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 PERSONALI

Codice	Descrizione
DPIA-001	**modello multimodale per Predire la RIsposta alla chemio-radioterapia in pazienti con cancro della CErvice localmente avanzato – studio PRICE**
ELABORAZIONE DPIA PER	☒ **Nuova attività trattamento** ☐ Aggiornamento DPIA ☐ Revisione periodica DPIA

Attività	Struttura/Funzione	Responsabile	data	firma
Redazione	**Principal Investigator**	*Annarita Fanizzi*		
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
Centri partecipanti quali Titolari del trattamento	Non presente
RESPONSABILE DEL TRATTAMENTO	Non presente
COORDINATORE E SPERIMENTATORI	Coordinatore: PhD Annarita Fanizzi, Ricercatore Sanitario IRCCS Istituto tumori 'Giovanni Paolo II' - Laboratorio di Biostatistica e Bioinformatica All'interno della steering committee sono da considerarsi: Dott.ssa Raffaella Massafra, IRCCS Istituto tumori 'Giovanni Paolo II' - Laboratorio di Biostatistica e Bioinformatica Dott.ssa Maria Colomba Comes, IRCCS Istituto tumori 'Giovanni Paolo II' - Laboratorio di Biostatistica e Bioinformatica Dott.ssa Samantha Bove, IRCCS Istituto tumori 'Giovanni Paolo II' - Laboratorio di Biostatistica e Bioinformatica Prof. Gennaro Cormio, IRCCS Istituto tumori 'Giovanni Paolo II' – S.C. di Ginecologica Oncologica Clinicizzata Dott. Oreste Donato, IRCCS Istituto tumori 'Giovanni Paolo II' – S.S.D. Radiologia Diagnostica. Dott. Alfredo Zito, IRCCS Istituto tumori 'Giovanni Paolo II' – S.C. Anatomia Patologica. Dott.ssa Annalisa Nardone, Istituto tumori 'Giovanni Paolo II' – S.C. di Radioterapia.

DATA PROTECTION IMPACT ASSESSMENT

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

	Dott. Emanuel Naglieri, Istituto tumori 'Giovanni Paolo II' – S.S. Oncologia Medica dell'apparato urogenitale.

MODALITA' CONDUZIONE	☒	DPIA OBBLIGATORIA
	☐	DPIA VOLONTARIA

DATA PROTECTION IMPACT ASSESSMENT

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

INDICE

Sommario

Informazioni sulla DPIA 7

ACCETTABILITA' DEL RISCHIO 8

1 Descrizione sistematica del trattamento 9

1.1.1 Contesto 9

1.2 Panoramica del trattamento 9

1.2.1 Quale è il trattamento in considerazione? 9
1.2.2 Quali sono le responsabilità connesse al trattamento? 14
1.2.3 Ci sono standard applicabili al trattamento? 14

1.3 Dati, processi e risorse di supporto 15

1.3.1 Quali sono i dati trattati e gli asset a supporto? 15

1.4 Finalità del trattamento 17

2 Principi Fondamentali 18

2.1 Valutazione della necessità e proporzionalità del trattamento del trattamento 18

2.1.1 Gli scopi del trattamento sono specifici, espliciti e legittimi? 18
2.1.2 Quale è il flusso dei dati durante il ciclo di vita del trattamento? 19
2.1.3 Quali sono le basi legali che rendono lecito il trattamento? 19
2.1.4 I dati sono esatti e aggiornati? 20
2.1.5 Qual è il periodo di conservazione dei dati? 20

2.2 Misure a tutela dei diritti degli interessati 20

2.2.1 Come sono informati del trattamento gli interessati? 20
2.2.2 Ove applicabile: come si ottiene il consenso degli interessati? 21
2.2.3 Come fanno gli interessati, o loro aventi diritto, a esercitare i loro diritti? 21
2.2.4 In caso di trasferimento di dati al di fuori dell'Unione europea, i dati godono di una protezione equivalente? 22

2.3 Misure esistenti o pianificate per la protezione del dato 22

3 Rischi 24
3.1 Panoramica dei rischi per diritti e libertà 24
3.2 Accesso illegittimo ai dati 27
3.2.1 Quali potrebbero essere i principali impatti sugli interessati se il rischio si dovesse concretizzare? 27
3.2.2 Quali sono le principali minacce che potrebbero concretizzare il rischio? 27
3.2.3 Quali sono le fonti di rischio? 27
3.2.4 Quali misure fra quelle individuate contribuiscono a mitigare il rischio? 27
3.2.5 Come stimereste la gravità del rischio, specialmente alla luce degli impatti potenziali e delle misure pianificate? 27
3.2.6 Come stimereste la probabilità del rischio, specialmente con riguardo alle minacce, alle fonti di rischio e alle misure pianificate? 27
3.3 Modifiche indesiderate dei dati 28
3.3.1 Quali sarebbero i principali impatti sugli interessati se il rischio si dovesse concretizzare? 28
3.3.2 Quali sono le principali minacce che potrebbero consentire la concretizzazione del rischio? 28
3.3.3 Quali sono le fonti di rischio? 28
3.3.4 Quali misure, fra quelle individuate, contribuiscono a mitigare il rischio? 28
3.3.5 Come stimereste la gravità del rischio, in particolare alla luce degli impatti potenziali e delle misure pianificate? 28
3.3.6 Come stimereste la probabilità del rischio, specialmente con riguardo a minacce, fonti di rischio e misure pianificate? 28
3.4 Perdita di dati 28
3.4.1 Quali potrebbero essere gli impatti principali sugli interessati se il rischio dovesse concretizzarsi? 28
3.4.2 Quali sono le principali minacce che potrebbero consentire la materializzazione del rischio? 29
3.4.3 Quali sono le fonti di rischio? 29
3.4.4 Quali misure, fra quelle individuate, contribuiscono a mitigare il rischio? 29
3.4.5 Come stimereste la gravità del rischio, specialmente alla luce degli impatti potenziali e delle misure pianificate? 29
3.5 METRICHE PER ANALISI RISCHIO 29

4 Panoramica dei rischi 31

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 osservazionale retrospettivo e prospettico 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.

Il presente documento contiene:
a) una descrizione 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, è risultato:
☒ BASSO ☐ MEDIO ☐ ALTO

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

1 Descrizione sistematica del trattamento

1.1.1 Contesto

1.2 Panoramica del trattamento

1.2.1 Quale è il trattamento in considerazione?

Il disegno dello studio prevede il collezionamento dati sia retrospettivo che prospettico. I principali criteri di inclusione sono:

- diagnosi confermata di carcinoma della cervice localmente avanzato,
- disponibilità di immagini di RM al basale (T0) e post-CRT (T1), di vetrini bioptici H&E digitalizzati e di dati clinici completi
- follow-up minimo di 12 mesi.

La disponibilità di immagine ecografiche al tempo T0 e al tempo T1 non è requisito essenziale di inclusione. Tuttavia, tale esame sarà ugualmente collezionato qualora presente per future valutazioni del potere predittivo delle stesse.
Per risposta alla CRT è valuta a 12 mesi dal termine della CRT, come da pratica clinica, e rappresenta il ground truth che sarà utilizzato come outcome dei modelli computazionali che si andranno a sviluppare. Pazienti che non hanno manifestato una progressione di malattia o una recidiva, sono intese come pazienti che hanno risposto alla terapia. In caso contrario i pazienti saranno considerati non rispondenti.
La raccolta della coorte retrospettiva nei primi 24 mesi dall'inizio del progetto. Il reclutamento della coorte prospettica durerà 24 mesi dall'inizio del progetto. Tuttavia, sui dati retrospettivi raccolti entro i primi 8 mesi, saranno avviate le attività di analisi di dati e immagini.
L'unità operativa di Oncologia Ginecologica clinicizzata, diretta dal Prof. Gennaro Cormio, collaborerà allo sviluppo del progetto nella fase di reclutamento della coorte retrospettiva e prospettica selezionando i casi secondo i criteri di inclusione definiti, nonché nella successiva discussione critica dei risultati ottenuti.

Per le finalità dello studio saranno collezionate le immagini di risonanza magnetica (RM) pre- e post-trattamento CRT. Le sequenze di imaging comprendono: immagini pesate in T2 e immagini pesate in diffusione (DWI). Collaborerà allo sviluppo del progetto, l'unità operativa di Radiologia Diagnostica diretta dal dott. Oreste Donato per la fase di collezionamento delle immagini di RM e della successiva fase di selezione delle regioni di interesse (ROI) sulle stesse immagini.
Saranno inoltre collezionate le immagini dei vetrini digitalizzati del tessuto bioptico. Il vetrino più rappresentativo in ematossilina/eosina FFPE (Formalin-Fixed Paraffin embedded) del campione biologico bioptico sarà selezionato dall'anatomo-patologo di riferimento e digitalizzato mediante uno speciale scanner digitale, già presente presso il nostro Istituto. Collaborerà allo sviluppo del progetto, l'unità operativa di Anatomia Patologica diretta dal dott. Alfonso Zito per la fase di collezionamento dei VD e della successiva fase di selezione delle regioni di interesse (ROI) sulle stesse immagini.
Le immagini RM e VD saranno accompagnate da dati demografici e clinici (età, storia ginecologica, ecc.), come dettagliato nella CRF allegata la progetto, che potranno fornire ulteriori parametri di analisi. I dati clinici, riferiti a tali pazienti, saranno raccolti direttamente dalle cartelle cliniche e collezionati in un opportuno database crittografato. Tutti i dati raccolti saranno pseudonimizzati al momento del collezionamento. In particolare, ad ogni paziente elegibile sarà associato un ID univoco al momento della raccolta delle caratteristiche cliniche. Le immagini radiologiche saranno estratte dal sistema informativo RIS PACS in essere presso il nostro Istituto. Tale sistema consente l'estrazione in formato DICOM senza dati sensibili delle immagini di interesse e crea uno pseudonimo che sarà linkato con l'ID del paziente di riferimento, una chiave di associazione. Analogamente, i vetrini digitalizzati sono già catalogati e archiviati dagli operatori dell'anatomia patologica con uno pseudonimo che sarà anch'esso linkato all'ID dello stesso paziente.
Alle immagini di RM e digital pathology verranno applicate opportune tecniche di normalizzazione per ridurre le distorsioni derivanti dalle diverse risoluzioni e caratteristiche tecniche di acquisizione delle immagini. Un radiologo e un patologo identificheranno una ROI rispettivamente sulle immagini RM e sul VD.
Saranno in seguito modellate reti neurali convoluzionali (CNN) pre-addestrate utilizzate come estrattori di caratteristiche quantitative dalle immagini collezionate. Per ogni ROI estratta rispettivamente dalle RM e dai VD, saranno estratte le caratteristiche quantitative valutando diverse architetture di reti pre-addestrate sfruttando l'approccio del transfer learning, diffusamente utilizzato nella comunità scientifica per estrarre caratteristiche quantitative dalle immagini superando il limitate tipico della ridotta disponibilità dei dati di uno studio monocentrico.

L'approccio del transfer learning consente l'estrazione di migliaia di caratteristiche, numero che varia in base al livello di profondità della rete da cui si estraggono le stesse. Necessaria risulterà, dunque, una importante analisi di riduzione della dimensionalità. Opportune tecniche saranno di riduzione della dimensionalità saranno dunque verificate e valutate al fine di individuare un sottoinsieme significativo funzionale al raggiungimento di prestazioni accurate e robuste.
Il sottoinsieme di caratteristiche significative selezionate sia dalle immagini di risonanza magnetica che dalle immagini di vetrini digitalizzati saranno dunque utilizzati dapprima separatamente per addestrare algoritmi di machine learning in grado di predire in maniera accurata la risposta alla CRT. In seguito, sarà sviluppato un modello multimodale in grado di migliorare le performance dei singoli modelli. Saranno pertanto implementati e valutati diverse tecniche ensemble, noti allo stato dell'arte, per analizzare congiuntamente i dati multimodali collezionati.
Infine, saranno applicate tecniche di intelligenza artificiale intuitiva (XAI) per visualizzare e comprendere il contributo delle caratteristiche radiomiche (ad esempio, le mappe di attivazione) e cliniche (ad esempio, il valore di Shapley) più fortemente associate agli esiti di interesse.
Nello specifico, con riferimento a OR 2, saranno sviluppati modelli di classificazione binari con il fine ultimo di discriminare pazienti responder e non responder. Saranno dapprima utilizzate le immagini CT, i vetrini bioptici digitalizzati e i dati clinici collezionai al momento della diagnosi (baseline), quindi prima della terapia; un primo modello di classificazione binaria multimodale sarà quindi sviluppato con il fine ultimo di discriminare precocemente i pazienti responder e no-responder alla CRT (Mod1). Tale modello ha lo scopo di fornire al clinico un primo allert sulla probabilità di non risposta, consentendogli eventualmente di definire un piano terapeutico più efficace. Un secondo modello predittivo della risposta alla CRT sarà sviluppato utilizzando oltre alle informazioni baseline collezionate, anche la RM post-trattamento (Mod2). Tale modello consentirà al clinico di definire un piano terapeutico aggiuntivo o di monitoraggio più inteso alle pazienti risultate non rispondenti.
Con riferimento a OR 4, saranno sviluppati modelli basati sullo stesso fremowork di analisi definito per OR2 per predire gli esiti di lungo periodo, come sopravvivenza libera da recidiva (DFS) e la sopravvivenza globale (OS). Saranno quindi sviluppati machine learning survival model utilizzando caratteristiche quantitative estratte dalle immagini radiologiche e di digital pathology.

Metodologie e Analisi statistica
In questo progetto, verranno utilizzate CNN pre-addestrate per l'estrazione di caratteristiche quantitative dalle immagini di RM e dalle immagini dei vetrini digitalizzati. Per set di dati limitati

nell'imaging medico, viene spesso utilizzata la tecnica del transfer learning. Tale tecnica, consiste nel pre-addestramento della rete neurale convoluzionale (CNN) per aggirare il requisito della mole massiccia di dati necessari per il processo di addestramento da zero di una rete convoluzionale. Le CNN sono state precedentemente addestrate (CNN pre-addestrate) su un numero enorme (milioni) di immagini naturali non mediche per apprendere come estrarre automaticamente caratteristiche di diverso tipo che sono, di basso livello caratteristiche, ad esempio bordi e punti, che rappresentano strutture locali dell'immagine e caratteristiche di alto livello, ad esempio forme e oggetti, che catturano caratteristiche globali dalle immagini. Le conoscenze che le reti hanno acquisito durante la formazione possono essere trasferite e applicate su immagini mai inedite attraverso diversi campi di ricerca (transfer learning). Nel corso dello studio saranno valutate diverse architetture di reti pre-addestrate, tra cui AlexNet, ResNet-50, Inception-V3, Xception, GoogLeNet. L'approccio che si intende implementare genera una pleura di caratteristiche quantitative per ciascuna tipologia di immagine considerata e pertanto, risulta necessaria una importante fase di feature selection e riduzione della dimensionalità.
Un sottoinsieme di caratteristiche significative sarà opportunamente selezionato, valutando diversi metodi noti allo stato dell'arte. Quindi, algoritmi di machine learning, noti allo stato dell'arte tra cui Random Forest, XGBoost, Support Vector Machine o Artificial Neural Networks, saranno addestrati sul sottoinsieme delle caratteristiche estratte dapprima considerando separatamente le caratteristiche radiomiche e quelle di digital pathology, e in seguito saranno elaborate congiuntamente anche con le caratteristiche clinico-patologiche collezionate al fine di definire un modello accurato e robusto predittivo della risposta alla CRT.
Infine, Le tecniche di Intelligenza Artificiale Esplicita (XAI) saranno applicate per visualizzare e comprendere il contributo delle caratteristiche quantitative estratte dalle immagini e cliniche (più fortemente associate agli esiti di interesse. Tecniche come Grad-CAM o LIME saranno applicate per identificare le aree delle immagini che contribuiscono maggiormente alla previsione della presenza di mutazioni e tecniche di interpretabilità per determinare l'influenza di ciascuna caratteristica clinica.
Con riferimento all'obiettivo secondario OR4, attualmente è disponibile un'ampia varietà di tecniche, che vanno dalle classiche curve descrittive di Kaplan-Meier non parametriche alle estensioni del modello inferenziale semi-parametrico di Cox. Un limite degli algoritmi classici è la difficoltà di modellare l'alta dimensionalità. Recentemente, le tecniche di apprendimento automatico applicate ai compiti di sopravvivenza consentono di superare questo problema. Infatti, la regressione di Cox classica è un modello parametrico basato su una stima probabilistica le cui

prestazioni di previsione dipendono dai parametri associati a ciascuna caratteristica. Pertanto, la difficoltà di individuare un modello probabilistico accurato aumenta parallelamente all'incremento del numero di caratteristiche considerate. Al contrario, i modelli di sopravvivenza ad apprendimento automatico, come i modelli di sopravvivenza random forest e gradient boosting, sono metodi non parametrici le cui prestazioni dipendono dalla dimensione dell'insieme di addestramento. In pratica, questi ultimi modelli non impongono alcuna ipotesi sulla distribuzione probabilistica, permettendo così di modellare adeguatamente le non linearità e gli effetti di interazione in un approccio data driven. Se, da un lato, queste limitazioni sono perseguite per ottenere la spiegabilità dei modelli di sopravvivenza black-box machine learning, dall'altro il loro superamento garantisce prestazioni più elevate.

Analisi statistiche: L'associazione tra le variabili cliniche e la risposta alla terapia misurata su una scala nominale binaria (Responder/No-respondr) sarà valutata attraverso l'implementazione di test statistici appropriati parametrici e non paramentrici.
Con riferimento agli esiti di lungo periodo, modelli di overall survival e progression-free survival saranno implementati.
I modelli sviluppati saranno opportunamente validati mediante tecniche di cross-validation sulla casistica retrospettiva, e in seguito su validati in maniera indipendente sulla casistica prospettica. Le prestazioni del modello saranno misurate in termini di area sotto la curva ROC (AUC), accuratezza, precisione, sensitivity e specificity.
Per misurare le prestazioni del modelli di sopravvivenza, obiettivo secondario dello studio, sarà utilizzato l'indice di concordanza (C-index) viene utilizzato.

Numerosità campionaria: Studi riportano che un modello predittivo è definito moderatamente accurato se raggiunge un'accuratezza tra il 70-90% (10.1126/science.328761). In relazione a tale dato, la numerosità minima campionaria del campione di validazione necessaria al raggiungimento di un modello di previsione moderatamente accurato pari all'80%, fissato un errore alfa tollerato pari a 0.05 e una precisione della stima di 0.15, deve essere non inferiore a 27 pazienti. Pertanto, considerando uno schema di hold-out validation 30:70, la numerosità campionaria minima del set di training deve essere di 63. Saranno dunque collezionati in retrospettivo almeno 100 pazienti in totale.
Per la coorte prospettica, una volta avviato lo studio, saranno reclutati consecutivamente tutti i pazienti elegibili che possono maturare il follow-up minimo di 12 mesi entro il termine dello studio.

Sulla base di stime di proiezione interne effettuate dal clinico dell'unità collaborante di Ginecologia Oncologica, nell'arco dei 24 mesi dall'inizio del progetto, sarà possibile reclutare 30 casi.
Pertanto, tra coorte retrospettiva e prospettica saranno reclutati almeno 130 pazienti complessivamente.
L'accuratezza globale del modello rimane comunque garantita dal calcolo della numerosità campionaria effettuata sulla coorte retrospettiva, in quanto, tale calcolo, definisce la numerosità campionaria minima necessaria al raggiungimento del livello di accuratezza prefissato, avendo definito a probabilità di errore di primo tipo e il livello di precisione.

Tipologia di Studio

Trattasi di studio osservazionale retrospettivo e prospettico non interventistico, 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 le 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?

….indicare….

Tipologia di dati personali	Categoria interessati
☒ **Dati identificativi comuni (es. nome, cognome, indirizzo)** ☒ **Dati di contatto (recapiti email, telefono, cellulare, etc.)** ☒ **Dati sanitari raccolti da archivi cartacei** ☒ **Dati raccolti da strumenti 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 sulla localizzazione ☐ Dati sulla solvibilità economica	• Pazienti in parte 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 biometrici ☐ Dati "giudiziari" (diritto penale)	
☐ dati soggetti a maggior tutela: dati relativi alle infezioni da HIV, all'uso di sostanze stupefacenti, psicotrope e di alcool, alle prestazioni erogate alle donne che si sottopongono ad interventi di interruzione volontaria della gravidanza o che decidono di partorire in anonimato, ad atti di violenza sessuale o di pedofilia, ai servizi offerti dai consultori familiari (l. 15 febbraio 1996, n. 66; l. 3 agosto 1998, n. 269; l. 6 febbraio 2006, n. 38; l. 5 giugno 1990, n. 135; d.P.R. 9 ottobre 1990, n. 309; l. 22 maggio 1978, n. 194; d.m. 16 luglio 2001, n. 349; l. 29 luglio 1975, n. 405)	
☒ **Altro:** età alla diagnosi, immagini mammografiche/ecografiche diagnostiche, vetrini istologici digitalizzati, campioni biologici	

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 delegati al trattamento di dati personali da parte del P.I. tramite apposito atto di nomina individuale
Soggetti esterni	NA
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), QuPath, 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 risulta necessario per le seguenti finalità dello Studio:

OR1 – Raccolta dati clinici, immagini di risonanza magnetica (RM) pre- e post-trattamento e VD bioptici

- OR 1.1 – Raccolta dati retrospettiva
- OR 1.2 – Raccolta dati prospettivo

OR 2- Sviluppo di un modello predittivo della recidiva e progressione di malattia a seguito della CRT

- OR 2.1 - Analisi delle immagini di RM pre-trattamento e post-trattamento
- OR 2.2 – Analisi delle immagini di VD della biopsia
- OR 2.3 – Definizione di un modello predittivo multimodale

OR 3 – Validazione dei modelli predittivi unimodali e multimodale sviluppati

OR 3.1 - Validazione dei modelli mediante tecniche di cross-validation sulla coorte retrospettiva collezionata
OR 3.2 - Validazione dei modelli sulla coorte prospettica collezionata
OR 3.4 - Analisi critica dei risultati ottenuti
OR 4 – Generalizzazione del modello OR2 di prevedere gli esiti clinici a lungo periodo
OR 4.1. Sviluppo modello predittivo di un modello predittivo degli esiti di lungo periodo, come la sopravvivenza libera da recidiva (DFS) e la sopravvivenza globale (OS)

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, unitamente alla presente VIP, l'informativa Privacy sullo studio in parola. 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; successivamente si provvede all'annotazione in un file di excel protetto da cifratura. I dati e le immagini, sui quali si avvieranno le attività di ricerca e Studio, saranno precedentemente pseudonimizzati. In un file separato (tabella di transcodifica) saranno conservate le associazioni codice pseudonimizzato (generatore casuale) e nome/cognome del paziente per l'eventuale necessità di dover rintracciare il paziente. Su tali dati verranno effettate le attività delle elaborazioni statistiche peculiari dello Studio.

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 ed i campioni biologici 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 3 anni dalla conclusione dello Studio, decorsi i quali saranno anonimizzati.
Tabella di corrispondenza Codice ID/Paziente	La tabella di corrispondenza sarà cancellata in modalità permanente decorsi 3 anni dalla conclusione 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. 6, par. 1, lett. A) e 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 28 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, o loro aventi diritto, a esercitare i loro diritti?

I diritti degli interessati di cui agli artt. 15-22 del GDPR sono garantiti nelle modalità indicate nell'informativa ex artt. 13-14 del GDPR rese al momento dell'arruolamento o della pubblicazione. 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, ivi comprese le valutazioni d'impatto sulla protezione dati.

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.

2.3 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. esecuzione di algoritmi di hashing non reversibili a chiave e/o generazione manuale di pseudonimo
  2. tabella fisica di associazione pseudonimo/identità custodita in armadio a chiave dal PI e solo da questi accessibile
- 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:

Per la custodia e la sicurezza dei dati genetici e dei campioni biologici sono adottate le seguenti

cautele:
a) l'accesso ai locali avviene previa identificazione delle persone, preventivamente autorizzate, che accedono a qualunque titolo dopo l'orario di chiusura;
b) la conservazione, l'utilizzo e il trasporto dei campioni biologici avvengono con modalità volte anche a garantirne la qualità, l'integrità, la disponibilità e la tracciabilità;
c) la consultazione dei dati genetici e biologici trattati con strumenti elettronici è consentita tramite sistemi di autenticazione multi-fattore;
d) i dati genetici e i campioni biologici contenuti in banche dati, sono trattati con tecniche di cifratura /pseudonimizzazione.

Con specifico riferimento alle operazioni di elaborazione dei dati dello Studio memorizzati in database centralizzato presso l'IRCCS BARI, sono implementate le seguenti misure di garanzia:

a) sistemi di autenticazione e di autorizzazione per il personale preposto al trattamento in funzione dei ruoli ricoperti e delle esigenze di accesso e trattamento;
b) procedure per la verifica periodica della qualità e coerenza delle credenziali di autenticazione e dei profili di autorizzazione assegnati ai soggetti designati al trattamento (sperimentatori);
c) sistemi di audit log per il controllo degli accessi al database e per il rilevamento di eventuali anomalie.

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?

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 assessment, 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 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

3.4.2 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

3.4.3 Quali sono le fonti di rischio?

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

3.4.4 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*.

3.4.5 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 presente 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……