IRCCS "Giovanni Paolo II"
PugliaSalute

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

VALUTAZIONE DI IMPATTO SULLA PROTEZIONE DEI DATI STUDI RETROSPETTIVI

Codice	Descrizione
DPIA-001	**Studio retrospettivo pazienti affette da carcinoma della mammella sottoposte a terapia adiuvante: sviluppo di sistemi di supporto alle decisioni terapeutiche**
ELABORAZIONE DPIA PER	☒ **Nuova attività trattamento** ☐ Aggiornamento DPIA ☐ Revisione periodica DPIA

Attività	Struttura/Funzione	Responsabile	data	firma
Redazione	**Principal Investigator**	Raffaella Massafra		
Verifica	**DPO**	Iris Mannarini		
Approvazione	**Direttore Generale**	Alessandro Delle Donne		

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

SOGGETTI COINVOLTI NELLO STUDIO	
TITOLARE promotore	**IRCCS ISTITUTO TUMORI GIOVANNI PAOLO II DI BARI ((Principal Investigator: Dott.ssa Raffaella Massafra)**
Centri partecipanti quali Titolari del trattamento	**Non presenti**
RESPONSABILE DEL TRATTAMENTO	**Non presente**
COORDINATORE E SPERIMENTATORI	**All'interno della steering committee sono da considerarsi:** Dott.ssa Raffaella Massafra, SSD Fisica Sanitaria, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott.ssa Annarita Fanizzi, SSD Fisica Sanitaria, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott. Vito Lorusso, U.O.C Oncologia Medica, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott. Francesco Giotta, U.O.C Oncologia Medica, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott. Agnese Latorre, U.O.C Oncologia Medica, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott. Alfredo Zito, U.O.C. Anatomia Patologica, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott. Daniele La Forgia, U.OC. Radiologia Senologica, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott.ssa Samantha Bove, SSD Fisica Sanitaria, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari Dott.ssa Maria Colomba Comes, SSD Fisica Sanitaria, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

FASI DPIA	☒	**Conduzione DPIA**	
	☐	Parere del DPO	
	☐	Validazione del Titolare	
	☐	Consultazione Preventiva	
	☐	Revisione DPIA	
MODALITA' CONDUZIONE	☒	DPIA OBBLIGATORIA	
	☐	DPIA VOLONTARIA	

IRCCS "Giovanni Paolo II"
PugliaSalute

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

Sommario

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

Informazioni sulla DPIA

La DPIA, acronimo di *Data Protection Impact Assessment*, è una valutazione preliminare, eseguita dal Titolare del trattamento dei dati personali, relativa agli impatti a cui andrebbe incontro un trattamento laddove dovessero essere violate le misure di protezione dei dati. In linea con l'approccio basato sul rischio adottato dal Regolamento generale sulla protezione dei dati, è necessario realizzare una valutazione d'impatto sulla protezione dei dati soltanto quando la tipologia di trattamento" può presentare un rischio elevato per i diritti e le libertà delle persone fisiche" (articolo 35 del Regolamento 2016/679).

Ai sensi dell'articolo 35, paragrafo 3 del Regolamento 2016/679 la valutazione è effettuata nei casi in cui un trattamento può presentare rischi elevati, ossia quando:
a. una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche;
b. il trattamento, su larga scala, di categorie particolari di dati personali di cui all'articolo 9, paragrafo 1, o di dati relativi a condanne penali e a reati di cui all'articolo 10;
c. la sorveglianza sistematica su larga scala di una zona accessibile al pubblico.

In particolare, preso atto della tipologia di Studio (retrospettivo) in argomento, è condotta e resa pubblica una valutazione d'impatto sulla protezione dei dati, in forza dell'art. 35 del Reg. UE 2016/679 riguardo al trattamento dei dati ai sensi e per gli effetti del combinato disposto degli artt. 9, par. 2, lett. j) del GDPR, 110 e 110 bis, comma 4 del Codice Privacy.

La presente valutazione contiene:
a) una descrizione sistematica dei trattamenti previsti e delle finalità del trattamento

b) una valutazione della necessità e proporzionalità dei trattamenti in relazione alle finalità;
c) una valutazione dei rischi per i diritti e le libertà degli interessati;
d) le misure previste per affrontare i rischi, includendo le garanzie, le misure di sicurezza e i meccanismi per garantire la protezione dei dati personali e dimostrare la conformità al presente regolamento, tenuto conto dei diritti e degli interessi legittimi degli interessati e delle altre persone in questione

ACCETTABILITA' DEL RISCHIO

Tenuto conto della natura, del contesto, delle finalità e dell'ambito di applicazione del trattamento in esame, il **livello di rischio residuo**, considerato accettabile indicato dal Titolare, sentito anche il parere del DPO, è risultato ☒ BASSO ☐ MEDIO ☐ ALTO

Di seguito sono illustrati i dettagli della valutazione d'impatto sulla protezione dei dati.

1 Descrizione sistematica del trattamento

1.1 Contesto

Il cancro al seno è una delle principali cause di morte in tutto il mondo. Per questo motivo, la gestione delle pazienti con carcinoma mammario è da sempre un argomento di grande interesse all'interno della comunità scientifica. La chirurgia di vario tipo seguita da terapie adiuvanti, quali chemioterapia, ormonoterapia, radioterapia da sola o in associazione, rappresenta il trattamento primario del carcinoma mammario. Nonostante siano stati compiuti grandi progressi per migliorare la sopravvivenza del cancro, la scelta di quale terapia adiuvante deve essere eseguita per prevenire l'insorgenza di eventi di malattia dopo il tumore primario, come recidiva, metastasi, controlaterale e secondo tumore, rimane ancora impegnativo. Gli esperti clinici prendono le loro decisioni per ogni singolo paziente seguendo le linee guida pertinenti, dopo aver raccolto e valutato una serie di misurazioni di parametri clinici e istologici. Finora, diversi lavori di ricerca hanno dimostrato il ruolo centrale dei sottotipi di carcinoma mammario sia sulla prognosi del tumore che sull'efficacia della terapia. A seconda del sottotipo individuato, possono essere selezionati i pazienti eleggibili per uno specifico trattamento adiuvante dopo l'intervento chirurgico, risparmiando così alcuni altri pazienti da trattamenti non necessari e/o potenzialmente tossici. Tuttavia, a causa dell'eterogeneità molecolare di questa malattia, è molto difficile prevedere l'esito e l'efficacia della terapia adiuvante su misura per ogni singolo paziente. All'interno di questo scenario emergente, è urgente la necessità di modelli predittivi che facciano un compromesso tra previsioni affidabili dei risultati della terapia e rapporto costo-efficacia. Negli ultimi anni, grazie ai grandi progressi nel campo dell'intelligenza artificiale applicata alla biomedicina, la progettazione e lo sviluppo di modelli di apprendimento automatico per supportare i processi decisionali clinici nel trattamento del cancro al seno, sono stati ampiamente studiati nello stato dell'arte arte.

1.2 Panoramica del trattamento

1.2.1 Quale è il trattamento in considerazione?

Lo studio non prevede alcuna modifica dell'iter diagnostico e terapeutico del paziente. Lo studio prevede l'arruolamento di pazienti trattati secondo pratica clinica. Al fine del protocollo, non viene richiesta nessuna procedura aggiuntiva rispetto alla pratica clinica, né ulteriori valutazioni diagnostiche o prelievi di campioni biologici. Anche la gestione dei pazienti, il follow-up e gli eventuali trattamenti rispecchieranno la normale pratica clinica.

Obiettivo dell'attività di ricerca è evidenziare in retrospettivo le relazioni latenti tra le caratteristiche, cliniche istologiche e terapeutiche e il follow-up delle pazienti sottoposte a terapia neoadiuvante. Inoltre, modelli di supporto alle decisioni saranno implementati mediante tecniche di intelligenza artificiale.
I dati clinici retrospettivi, riferiti a pazienti con carcinoma mammario sottoposti a terapia adiuvante, saranno raccolti direttamente dalle cartelle cliniche. I dati clinici pre-trattamento includeranno l'età alla diagnosi, il grado istologico, il tipo istologico, l'espressione del recettore degli estrogeni (ER), l'espressione del recettore del progesterone (PgR), il marker cellulare di proliferazione (Ki67), il punteggio del recettore 2 del fattore di crescita epidermico umano (HER2-neu), la multifocalità del tumore, lo stato clinico dei linfonodi ascellari, lo stato BRCA1/2. Altre informazioni clinico-patologiche disponibili in cartella cliniche e ritenute di interesse da parte dei clinici in corso di studio saranno altresì collezionate. Saranno inclusi altri dati relativi al tipo di intervento chirurgico eseguito e allo schema farmacologico seguito (ad es. Antracicline, trastuzumab, pertuzumab, taxano). Delle stesse pazienti sarà collezionato il follow-up alle successive rivalutazioni di controllo.
Sanno implementati studi di associazione e di sopravvivenza sia mediante tecniche allo stato dell'arte che mediante approcci di machine learning al fine di evidenziare caratteristiche significative del fenomeno di evoluzione della malattia dopo primo tumore della mammella in

pazienti sottoposte a chemioterapia neoadiuvante. Saranno utilizzati opportuni test e modelli statistici parametri e non parametrici per la valutazione di associazioni univariate e multivariate tra le caratteristiche dei pazienti osservati e del relativo percorso di cura, rispetto al follow-up di trattamento. Inoltre, sulla scorta del dominio di conoscenza acquisito, verranno sviluppati algoritmi di machine learning e sistemi di intelligenza artificiale. L'intelligenza artificiale è la scienza che sviluppa l'architettura necessaria affinché le macchine simulino la capacità di ragionamento all'essere umano, come ad esempio, le reti neurali, sistemi informatici che cercano di simulare le reti neuronali biologiche. Nell'ambito degli studi di pattern recognition e teoria dell'apprendimento computazionale gli algoritmi di machine learning, invece, sono metodi di analisi dei dati che mirano ad automatizzare la creazione di modelli statistici analitici e consentono alle macchine intelligenti di migliorarsi con il tempo, esattamente come avviene con il cervello umano. Speciali algoritmi sono addestrati iterativamente su una base dati fornita dall'uomo e 'imparano' a prendere decisioni ed effettuare predizioni senza ulteriori interventi dell'uomo. Tali modelli saranno addestrati sulle informazioni di varia natura estratte dalla storia clinica e dati anagrafici dei pazienti che hanno concluso un percorso e per i quali pertanto è noto il follow-up terapeutico, diagnostico e operatorio. I ricercatori analizzeranno il flusso informativo in input (percorsi e follow-up) e l'attendibilità dei modelli predittivi generati, al fine di individuare eventuali criticità e suggerire eventuali correttivi.

L'associazione tra le variabili cliniche e gli esiti richiesti sarà valutata attraverso l'implementazione di test statistici appropriati. L'integrazione delle informazioni raccolte e il loro sfruttamento pratico come supporto ai processi decisionali sarà effettuata tramite modelli statistici multivariati, algoritmi di intelligenza artificiale inclusi tecniche di apprendimento automatico, in particolare classificatori standard (rispettivamente, Random Forest, Support Vector Machine) o Artificial Neural Networks.

Le tecniche di selezione delle caratteristiche saranno annidate all'interno dei modelli di classificazione per identificare le caratteristiche più significative per le attività di previsione. Le metriche standard, come accuratezza, sensibilità e specificità, nonché l'AUC verranno calcolate per valutare le prestazioni del modello utilizzando schemi di convalida incrociata e test

indipendenti. Per ciascun paziente individualmente, il modello di AI implementato effettuerà una scelta in accordo con l'esito da prevedere. Tale scelta sarà chiarita da metodi di intelligenza artificiale spiegabili, come SHapley Additive exPlanations (SHAP), basati sul calcolo dei valori di Shapley, e Local Interpretable Model-agnostic Explanations (LIME). In sostanza, ogni paziente sarà rappresentato da un vettore di importanza, dove ogni elemento rappresenterà il peso di come ciascuna caratteristica, se considerata da sola o in collaborazione con tutte le altre caratteristiche considerate, contribuirà alla scelta finale del modello di AI. Al contrario dell'output di una tecnica di selezione delle caratteristiche, ovvero un unico vettore di caratteristiche importanti stimate sul training set e utilizzato per tutti i pazienti del set di test, i modelli XAI restituiscono più vettori di caratteristiche importanti, uno per ogni paziente, che possono variano da paziente a paziente.
Le numerosità considerate, sulla base dei risultati degli studi clinici riportati in letteratura, sono sufficienti per garantire il raggiungimento della potenza statistica.
Considerando i risultati di una prima ricognizione sui dati dei pazienti seguiti dalla Breast Unit del nostro Istituto, al fine di rendere più robuste le analisi è possibile aumentare la numerosità campionaria a 1000 pazienti con i requisiti sopra descritti.

Principali criteri di inclusione:

Pazienti affetti da carcinoma della mammella che:

- abbiano effettuato terapia adiuvante,
- non siano metastatiche ad-inizio,
- un follow-up di almeno 5 anni.

Principali criteri di esclusione:

Pazienti affetti da carcinoma della mammella che:

- hanno effettuato terapia neoadiuvante,
- metastatiche ab-inizio,
- con meno di 5 anni di follow-up.

Potranno anche essere arruolati pazienti deceduti al momento dell'arruolamento nel presente studio, purché vengano rispettati i criteri di inclusione.

I dati verranno anonimizzati e raccolti su un database specificamente preparato. Quindi verrà effettuata la valutazione statistica dei risultati.

I risultati di questo studio, se di interesse, potranno costituire la base per uno studio di validazione di più grandi dimensioni e multicentrico. In tal caso, sarà predisposto un data transfer agreement e tutti i dati raccolti saranno resi anonimi prima del trasferimento tra più istituzioni.
Si sottolinea che, lo scopo dello studio è attualmente lo "sviluppo di un algoritmo predittivo" e non prevede alcuna implicazione nella pratica clinica.
Ciò nonostante, in linea con i principi di conoscibilità, non esclusività e non discriminazione algoritmica richiamati da ultimo anche dal Garante Privacy, in materia di utilizzo di algoritmi di intelligenza artificiale, si sottolinea quanto segue: (1) tutti i risultati della ricerca e i dettagli tecnici dei modelli implementati saranno resi pubblici mediante pubblicazioni scientifica e divulgazione in conferenze nazionali ed internazionali, (2) saranno collezionate anche informazioni riferite a soggetti deceduti e non contattabili in assenza delle quali il campione selezionato sarebbe incompleto creando di conseguenza, come rappresentato, possibili bias nello sviluppo dell'algoritmo, (3) fermi restando i vantaggi offerti dai sistemi automatizzati di computazione dei dati che si intendono implementare, tali operazioni non sono unicamente delegate ad elaborazioni automatizzate, ma al contrario, necessitano di una integrazione con l'intervento umano che, sulla base delle competenze ed expertise tecnico-specifiche, monitorano, correggono o mitigano le operazioni effettuate mediante algoritmi automatizzati. Pertanto, le tecniche di intelligenza artificiale e di apprendimento automatizzato utilizzate nell'ambito delle attività di elaborazione dello Studio non portano a un processo decisionale automatizzato.

Tipologia di Studio

Trattasi di studio osservazionale retrospettivo monocentrico.

1.2.2 Quali sono le responsabilità connesse al trattamento?

Gli sperimentatori coinvolti nello Studio sono appositamente autorizzati al trattamento dei dati, ai sensi dell'art. 29 del Reg. UE 2016/679 e dell'art. 2 quaterdecies del Dlgs 196/2003, così come novellato dal D.lgs 101/2018. Nell'ambito dello Studio non risultano designati soggetti terzi in qualità di Responsabili del trattamento dati ai sensi dell'art. 28 del GDPR.

1.2.3 Ci sono standard applicabili al trattamento?

- La linea guida di Buona Pratica Clinica [Good Clinical Practice (GCP)] è uno standard internazionale di etica e qualità scientifica per progettare, condurre, registrare e relazionare gli studi clinici che coinvolgono soggetti umani. La GCP ha l'obiettivo di fornire, in conformità con i principi per la tutela dei diritti dell'uomo stabiliti dalla Dichiarazione di Helsinki, uno standard comune ad Unione Europea, Giappone e Stati Uniti per facilitare la mutua accettazione dei dati clinici da parte delle autorità regolatorie di queste aree geografiche;

- La linea guida recepita dall'Italia (G.U.R.I. n.191 del 18 agosto 1997) è stata messa a punto sulla base delle GCP attualmente adottate da Unione Europea, Giappone e Stati Uniti, oltre che da Australia, Canada, Paesi Nordici e dall'Organizzazione Mondiale della Sanità (OMS);

- Il trattamento di dati personali per scopi di ricerca scientifica è effettuato nel rispetto del Regolamento UE 2016/679, del Codice, delle Prescrizioni relative al trattamento dei dati genetici e delle Prescrizioni relative al trattamento dei dati personali effettuato per scopi di ricerca scientifica, nonché delle Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica, che costituiscono condizione essenziale di liceità e correttezza dei trattamenti.

1.3 Dati, processi e risorse di supporto

1.3.1 Quali sono i dati trattati e gli asset a supporto?

Tipologia di dati personali	Categoria interessati

☒ **Dati identificativi comuni** (es. nome, cognome,indirizzo) ☒ **Dati di contatto** (recapiti email, telefono, cellulare, etc.) ☒ **Dati sanitari già presenti negli archivi** ☒ Dati raccolti da archivi cartacei ☒ Dati raccolti da archivi informatici ☐ Credenziali di autenticazioni, chiavi di accesso ☐ Dati raccolti da strumenti audiovisivi, videosorveglianza ☐ Dati raccolti da tecnologie traccianti e/o di monitoraggio ☐ Dati raccolti da tecnologie IoT ☐ Dati su abitudini di vita, consumi e comportamento ☐ Dati su familiari/stato familiari ☐ Dati bancari ☐ Dati sulla localizzazione ☐ Dati sulla solvibilità economica	• Pazienti deceduti o non reperibili • Pazienti in vita (in follow-up presso il nostro Istituto)
☐ Appartenenza sindacale ☐ Convinzioni politiche, religiose o filosofiche ☐ Origine razziale o etnica ☒ **Dati sulla salute** ☐ Orientamento e vita sessuale ☐ Dati genetici ☐ Dati biometrici ☐ Dati "giudiziari" (diritto penale)	
☐ dati soggetti a maggior tutela	

Altro: ……	

COMPONENTI ORGANIZZATIVE	
Soggetti interni	Lo staff dello studio è composto dal Principal Investigator, ricercatori e data manager opportunamente individuati in fase di sottomissione dello studio e nel corso dello stesso. Al Principal Investigator viene conferita la delega per la gestione delle attività di trattamento dei dati personali per i compiti relativi alla protezione dei dati personali necessari per la conduzione dello studio. Gli altri componenti dello staff sono autorizzati al trattamento di dati personali da parte del P.I. tramite apposito atto di nomina individuale.
Soggetti esterni	Non applicabile.
COMPONENTI TECNOLOGICHE	
Applicazioni	Per l'elaborazione dei dati sono utilizzati sistemi di office automation quali RedCap, Microsoft Word, Excel, Python (versione >3.9.2), Matlab (versione >2023a), utilizzati esclusivamente presso il Centro Promotore (IRCCS Istituto Tumori di Bari)
Infrastrutture ICT	Per la conservazione dei dati in formato elettronico sono utilizzati i sistemi di storage aziendali opportunamente protetti sia per quanto riguarda l'accesso fisico che l'accesso ai database che sono opportunamente criptati secondo le regole tecniche usuali e politiche di backup specifiche.
Reti informatiche	I computer utilizzati per il trattamento dei dati si trovano su rete dedicata e messa in sicurezza su apposita VLAN.
COMPONENTI FISICHE	
Asset	Per l'elaborazione dei dati sono utilizzati appositi strumenti software aziendali. I PC su cui sono installati tali software sono muniti di idonei sistemi di autenticazione, autorizzazione e tracciabilità delle operazioni
Sedi	Il trattamento dei dati avviene attraverso postazioni di lavoro presso la sede aziendale della ricerca scientifica con accesso riservato
Archivi	I dati personali sono conservati in sicurezza presso l'archivio corrente aziendale o su IaaS (Infrastructure as a Service) in cloud opportunamente protetta.

1.4 Finalità del trattamento

Il trattamento dei dati personali identificativi risulta necessario per le seguenti finalità dello Studio:

- Studi di sopravvivenza
- Analisi statistiche di natura osservazionale finalizzati all'individuazione di evidenze statistiche in relazione all'esito di trattamento.
- Sviluppo di sistemi di supporto alle decisioni terapeutiche per la previsione della ricorrenza di malattia.

2 Principi Fondamentali

2.1 Valutazione della necessità e proporzionalità del trattamento del trattamento

Il trattamento è effettuato nel rispetto delle prescrizioni previste dall'art. 5 del GDPR e pertanto saranno trattati secondo i principi di:

1. liceità, correttezza e trasparenza
2. limitazione della finalità
3. minimizzazione dei dati
4. esattezza
5. limitazione della conservazione
6. integrità e riservatezza

Lo Studio in argomento comporta il trattamento di dati personali riconducibili allo stato di salute degli assistiti in cura presso l'IRCCS, secondo i criteri di inclusione dello Studio.

2.1.1 Gli scopi del trattamento sono specifici, espliciti e legittimi?

Il trattamento correlato allo Studio è effettuato nel rispetto del principio di liceità e trasparenza. A tal proposito è stata predisposta e pubblicata sul sito internet istituzionale, l'informativa Privacy sugli studi retrospettivi. Lo scopo dello Studio è esplicito ed è descritto dettagliatamente nella

documentazione di presentazione del medesimo Studio approvato dal Comitato Etico competente.

2.1.2 Quale è il flusso dei dati durante il ciclo di vita del trattamento?

Il ciclo di vita del dato ha origine dall'acquisizione dei dati relativi alla salute dalla documentazione sanitaria e archivi presenti presso le Unità Operative dell'IRCCS ai sensi dell'art. 110Bis, 4 comma, Cod. Privacy.

Su tali dati verranno effettuate le attività delle elaborazioni statistiche peculiari dello Studio.

I dati clinici sono introdotti su sistema informativo di raccolta delle eCRF "RedCAP" ospitato sulla IAAS di InnovaPuglia (Cloud Sanità) che espone la propria interfaccia web su Internet con protocollo https e separa la Web Application dal server database.

La Web Application è sita nel layer DMZ, non raccoglie dati, ma consente l'accesso agli utenti autorizzati ed è protetta dalla rete internet esterna tramite firewall. Il database che colleziona i dati è accessibile esclusivamente dalla DMZ, la connessione è protetta da un apposito firewall.

I dati relativi all'identità del paziente sono sottoposti a pseudonimizzazione eseguita secondo una delle due procedure descritte nel paragrafo relativo alle misure tecniche.

2.1.3 Quali sono le basi legali che rendono lecito il trattamento?

Basi giuridiche del trattamento di dati.

Paziente in vita e rintracciabile

Art. 6 par. 1, lett. A) e Art. 9 par. 2 lett. a) del GDPR (**acquisizione del consenso**).

Pazienti deceduti o non rintracciabili

Art. 9 par. 2 lett. j) del GDPR e artt. 110-110 bis c. 4 del d.lgs 196/03 e Aut. Gen. 9/2016 e ss aggiornamenti:

Il trattamento è necessario a fini di ricerca scientifica in campo medico, biomedico o epidemiologico, effettuata in base a disposizioni di legge o di regolamento o al diritto dell'Unione europea, ed è condotta e resa pubblica una valutazione d'impatto sulla protezione dei dati.

L'interessato è qualificato come non rintracciabile dopo almeno 3 tentativi (tracciati) di contatto non riusciti.

L'interessato deceduto viene rilevato dalla CC (in caso di decesso durante il periodo di degenza) o dal sistema TS (tessera sanitaria).

Ulteriori garanzie:
art. 8, comma 5-bis del d.lgs. n. 288 del 2003

2.1.4 I dati sono esatti e aggiornati?

I dati personali sono acquisiti dagli archivi aziendali con ulteriori controlli interni in caso di omonimie o omocodie.

2.1.5 Qual è il periodo di conservazione dei dati?

Tipologia di dati personali	Tempi di conservazione
Dati pseudonimizzati	I dati pseudonimizzati saranno conservati per la durata dello Studio, decorsi i quali saranno anonimizzati.
Tabella di corrispondenza Codice ID/Paziente	La tabella di corrispondenza sarà cancellata in modalità permanente al termine della durata dello Studio in parola.

2.2 Misure a tutela dei diritti degli interessati

2.2.1 Come sono informati del trattamento gli interessati?

Con riferimento ai pazienti viventi, saranno rese le informazioni sul trattamento dei dati **ai sensi dell'art. 13** del Reg. UE 2016/679, nella fase di arruolamento.
A beneficio dei pazienti deceduti (o per quelli irreperibili) sono pubblicate nell'apposita sezione del sito internet istituzionale, le informazioni sul trattamento dei dati, **ai sensi dell'art. 14, par. 5, lett. b)** del Reg. UE 2016/679. È altresì pubblicato l'informativa al trattamento dei dati personali con relativa valutazione d'impatto.

2.2.2 Ove applicabile: come si ottiene il consenso degli interessati?

Paziente in vita e rintracciabile
Art. 9 par. 2 lett. a) del GDPR
Il consenso, ove possibile, è raccolto dal paziente in fase di arruolamento, tramite l'acquisizione di firma autografa su modello "Consenso Informato Privacy specifico della Ricerca rev 3.0 del 30 ottobre 2024".

Paziente deceduti/non rintracciabili

Per tale categoria di pazienti, il consenso non può essere raccolto pertanto per il nostro IRCCS ci si avvale dell'art. **110 e 110 bis, comma 4 del Codice Privacy e art. 9, par 2, lett j), GDPR.**

2.2.3 Come fanno gli interessati a esercitare i loro diritti?

I diritti dei pazienti in vita e/o ricontattabili di cui agli artt. 15-22 del GDPR sono sempre garantiti nelle modalità indicate nell'informativa ex artt. 13-14 del GDPR rese al momento dell'arruolamento.
Altresì sono resi disponibili sul sito internet istituzionale (https://www.sanita.puglia.it/web/irccs/privacy1) i modelli da poter utilizzare per l'esercizio di tali diritti.
I diritti di cui agli articoli da 15 a 22 del Reg. UE 2016/679 riferiti ai dati personali concernenti persone decedute possono essere esercitati da chi ha un interesse proprio, o agisce quale avente

diritto o per ragioni familiari meritevoli di protezione. Le informazioni sul trattamento dei dati circa gli Studi condotti in assenza del consenso sono rese pubbliche sul sito internet istituzionale, nell'apposita sezione dedicata alla ricerca scientifica, unitamente alla valutazione di impatto sulla protezione dei dati personali.

2.2.4 In caso di trasferimento di dati al di fuori dell'Unione europea, i dati godono di una protezione equivalente?

I dati personali non saranno trasferiti verso Paesi Terzi extra UE.

Misure esistenti o pianificate per la protezione del dato

• **garanzie** (adozione di tecniche di pseudonimizzazione, minimizzazione, implementazione della privacy by design e by default, previsione di procedure volte a testare, verificare e valutare l'efficacia delle garanzie e misure adottate).
• **misure di sicurezza organizzative** (es: norme e procedure che disciplinano l'aspetto organizzativo della sicurezza)
• **misure di sicurezze fisiche** (es: misure di protezione di aree, apparecchiature, dati)
• **misure di sicurezza logiche** (backup, piano di continuità operativa, piano di disaster recovery) sia in relazione al corretto utilizzo degli strumenti elettronici, sia in relazione alla loro gestione e manutenzione

Di seguito le principali misure tecniche applicate, ai sensi dell'art. 32 del Reg. UE 2016/679:

- Endpoint protection: Antivirus e *firewall* sulle singole postazioni di lavoro costantemente aggiornati mediante server ed associazioni a dominio. L'IRCCS ha acquisito un sistema di sicurezza integrato che comprende la gestione del *firewall* e del SOC.
  Per il monitoraggio e il controllo della rete viene utilizzato lo strumento Manage Engine Central che individua, gestisce e tiene traccia delle risorse poste sotto il dominio. Il SOC funge da primo soccorso in caso di incidente di sicurezza. Si possono eseguire operazioni come: isolare gli *endpoint,* terminare i processi dannosi, impedire l'esecuzione di processi dannosi ed eliminare i *files*.
- Implementazione di un Piano Operativo del servizio di sicurezza;
- Adozione del *cloud* di Regione Puglia, gestito dalla società *in-house* Innovapuglia, come *cloud* aziendale. Per i dati migrati sui menzionati *cloud* sono garantiti ridondanza dei dati e *backup*.

- Accesso alla postazione di lavoro mediante password a dominio aggiornata secondo i criteri di sicurezza adeguati al trattamento dei dati sensibili.
- Collocazione del database su postazione di lavoro isolata dalla rete. L'accesso ai locali come i centri di controllo è consentito solo al personale di manutenzione che detiene le chiavi di accesso. In tema di sicurezza fisica, viene altresì garantita la continuità elettrica da parte dei sistemi UPS, la refrigerazione attraverso impianti centralizzati e sistemi SPLIT local.
- Database criptato e protetto da password adeguato al trattamento dei dati sensibili
- Tecniche di pseudonimizzazione dell'identità dei pazienti realizzate alla volta con:
    1. tabella fisica di associazione pseudonimo/identità custodita in armadio a chiave dal PI e solo da questi accessibile
    2. esecuzione di algoritmi di hashing non reversibili a chiave
- Registrazione dei log di accesso al server applicativo e *database*. È altresì prevista la possibilità di verificare i log dall'Event Viewer di ciascuna postazione degli utenti abilitati come Amministratori di Sistema.
- Aggiornamento costante dei sistemi operativi e dei software di sistema e di ambiente. Predisposizione di un *asset inventory* tecnologico attraverso lo strumento di Manage Engine Central che individua, gestisce e tiene traccia delle risorse poste sotto dominio.
- Backup quotidiano della base dei dati su supporto ottico custodito separatamente in armadio ad accesso fisico ad uso esclusivo del PI. È presente un data center virtuale con servizi Backup As A Service presso Innovapuglia e PSN.
- Utilizzo di utenze nominative
- Meccanismi di identificazione ed autenticazione degli utenti
- Classificazione strutturata delle informazioni che tenga conto delle informazioni riservate/contenenti particolari categorie di dati *ex* art. 9 GDPR, attraverso sw dotati di certificazione di sicurezza.
- Password Policy adeguate al trattamento dei dati sensibili. L'ente ha sviluppato una policy sull'assegnazione delle password e che prescriva come tutte le macchine sotto dominio dell'ente richiedano periodicamente l'aggiornamento delle password.
- Erogazione di contenuti formativi per i dipendenti dell'ente che operano nel campo della ricerca.

Misure di sicurezza specifiche per campioni biologici:
Non applicabile.

3 Rischi

3.1 Panoramica dei rischi per diritti e libertà

Il processo di ***valutazione del rischio*** parte dalla determinazione dell'impatto sull'interessato (cioè sulla persona fisica a cui il dato si riferisce) in caso di distruzione, perdita, modifica, divulgazione non autorizzata o altri avvenimenti negativi che possono compromettere la sicurezza del trattamento.

L'impatto derivante dalla perdita di una o più delle caratteristiche della sicurezza delle informazioni, ossia riservatezza, integrità e disponibilità, rappresenta la gravità del danno diretto o indiretto causato agli interessati.

Nel valutare i rischi per le libertà e diritti degli interessati, però, come suggerisce la norma ISO/IEC 29134 si dovrebbero considerare anche altri aspetti, oltre alla sicurezza dei dati; e che pertanto devono essere considerati gli effetti complessivi del trattamento.

I rischi pertanto sono identificati in base ai seguenti quattro parametri:

1) conformità ai principi applicabili al trattamento dei dati (art. 5 del Reg. UE 2016/679)
2) riservatezza
3) integrità
4) disponibilità.

A tal fine, nella determinazione del livello di impatto sono incluse valutazioni sulle possibili conseguenze derivanti da mancanza di trasparenza, mancato rispetto dei tempi di conservazione dei dati, o dalla violazione degli altri principi fondamentali applicabili alla protezione dei dati personali.

- **Quali sono le principali minacce che potrebbero concretizzare il rischio?**

Una minaccia potrebbe concretizzarsi solo al momento dell'acquisizione dei dati durante la consultazione della documentazione sanitaria, che però è effettuata da personale esercente la professione sanitaria, tenuta al segreto professionale, ed istruita in materia di protezione dei dati

personali.

- **Quali sono le fonti di rischio?**

Una fonte di rischio potrebbe essere rappresentata dalla tabella di transcodifica che è gestita separatamente e che se sottratta insieme al database centralizzato dello Studio, consentirebbe di risalire allo stato di salute ed alle patologie dei soggetti inclusi nello Studio.

Non si ravvisano rischi per l'assistito in merito alla perdita di disponibilità del dato in quanto, in caso di evento avverso, non saranno compromessi i dati acquisiti e conservati per finalità di diagnosi, assistenza e cura. Anche in caso di perdita di integrità, non saranno compromessi dati acquisiti e conservati per finalità di diagnosi, assistenza e cura, ma solo per la finalità dello Studio.

- **Quali misure fra quelle individuate contribuiscono a mitigare il rischio?**

Oltre alle istruzioni operative fornite agli sperimentatori, è implementato un sistema crittografico sull'archivio centralizzato che prevede crittografia AES 256 bit con 14 round o cicli di elaborazione crittografica.

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

Il rischio residuo calcolato, dopo l'adozione delle misure di sicurezza pianificate e di quelle tecniche generali dell'Ente, visto anche il rispetto del principio di non esclusività della decisione algoritmica in quanto la validazione resta degli specialisti che intervengono nelle fasi di studio (radiologi e ricercatori clinici) anche al fine di correggere eventuali output non conformi, considerata l'applicazione delle misure di sicurezza dirette sul dato come la pseudonimizzazione e cifratura, nonché dell'assenza del trasferimento dei dati personali verso Paesi extra UE è BASSO.

Le fonti di rischio possono essere categorizzate in:

- **Violazioni dei principi applicabili ai trattamenti di dati personali**
- **Minacce alla sicurezza dei trattamenti**
- **Eventi con danni fisici/materiali**
- **Eventi naturali**

- **Perdita o indisponibilità di servizi essenziali**
- **Compromissione di dati e informazioni**
- **Problemi tecnici**
- **Azioni non autorizzate**
- **Compromissione di funzioni / servizi per errori o azioni malevole**

Il livello di rischio è direttamente proporzionale alla probabilità che si verifichino le diverse minacce e alla gravità dell'impatto per gli interessati. Può essere mitigato con l'applicazione delle necessarie misure di mitigazione.

Se l'applicazione delle misure di mitigazione riduce il livello di rischio, fornendo un primo livello di rischio residuo, il governo dei processi e il presidio di controlli efficaci può fornire un ulteriore livello di ponderazione. Ecco perché oltre alle specifiche contromisure, la metodologia utilizzata inserisce, mediante un self assesment, degli obiettivi di controllo specifici per diverse categorie e ambiti, e dei controlli sullo svolgimento del processo di Valutazione di impatto.

Per la data protection si fa riferimento ai controlli della ISO/IEC 29151, estensione di quelli della ISO/IEC 27001 Annex A, a quelli della ISO/IEC 27701:2019 e della ISDP10003:2018.

3.2 Accesso illegittimo ai dati

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

Danno immateriale, perdita dignità, perdita di controllo sui propri dati personali, irritazione, perdita della fiducia nella sanità pubblica, perdita finanziaria

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

Accessi esterni non autorizzati, Uso improprio del software, Corruzione dei dati, Comunicazione illegale dei dati e dei documenti, Uso non autorizzato dei dati, attacco hacker.

3.2.3 Quali sono le fonti di rischio?

Fonti di rischio umane interne, fonti di rischio umane esterne

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

Crittografia AES 256 bit sugli archivi elettronici dello Studio e dei relativi backup, Controllo degli accessi logici, Tracciabilità, Lotta contro il malware, Gestione postazioni, Politica di tutela della privacy, Firewalling, EDR, registrazione dei log di accesso al server mediante applicativo e *database*.

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

Significativa. La gravità del rischio potenziale di accesso illecito ai dati è stimata come ALTA, in considerazione della tipologia di dati raccolti.

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

La probabilità di accadimento è stimata come BASSA in considerazione delle misure di garanzia implementate con particolare riferimento alle tecniche di pseudonimizzazione e crittografia applicate, oltre che a tutte le misure di natura tecnica e organizzativa implementate dall'ente.

3.3 Modifiche indesiderate dei dati

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

Modifiche ai dati raccolti per finalità di ricerca non comportano un impatto diretto all'interessato.

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

Accessi esterni non autorizzati, Azione di virus informativi o di codici malefici, Uso non autorizzato dei dati, Sabotaggio, Alterazione dolosa o colposa dati

3.3.3 Quali sono le fonti di rischio?

Fonti di rischio umane interne, fonti di rischio umane esterne

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

Crittografia, Controllo degli accessi logici, Tracciabilità, Lotta contro il malware, Gestione postazioni, Backup, Manutenzione, Politica di tutela della privacy, Firewalling, EDR, Registrazione dei log di accesso

al server applicativo e *database*.

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

Limitata. La gravità del rischio potenziale di modifica illecita dei dati è stimata come BASSA, in considerazione della presenza di dati originali già raccolti per finalità di diagnosi e cura.

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

La probabilità di accadimento è stimata come BASSA in considerazione delle misure di garanzia implementate.

3.4 Perdita di dati

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

Perdita di fiducia, irritazione, perdita reputazione, perdita di controllo sui propri dati personali

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

Azione di virus informativi o di codici malefici, Sabotaggio, attacco hacker, Uso non autorizzato dei dati, Uso improprio del software, Accessi esterni non autorizzati

3.4.2 Quali sono le fonti di rischio?

Fonti di origine naturale, fonti di rischio umane esterne, fonti di rischio umane interne

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

Backup, Disaster Recovery plan, Manutenzione, Politica di tutela della privacy, Controllo degli accessi logici, Crittografia, Tracciabilità, Lotta contro il malware, Firewalling, EDR, Registrazione dei log di accesso al server applicativo e *database*.

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

La probabilità di accadimento è stimata come BASSA in considerazione delle misure di garanzia implementate.

3.5 METRICHE PER ANALISI RISCHIO

Valori dei livelli di rischio

Livello	Descrizione
BASSO	Il rischio per gli interessati è accettabile dall'organizzazione mediante misure organizzative e tecniche idonee, ma deve continuare ad essere monitorato per controllare che cambiamenti non incrementino il livello di rischio
MEDIO	Il rischio medio per gli interessati potrebbe essere accettabile ma l'adozione delle misure tecnico-organizzative deve essere monitorata su base regolare, e il trattamento può essere sottoposto a ulteriori considerazioni
ALTO	Il rischio per le persone interessate al trattamento è ad un livello non accettabile enecessita un rafforzamento delle misure di mitigazione
ELEVATO	Il rischio per gli interessati si presenta elevato o molto critico, mantenendo un livello non accettabile per l'organizzazione e necessitando l'aggiunta di ulteriori controlli a prevenzione/mitigazione dello stesso

Valori dei livelli di probabilità

Livello	Descrizione
BASSO	Evento/Minaccia poco probabile/frequente, o raro; è improbabile che

	la minaccia si concretizzi in condizioni normali o può verificarsi con frequenza inferiore rispetto alle tendenze riportate da studi, ricerche, statistiche di settore
MEDIO	Evento/Minaccia possibile; è un evento che si è già verificato o che può verificarsi con frequenza in media con le tendenze riportate da studi, ricerche, statistiche di settore
ALTO	Evento/Minaccia probabile; è un evento che si è già verificato o che può verificarsi con frequenza superiore rispetto alla media con riferimento alle tendenze riportate da studi, ricerche, statistiche di settore

Valori dei livelli di impatto

Livello	Descrizione
IRRILEVANTE	Gli interessati possono incontrare alcuni piccoli inconvenienti, che supereranno senza troppi problemi
LIMITATO	Gli interessati possono incontrare disagi significativi, che riusciranno comunque a superare a dispetto di alcuni problemi
SIGNIFICATIVO	Gli interessati possono incontrare conseguenze significative, che dovrebbero essere ingrado di superare anche se con gravi difficoltà
CRITICO	Gli interessati possono avere conseguenze gravi, o addirittura irreversibili, che potrebbero non superare

4 Panoramica dei rischi

Rischio Privacy	Descrizione delle conseguenze per gli interessati derivanti dalla vulnerabilità del trattamento	Livello di impatto
Perdita dei dati personali	**La perdita dei dati potrebbe comportare un danno agli interessati in termini di perdita di controllo sui propri dati**	**MEDIO**
Distruzione non autorizzata o indisponibilità	**La distruzione dei dati o l'indisponibilità degli archivi dello Studio non comporta un impatto diretto sugli interessati**	**BASSO**
Modifica non autorizzata	**La modifica dei dati per finalità di ricerca non comporta un impatto diretto sugli interessati**	**BASSO**
Divulgazione non autorizzata	**La divulgazione di dati personali dei partecipanti allo Studio potrebbe comportare un danno rilevante agli interessati**	**ALTO**
Accesso ai dati non autorizzato	**L'accesso illecito ai dati personali dei partecipanti allo Studio potrebbe comportare un danno rilevante agli interessati**	**ALTO**

Eccessiva raccolta di dati personali	**Utilizzare più dati personali del dovuto implicherebbe un'esposizione di dati personali all'utilizzo per scopi non pertinenti e non compatibili**	**BASSO**
Collegamenti o raffronti inappropriati o non autorizzati a dati personali	**Collegamenti o raffronti con altre banche dati potrebbe comportare danni immateriali agli interessati**	**BASSO**

Rischio Privacy	**Descrizione delle conseguenze per gli interessati derivanti dalla vulnerabilità del trattamento**	**Livello di impatto**
Perdita di controllo dei dati da parte degli interessati	**La mancanza di trasparenza e sicurezza dei trattamenti potrebbe comportare un impatto per gli interessati**	**BASSO**
Riuso per finalità diverse dei dati personali senza la consapevolezza e/o il consenso degli interessati	**I dati personali potrebbero essere utilizzati per altre finalità sconosciute all'interessato con danno immateriale agli interessati (mancanza di trasparenza e consenso)**	**BASSO**
Disequità o difettosità dell'elaborazione o del processo	**In caso di errata elaborazione delle informazioni, errori di registrazione etc. gli interessati potrebbero subire nocumento**	**BASSO**

Conservazione immotivatamente prolungata dei datipersonali	**La conservazione dei dati oltre il periodo prestabilito e motivato potrebbe comportare un danno immateriale agli interessati**	**BASSO**
Inesattezza o perdita di qualità dei dati personali	**Eventuali inesattezze o perdita della qualità dei dati raccolti non presenta un impatto diretto sui pazienti**	**BASSO**
Re-identificazione dei soggetti interessati	**Il processo di anonimizzazione potrebbe non eliminare la probabilità di re-identificazione dei partecipanti allo Studio, con particolare riferimento a malattie rare, con conseguente nocumento agli interessati**	**BASSO**

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

CATEGORIE DI MINACCE CONSIDERATE	Livello MAX Prob.
Minacce alla conformità del trattamento	BASSO
Eventi con danni fisici	BASSO
Eventi naturali	BASSO
Indisponibilità dei servizi essenziali	BASSO
Violazioni di dati per azioni deliberate	MEDIO
Problemi tecnici	BASSO
Violazioni di dati per azioni involontarie	BASSO

CATEGORIE DI MINACCE	EFFICACIA MISURA ESISTENTE
Minacce alla conformità del trattamento	MISURE ESISTENTI ADEGUATE
Eventi con danni fisici/materiali/immateriali	MISURE ESISTENTI ADEGUATE
Eventi Naturali	MISURE ESISTENTI ADEGUATE
Indisponibilità di Servizi essenziali	MISURE ESISTENTI ADEGUATE
Compromissione di dati e informazioni per azioni deliberate	MISURE ESISTENTI ADEGUATE

Problemi tecnici	**MISURE ESISTENTI ADEGUATE**
Compromissione di dati o servizi per azioni involontarie	**MISURE ESISTENTI ADEGUATE**

A seguito della ponderazione del livello di rischio calcolata mediante l'applicazione della mitigazione delle misure tecniche ed organizzative, il **rischio residuo** risulta **BASSO**, pertanto

ACCETTABILE ☒	**NON ACCETTABILE** ☐

DATA PROTECTION IMPACT ASSESSMENT

ai sensi dell'art. 35 del Reg. UE 2016/679

Il Titolare del trattamento, in persona del direttore Generale *pro tempore*, preso atto delle valutazioni sopra riportate in ordine all'analisi del potenziale impatto per i diritti e le libertà degli interessati, con l'adozione della VIP, dispone che il documento:

a) sia reso pubblico sul sito internet istituzionale nell'apposita sezione della Ricerca Scientifica (pubblicazione obbligatoria se lo Studio rientra nell'ambito del programma di ricerca nazionale, ai sensi dell'articolo 12-bis del decreto legislativo 30 dicembre 1992, n. 502)
b) sia resa disponibile agli interessati, su istanza dei medesimi.

Data……..

Firma del Direttore Generale……