento:

Dipartimento di Scienze Biomediche, Consiglio Nazionale delle Ricerche (DSB-CNR), Roma e
Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche (IRGB-CNR), Monserrato (CA)

8.2.3 Responsabili della ricerca:

Antonio Cerasa, Direttore DSB-CNR, Legale Rappresentante

Marcella Devoto, Associato IRGB-CNR, Responsabile Scientifica

I partners partecipano al raggiungimento degli obiettivi definiti attraverso l'assegnazione di questi ruoli:

8.2.4 Obiettivi specifici per gruppi partecipanti

01: Creazione infrastruttura per l'applicazione dell'analisi genomica alla medicina di precisione

Partecipanti: IBIOM, ITB, ICAR, IMATI, IGM, IAC, IBF, IRGB, CID Ethics

02: Utilizzo del profilo genomico per identificare individui con rischio elevato di sviluppare malattie cardiovascolari complesse (prevenzione primaria) o recidiva e progressione severa della malattia cardiovascolare (prevenzione secondaria)

Partecipanti: IRGB, IGB, Monzino, Humanitas

03: Utilizzo del profilo genomico come ausilio alla diagnosi e alla stratificazione fenotipica di pazienti affetti da cardiomiopatie e sindromi aritmogene

Partecipanti: IRGB, IBIOM, San Donato

04: Utilizzo del profilo genomico per l'identificazione di determinanti genetici associati alla progressione verso disturbi neurocognitivi maggiori in pazienti affetti da *Mild Cognitive Impairment* (MCI) e *Subjective Cognitive Decline* (SCD)

Partecipanti: IRGB, IBBC, IN, Fatebenefratelli, Neurologico BO, Besta

05: Identificazione di nuovi determinanti genetici ad alto impatto associati a demenze ad esordio precoce e demenze con elevata familiarità

Partecipanti: IRGB, IBBC, Fatebenefratelli, Neurologico BO, Besta

06: Caratterizzazione del carico mutazionale (*variant burden analysis*) in pazienti affetti da *Multiple System Atrophy* (MSA) per l'identificazione di determinanti genetici associati all'insorgenza della malattia e alla sua manifestazione in forme cerebellari (MSA-C) o parkinsoniane (MSA-P)

Partecipanti: IRGB, ITB, Neurologico BO, Besta

07: Caratterizzazione genetica di casi irrisolti di patologie degenerative neurologiche o neuromuscolari

Partecipanti: IGM, IBBC, IRIB

I singoli partecipanti all'attività di ricerca sono titolari autonomi del trattamento. Qualora necessario, saranno individuati specifici accordi per il trasferimento dei dati e dei materiali nonché nomine come responsabili esterni del trattamento ex art 28 GDPR. Tutti i trasferimenti di dati saranno sottoposti ai modelli di ToU e AUP introdotti ed in uso presso la piattaforma ELIXIR-IT al fine di garantire il rispetto delle procedure di sicurezza e la compliance con la disciplina del GDPR.

Una copia della documentazione è allegata al presente documento.

8.3 Modifiche alle modalità di trattamento dei dati

8.3.1 Questo trattamento apporterà nuove o significative modifiche alle modalità di trattamento dei dati personali, che potrebbero destare preoccupazioni dell'interessato?

Non si ravvisano modifiche sostanziali tali da ingenerare preoccupazioni o incertezze negli interessati. L'informativa fornita contiene tutti gli elementi necessari per garantire la comprensione del trattamento in ogni sua fase ed è resa disponibile anche a coloro che non sono direttamente contattabili. In conformità a quanto previsto dall'articolo 6, comma 3, delle Regole Deontologiche, il Titolare adotta adeguate forme di pubblicità, come la pubblicazione dell'informativa sul sito **regina.cnr.it**, e l'acquisto di spazi pubblicitari su un quotidiano a diffusione nazionale. Tali misure sono dirette a garantire la massima trasparenza delle finalità e delle modalità di trattamento, pur in presenza di un numero elevato di interessati e di una complessa organizzazione del progetto.

8.3.2 Questa iniziativa di trattamento apporterà nuove o significative modifiche alle modalità di consolidamento, interscambio, riferimenti incrociati, abbinamento di dati personali, provenienti da più sistemi di trattamento?

Nell'ambito del progetto REGINA, sono previste diverse fasi di elaborazione e analisi dei dati personali, che possono includere l'applicazione di sistemi di Intelligenza Artificiale (IA), con l'obiettivo di generare risultati aggregati e di natura statistica. Tali attività comportano l'integrazione, l'interscambio, il confronto incrociato e la combinazione di dati provenienti da più fonti, senza tuttavia apportare modifiche sostanziali tali da incidere negativamente sulla protezione dei dati personali.

In ogni passaggio, sono adottate misure tecniche e organizzative idonee a garantire la qualità e l'integrità dei dati raccolti. Eventuali trasferimenti o interscambi di dati tra diversi sistemi di trattamento sono disciplinati da specifici accordi (ad esempio, nomine a responsabile del trattamento, ai sensi degli articoli 26 e 28 del Regolamento (UE) 2016/679 – GDPR), con lo scopo di assicurare che l'intero processo, dalla raccolta fino alla produzione dei risultati statistici, avvenga nel pieno rispetto delle disposizioni vigenti in materia di protezione dei dati personali.

La finalità esclusivamente statistica e di ricerca scientifica, insieme all'adozione di procedure di pseudonimizzazione e/o aggregazione, contribuisce a minimizzare i rischi di reidentificazione degli interessati, in conformità ai principi di minimizzazione dei dati e limitazione delle finalità. Di conseguenza, benché vengano introdotti ulteriori step di elaborazione, non si determinano nuove o rilevanti modifiche tali da suscitare preoccupazioni in merito al rispetto dei diritti e delle libertà fondamentali degli interessati.

8.4 Modifiche alle procedure di trattamento dei dati

8.4.1 Questo trattamento potrà introdurre nuove modalità e procedure di raccolta dei dati, che non siano sufficientemente trasparenti?

L'informativa resa agli interessati contiene tutti gli elementi che descrivono le modalità di raccolta, pertanto, si ritiene che le procedure siano trasparenti e che non si possano verificare situazioni di intrusività nella sfera personale non conosciuta o non condivisa.

8.4.2 Questo trattamento introdurrà nuove o modificate modalità di conservazione dei dati non chiare o prolungate oltremodo?

I riferimenti riguardo ai tempi di conservazione sono espressamente indicati e resi conosciuti attraverso l'informativa e fanno riferimento al protocollo di studio o sono determinati dal massimario di scarto.

8.4.3 Questo trattamento modificherà le modalità di messa a disposizione dei dati?

No, il trattamento non prevede una modifica alle modalità con le quali i dati sono messi a disposizione. Si richiama per opportuno riferimento il punto 7.2.4 relativamente alla comunicazione o al trasferimento dei dati.

8.5 Esenzioni dalla applicazione delle disposizioni del GDPR (ex art.2, comma 2 GDPR)

8.5.1 L'attività di trattamento esula dall'ambito delle disposizioni legislative dell'Unione Europea?

Per i paesi al di fuori dello Spazio Economico Europeo (extra UE) i dati saranno trasferiti esclusivamente nel caso in cui sia stata emanata una decisione di adeguatezza o nei seguenti casi:

- se l'interessato è stato informato dal titolare dell'assenza di una decisione di adeguatezza e dei conseguenti rischi e ha espresso il proprio consenso al trasferimento;
- sulla base di accordi/contrattuali stipulati che forniscono garanzie adeguate agli interessati (CCS Clausole Contrattuali Standard);
- per i gruppi di imprese, il trasferimento deve avvenire sulla base di norme vincolanti di impresa che devono essere approvate dall'autorità competente (in Italia, il Garante della Privacy);
- il trasferimento è necessario per l'esecuzione di un contratto concluso su richiesta dell'interessato e del titolare;
- sulla base dell'adesione al DPF (Data Privacy Framework che consente una tutela adeguata degli interessati nei trasferimenti di dati personali tra Europa e Stati Uniti).

8.5.2 Si prevede di effettuare una consultazione istituzionale?

Alla luce delle modifiche apportate all'art. 110, comma 1, del d.lgs. 196/2003, dal decreto-legge 2 marzo 2024, n. 19, convertito con legge 29 aprile 2024, n. 56, non è previsto lo svolgimento di una consultazione istituzionale.

8.5.3 Le finalità del trattamento sono chiare e sufficientemente pubblicizzate?

Sì. Il trattamento dei dati personali, effettuato nell'ambito della ricerca, nelle sue diverse forme è descritto in maniera chiara e intellegibile nell'informativa ex art. 13 del GDPR. In tale documento sono elencate le finalità e le modalità del trattamento, così da garantire la trasparenza e la correttezza dei trattamenti effettuati da parte del titolare.

Tutti i documenti relativi alle attività di studio (informativa, DPIA, ecc.) sono pubblicati sul sito web <https://regina.cnr.it/>. Inoltre, il trattamento verrà reso noto al pubblico attraverso la pubblicazione di un avviso su un quotidiano a larga diffusione nazionale.

9 ESITO DELL'ANALISI PRELIMINARE DEI RISCHI

9.1 Identificazione preliminare dei rischi

La tabella seguente illustra i principali rischi afferenti alla protezione dei dati, che sono stati identificati in fase di valutazione preliminare.

	Descrizione del rischio	Valutazione preliminare di esposizione
Rischio 1	Accesso illegittimo ai dati	Medio
Rischio 2	Modifiche indesiderate dei dati	Trascurabile
Rischio 3	Perdita di dati	Trascurabile

9.2 Decisione su come procedere

Tenendo conto della tipologia di trattamento e in conformità alle indicazioni previste all'articolo 35, paragrafo 3 del Regolamento Generale sulla protezione dei dati – Regolamento (UE) 2016/679 – GDPR, è necessario effettuare la DPIA sul trattamento della ricerca in ogni sua declinazione.

9.3 Congruità con altre leggi, codici o regolamenti afferenti alla protezione dei dati

In relazione al provvedimento dell'Autorità Garante n. 146/2019, è stata effettuata una verifica di conformità, come parte di questa DPIA, secondo quanto illustrato nell'Appendice A, e si è giunti alla conclusione che l'attività di trattamento oggetto della presente DPIA è conforme alle prescrizioni del provvedimento indicato.

10 CONTENUTI ANALITICI DELLA DPIA

10.1 Descrizione analitica delle operazioni di trattamento, con indicazione delle finalità perseguite dal Titolare del Trattamento

I dati personali, ivi inclusi quelli appartenenti alle categorie particolari ai sensi dell'articolo 9 del Regolamento (UE) 2016/679 (GDPR) – nello specifico, dati genetici e relativi allo stato di salute dei volontari – sono raccolti e trattati con la finalità di svolgere studi di ricerca scientifica, le cui caratteristiche e modalità operative risultano già illustrate nel Capitolo 3 del presente documento.

10.2 Valutazione della necessità e proporzionalità delle operazioni di trattamento, in relazione alle finalità

La necessità e la proporzionalità delle operazioni di trattamento si valutano in maniera positiva in quanto sono presenti le seguenti misure:

- finalità determinate, esplicite e legittime;
- liceità del trattamento;
- dati personali adeguati, pertinenti e limitati a quanto necessario;
- limitazione della conservazione.

10.3 Misure esistenti o pianificate

10.3.1 Formazione e Sensibilizzazione del Personale

Tutti gli incaricati al trattamento sono regolarmente invitati a seguire corsi sulla protezione dei dati personali. Saranno adeguatamente informati del piano di gestione dei dati, della politica di accesso ai dati e del codice di condotta. Tutti i membri che accedono ai dati e cesseranno il rapporto di lavoro non avranno più accesso agli stessi tramite cancellazione utente.

10.3.2 Sicurezza dei canali informatici

Il trasferimento dei dati genetici in formato elettronico è effettuato mediante canali “sicuri” di comunicazione mediante i servizi offerti dalla piattaforma Compute di ELIXIR-IT.

I dati in transito sono protetti mediante crittografia conforme agli standard riconosciuti a livello internazionale sia di natura simmetrica (es. AES-256 o equivalenti) che di natura asimmetrica

10.3.3 Controlli di Accesso e Autenticazione Forte

Autenticazione a più fattori (MFA) per l'accesso ai sistemi.

Ruoli e profili di accesso basati sul “principio del minimo privilegio” (least privilege), concedendo a ciascun utente solo i permessi necessari per svolgere le proprie mansioni.

Soluzioni di Identity and Access Management (IAM) e aggiornamento costante dei diritti di accesso.

10.3.4 Crittografia e Pseudonimizzazione Robusta

Il progetto prevede l'utilizzo esclusivo di dati pseudonimizzati sin dall'origine. I database di associazione e le chiavi di decrittazione, che consentono di ricondurre i dati pseudonimizzati all'identità degli interessati, sono conservati dai soggetti che forniscono i dati (gli Istituti CNR e gli IRCCS coinvolti), i quali adottano misure tecniche e organizzative idonee a garantire la sicurezza delle informazioni.

I dati, sia in transito che a riposo, sono protetti mediante crittografia conforme agli standard riconosciuti a livello internazionale (es. AES-256 o equivalenti).

10.3.5 Controllo delle Versioni e Back-up Periodici

Si effettuano backup regolari, conservati in ambienti segregati e sicuri, per garantire la possibilità di recupero in caso di danneggiamento o manomissione.

Sono impostati backup giornalieri e data retention di almeno 30 gg su sistemi informatici separati

10.3.6 Audit Log e Tracciabilità degli Accessi

Il sistema informatico implementa un meccanismo di audit log per registrare tutte le attività svolte sui dati personali, consentendo di individuare eventuali accessi o utilizzi non autorizzati e di determinare le cause di eventuali incidenti di sicurezza.

I log files contengono informazioni dettagliate sugli accessi, le operazioni effettuate e i timestamp, in conformità al principio di responsabilizzazione.

Tali registrazioni sono conservate in modo sicuro per un periodo minimo di 6 mesi, come previsto dalle best practice di sicurezza informatica e dalle linee guida applicabili.

10.3.7 Valutazioni di Sicurezza e Aggiornamenti Regolari

Il sistema informatico è sottoposto a valutazioni regolari di sicurezza per identificare e mitigare eventuali vulnerabilità. Tali valutazioni includono:

- Audit periodici: Analisi programmate delle configurazioni di sicurezza, dei permessi di accesso e delle misure implementate per garantire la protezione dei dati personali.
- Penetration test: Test simulati per identificare potenziali falle di sicurezza, con cadenza almeno annuale o più frequente in caso di aggiornamenti significativi.
- Revisione delle politiche di sicurezza: Verifica regolare delle policy per assicurare che siano aggiornate rispetto agli standard tecnologici e normativi più recenti.
- Inoltre, il sistema viene mantenuto aggiornato mediante:
- Aggiornamenti software: Applicazione tempestiva di patch di sicurezza e aggiornamenti di sistema per prevenire l'esposizione a vulnerabilità conosciute.
- Monitoraggio continuo: Implementazione di strumenti di monitoraggio per rilevare attività sospette, anomalie o accessi non autorizzati in tempo reale.

10.3.8 Valutazione dei Fornitori e Contratti di Responsabilità

I fornitori sono selezionati sulla base di una valutazione delle loro capacità di protezione dei dati, incluse certificazioni (es. ISO/IEC 27001) e politiche di sicurezza.

I rapporti sono regolati da contratti conformi all'art. 28 GDPR, che includono: obbligo di trattare i dati su istruzioni del titolare, adozione di misure di sicurezza adeguate, notifica di data breach e autorizzazione per eventuali sub-responsabili.

I fornitori sono soggetti a audit regolari per verificare la conformità alle normative e agli obblighi contrattuali, con particolare attenzione ai dati sensibili come quelli genetici o sanitari.

10.3.9 Lotta contro il malware

Il sistema informatico adotta misure avanzate per la prevenzione, il rilevamento e la mitigazione del malware, in conformità con l'art. 32 GDPR. Queste misure includono:

- Soluzioni antivirus e antimalware: Software aggiornati regolarmente per proteggere da minacce conosciute e nuove varianti.
- Monitoraggio continuo: Sistemi di rilevamento delle intrusioni (IDS/IPS) e strumenti di analisi comportamentale per identificare attività sospette in tempo reale.
- Policy di utilizzo sicuro: Restrizioni su installazione di software non autorizzato, con gestione centralizzata degli aggiornamenti.
- Formazione periodica: Sensibilizzazione degli utenti sui rischi del malware, come phishing e allegati sospetti.
- Piani di risposta agli incidenti: Protocolli per contenere e risolvere rapidamente eventuali infezioni, con ripristino dei dati e report alle autorità, se necessario.

Le misure sono sottoposte a verifiche regolari per garantirne l'efficacia e l'aggiornamento rispetto alle minacce emergenti.

10.3.10 Partizionamento

Nel sistema informatico, i dati genetici e i dati sanitari sono trattati in maniera separata attraverso un approccio di partizionamento logico e fisico.

10.3.11 Minimizzazione dei dati

I dati trattati sono esclusivamente quelli necessari, previsti dal protocollo di studio.

10.4 Rischi

10.4.1 Accesso illegittimo ai dati

10.4.1.1 Quali potrebbero essere i principali impatti sugli interessati se il rischio si dovesse concretizzare?

Discriminazione e stigmatizzazione, Impatto psicologico ed emotivo, Riduzione della fiducia nel sistema di ricerca e cura, Stato di ansia e sfiducia causato dall'incertezza sull'accuratezza e la sicurezza dei propri dati genetici e sanitari.

10.4.1.2 Quali sono le principali minacce che potrebbero concretizzare il rischio?

Cyber-attacchi esterni (hacking), Minaccia interna (insider threat), Compromissione di fornitori o terze parti, Furto o smarrimento di dispositivi e supporti di memorizzazione, Carenze nelle misure di sicurezza e configurazioni errate, divulgazione accidentale di informazioni riservate, Malware

10.4.1.3 Quali sono le fonti di rischio?

Tecnologia e Infrastruttura IT, Fattori Organizzativi e di Governance, Fornitori e Terze Parti, Errori o comportamenti umani

10.4.1.4 Quali misure fra quelle individuate contribuiscono a mitigare il rischio?

Formazione e Sensibilizzazione del Personale, Sicurezza dei canali informatici, Controlli di Accesso e Autenticazione Forte, Crittografia e Pseudonimizzazione Robusta, Audit Log e Tracciabilità degli Accessi, Valutazioni di Sicurezza e Aggiornamenti Regolari, Valutazione dei Fornitori e Contratti di Responsabilità, Lotta contro il malware, Partizionamento, Minimizzazione dei dati

10.4.1.5 Come stimereste la gravità del rischio, specialmente alla luce degli impatti potenziali e delle misure pianificate?

Limitata, L'impatto sugli interessati in caso di accesso non autorizzato è valutato come a rischio limitato, grazie alla pseudonimizzazione dei dati e all'implementazione di rigorose misure di sicurezza informatica. Tra queste, è inclusa la conservazione separata del codice di pseudonimizzazione, che impedisce la diretta associazione dei dati con l'identità degli interessati.

10.4.1.6 Come stimereste la probabilità del rischio, specialmente con riguardo alle minacce, alle fonti di rischio e alle misure pianificate?

Limitata, Considerando:

Il settore di ricerca biomedica non particolarmente esposto a minacce sistematiche e note.

Le misure di sicurezza tecniche e organizzative pianificate o già in essere.

La probabilità del rischio è valutata come **Limitata**

10.4.2 Modifiche indesiderate dei dati

10.4.2.1 Quali sarebbero i principali impatti sugli interessati se il rischio si dovesse concretizzare?

Perdita di tempo dovuto alla necessità di ripetizione di azioni già svolte, Stato di ansia e sfiducia causato dall'incertezza sull'accuratezza e la sicurezza dei propri dati genetici e sanitari., Riduzione della fiducia nel sistema di ricerca e cura

10.4.2.2 Quali sono le principali minacce che potrebbero consentire la concretizzazione del rischio?

Cyber-attacchi esterni (hacking), Minaccia interna (insider threat), Vulnerabilità tecniche dei sistemi, Compromissione di fornitori o terze parti, configurazione errata di sistemi di sicurezza, Errore umano o gestione inadeguata, Piattaforme di condivisione file

10.4.2.3 Quali sono le fonti di rischio?

Tecnologia e Infrastruttura IT, Fattori Organizzativi e di Governance, Fornitori e Terze Parti, Errori o comportamenti umani

10.4.2.4 Quali misure, fra quelle individuate, contribuiscono a mitigare il rischio?

Formazione e Sensibilizzazione del Personale, Controlli di Accesso e Autenticazione Forte, Controllo delle Versioni e Back-up Periodici, Sicurezza dei canali informatici, Audit Log e Tracciabilità degli Accessi, Valutazioni di Sicurezza e Aggiornamenti Regolari, Valutazione dei Fornitori e Contratti di Responsabilità, Lotta contro il malware

10.4.3 Come stimereste la gravità del rischio, in particolare alla luce degli impatti potenziali e delle misure pianificate?

Trascurabile, L'impatto sugli interessati è limitato e le misure di sicurezza adottate rendono trascurabile il rischio.

10.4.4 Come stimereste la probabilità del rischio, specialmente con riguardo a minacce, fonti di rischio e misure pianificate?

Trascurabile, La probabilità che l'evento si verifichi è valutata come bassa, considerando le precauzioni e le misure di sicurezza adottate per proteggere i dati e gestire eventuali anomalie.

10.4.5 Perdita di dati

10.4.5.1 Quali potrebbero essere gli impatti principali sugli interessati se il rischio dovesse concretizzarsi?

Perdita di tempo dovuto alla necessità di ripetizione di azioni già svolte, Stato di ansia e sfiducia causato dall'incertezza sull'accuratezza e la sicurezza dei propri dati genetici e sanitari., Riduzione della fiducia nel sistema di ricerca e cura

10.4.5.2 Quali sono le principali minacce che potrebbero consentire la materializzazione del rischio?

Cyber-attacchi esterni (hacking), Minaccia interna (insider threat), Errore umano o gestione inadeguata, Compromissione di fornitori o terze parti, Furto o smarrimento di dispositivi e supporti di memorizzazione, Malware, Vulnerabilità tecniche dei sistemi, divulgazione accidentale di informazioni riservate, disastri ambientali/fisici

10.4.5.3 Quali sono le fonti di rischio?

Errori o comportamenti umani, Tecnologia e Infrastruttura IT, Fattori Organizzativi e di Governance, Fornitori e Terze Parti

10.4.5.4 Quali misure, fra quelle individuate, contribuiscono a mitigare il rischio?

Controllo delle Versioni e Back-up Periodici, Controlli di Accesso e Autenticazione Forte, Formazione e Sensibilizzazione del Personale, Audit Log e Tracciabilità degli Accessi, Sicurezza dei canali informatici, Valutazioni di Sicurezza e Aggiornamenti Regolari, Valutazione dei Fornitori e Contratti di Responsabilità, Lotta contro il malware

10.4.5.5 Come stimereste la gravità del rischio, specialmente alla luce degli impatti potenziali e delle misure pianificate?

Trascurabile. L'impatto sugli interessati è limitato e le misure di sicurezza adottate rendono trascurabile il rischio.

10.4.5.6 Come stimereste la probabilità del rischio, specialmente con riguardo alle minacce, alle fonti di rischio e alle misure pianificate?

Trascurabile, La probabilità che l'evento si verifichi è valutata come trascurabile, considerando le precauzioni e le misure di sicurezza adottate per proteggere i dati e gestire eventuali anomalie.

10.5 Panoramica dei rischi

Impatti potenziali

Discriminazione e stigmatizzazione
Impatto psicologico ed emotivo
Riduzione della fiducia nel
Stato di ansia e sfiducia c...
Perdita di tempo dovuto all

Minaccia

Cyber-attacchi esterni (hac...
Minaccia interna (insider t...
Compromissione di fornito...
Furto o smarrimento di disp...
Carenze nelle misure di sic...
divulgazione accidentale di...
Malware
Vulnerabilità tecniche dei ...
configurazione errata di si...
Errore umano o gestione in...
Piattaforme di condivisione...
disastri ambientali/fisici

Fonti

Tecnologia e Infrastruttura...
Fattori Organizzativi e di ...
Fornitori e Terze Parti
Errori o comportamenti um...

Misure

Formazione e Sensibilizzaz...
Sicurezza dei canali inform...
Controlli di Accesso e Aut...
Crittografia e Pseudonimizz...
Audit Log e Tracciabilità d...
Valutazioni di Sicurezza e ...
Valutazione dei Fornitori e ...
Lotta contro il malware
Partizionamento
Minimizzazione dei dati
Controllo delle Versioni e ...

Accesso illegittimo ai dati

Gravità : Limitata

Probabilità : Limitata

Modifiche indesiderate dei dati

Gravità : Trascurabile

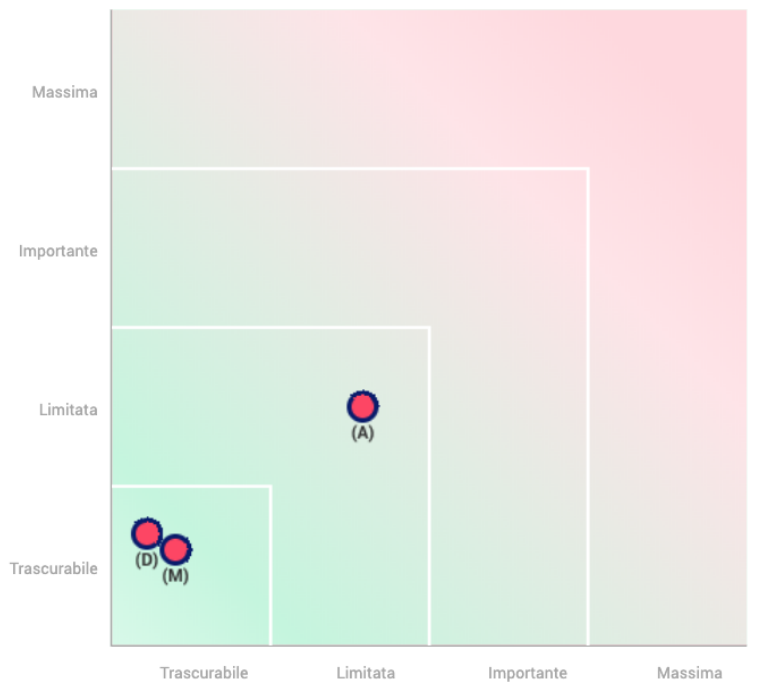
Probabilità : Trascurabile

Perdita di dati

Gravità : Trascurabile

Probabilità : Trascurabile

Gravità del rischio



Probabilità del rischio

- Misure pianificate o esistenti
- Con le misure correttive implementate
- (A)ccesso illegittimo ai dati
- (M)odifiche indesiderate dei dati
- (P)erdita di dati

18/01/24

10.6 Indicare le salvaguardie previste se i dati personali sono trasferiti o comunicati a paesi terzi

I dati non saranno comunicati o trasferiti a paesi terzi

10.7 Eventuale coinvolgimento del DP

La presente DPIA è stata sottoposta all'attenzione del DPO.

Il DPO ha espresso parere sul documento.

11 REVISIONE E AGGIORNAMENTO, CON RIESAME DI CONGRUITÀ CON LE ESIGENZE DI PROTEZIONE DEI DATI - ART 35 GDPR

Il riesame di congruità è condizionato a eventuali modifiche organizzative e normative che possono determinarsi nel corso dello studio.

11.1 Data entro la quale deve essere condotto il riesame di congruità

La revisione della DPIA verrà effettuata tempestivamente nei casi specifici in cui è necessaria un'azione correttiva o di miglioramento delle modalità di trattamento.

Il riesame sarà effettuato di norma con cadenza annuale, tale periodo può variare in rapporto alle eventuali criticità rilevate durante le fasi del trattamento o a seguito di interventi normativi, regolatori, audit.

Anche in questa fase verrà chiesto il parere del DPO.

12 APPENDICE A - Lista di controllo della congruità del trattamento previsto con le esigenze di protezione dei dati

Domanda	Risposta
1. Quali tipologie di dati personali vengono trattate?	Dati personali (età, sesso ecc...), dati clinici (diagnosi, trattamenti sanitari, trattamenti farmacologici, esiti di esami di laboratorio, esiti di visite cliniche Specialistiche, ecc...), dati genetici
2. Se vengono trattati speciali categorie di dati, elencati all'articolo 9 comma 1 GDPR, sulla base di quanto illustrato nella DPIA, esiste una motivazione legittima per il trattamento?	Sì, quale parte fondamentale del processo di conduzione dell'attività di ricerca.
3. Vi sono aspetti afferenti al rispetto dell'articolo 2, comma 2, del regolamento, che protegge i diritti fondamentali e le libertà delle persone fisiche, ed in particolare il loro diritto alla protezione dei dati personali, che non siano trattati in questa DPIA?	NO
4. Tutti i dati personali che verranno trattati sono coperti da garanzie di riservatezza? Se sì, come viene garantita?	I dati vengono trattati solo da soggetti autorizzati inoltre ogni caso è pseudonimizzato o anonimizzato.
5. Come viene offerta agli interessati l'informativa in merito al fatto che i loro dati personali verranno raccolti e trattati?	L'informativa viene inoltre pubblicata sul sito regina.cnr.it
6. Il trattamento dei dati comporta l'utilizzo di dati personali già raccolti, che verranno utilizzati per finalità secondarie?	Trattandosi di uno studio retrospettivo, saranno utilizzati dati già acquisiti dall'ente. Il trattamento è compiutamente descritto nell'informativa degli studi da cui provengono i dati.
7. Quali procedure vengono adottate per verificare che le modalità di raccolta dei dati sono adeguate, coerenti e non eccessive, in relazione alle finalità per i quali i dati vengono trattati?	Al fine di ridurre il rischio per i diritti e le libertà fondamentali degli interessati è prevista la piena applicazione dei principi affermati dall'art. 5 GDPR. Nel dettaglio, il principio di minimizzazione, di pertinenza, di esattezza, ecc...
8. Con quali modalità viene verificata l'accuratezza dei dati personali raccolti e trattati?	L'accuratezza dei dati personali raccolti viene garantita mediante gli identificatori utilizzati per l'estrazione: nome, cognome, data di nascita e codice fiscale dei pazienti inclusi nello studio.
9. È stata effettuata una valutazione circa il fatto che il trattamento dei dati personali raccolti	Sì, tuttavia, in ragione delle misure di sicurezza adottate, non si ritiene che il trattamento effettuato, nelle varie fasi di studio, possa

potrebbe causare danni ai diritti e alle libertà agli interessati coinvolti?	causare danni ai diritti e alle libertà agli interessati coinvolti.
10. È stato stabilito un periodo massimo di conservazione dei dati?	I dati saranno conservati per un massimo di 10 anni dalla conclusione dello studio, dopo di che verranno eliminati o anonimizzati mediante la rimozione definitiva e irreversibile dell'associazione tra il codice assegnato ai dati e l'identità del partecipante.
11. Quali misure tecniche e organizzative di sicurezza sono state adottate per prevenire qualsivoglia trattamento di dati personali non autorizzato o illegittimo?	Accesso ai dati riservato – Solo personale avente diritto nell'ambito delle funzioni a cui è normalmente preposto accederà ai dati al fine di estrazione. Ognuna delle persone coinvolte nell'estrazione dei dati avrà accesso esclusivamente al sistema a cui è normalmente preposto. Pseudonimizzazione – il data set in input, a seguito del processo di pseudonimizzazione restituirà in output il data set pseudonimizzato o anonimizzato, modificato rispetto al dato originale, e il data set di transcodifica depositato su repository aziendale ad accesso riservato mediante password. Gestione postazioni: le postazioni utilizzate sono principalmente in dominio aziendale e le misure adottate sono quelle previste da regolamenti e policy aziendali. Controllo degli accessi fisici: l'accesso ai locali è consentito al solo personale che abbia necessità di accedere a dispositivi o attrezzature necessarie al trattamento, conservate in tali locali. Sicurezza dei documenti cartacei: in base alle istruzioni generali impartite dal Titolare, ogni singolo soggetto coinvolto nel trattamento della ricerca, condivide la documentazione prodotta con i soli appartenenti all'equipe di ricerca. Normalmente la documentazione cartacea riguarda i dati del progetto (es. Protocollo, parere CE AVEC) poiché i dati personali relativi ai soggetti coinvolti vengono trattati principalmente in modalità informatizzata. Sicurezza dei canali informatici: tutte le postazioni e i dispositivi aziendali/istituto sono equipaggiati con strumenti di antispam, antivirus, sistemi di monitoraggio degli apparati fisici di rete e server e gestione degli alert.

	I canali informatici, in particolare quelli comunicanti con l'esterno, sono criptati tramite chiavi asimmetriche (ED25519 o equivalenti) specifiche per ogni utente, con l'aggiunta di ulteriori chiavi asimmetriche utilizzate la crittazione dei dati di studio Gestione delle politiche di tutela della privacy: Il Titolare ha adottato Privacy Policy aziendali periodicamente revisionate, disponibili su sito web aziendale e condivise con tutto il personale. Il personale viene adeguatamente formato in merito alle attività di trattamento e alle misure di sicurezza da adottare.
12. È previsto il trasferimento di dati personali in un paese non facente parte dell'Unione europea? Se sì, quali provvedimenti sono stati adottati per garantire che i dati siano salvaguardati in modo appropriato?	NO