

Valutazione d'Impatto sulla Protezione dei dati (Data Protection Impact Assessment)



M/903/150-C
Rev. 4

La DPIA (Data Protection Impact Assessment) – o anche VIP (Valutazione d'Impatto Privacy) - è un processo (che esita in un documento) inteso a descrivere il trattamento di dati personali, valutarne la necessità e la proporzionalità, nonché contribuire a gestire i rischi per i diritti e le libertà delle persone fisiche derivanti dal trattamento, valutando detti rischi e determinando le misure per affrontarli. È strumento e conseguenza della responsabilizzazione del titolare, e si riferisce a un trattamento conosciuto analiticamente e descritto in ogni suo aspetto; essa, perciò, assume anche una valenza organizzativa, con positiva ricaduta sul piano operativo e logistico dello studio, in particolare se osservazionale (uno studio, cioè, che si risolve esclusivamente nella raccolta ed elaborazione di dati per lo più personali. La DPIA mette dunque a disposizione, in generale:

- una descrizione sistematica del trattamento;
- la esplicitazione delle finalità del trattamento;
- una valutazione della necessità e proporzionalità del trattamento;
- una valutazione dei rischi per i diritti e le libertà degli interessati;
- le misure tecniche e organizzative che il titolare ritiene di dover adottare allo scopo di mitigare tali rischi.

La DPIA è redatta dal preposto al trattamento e oggetto di parere da parte del Responsabile della protezione dei dati.

DESCRIZIONE DEL TRATTAMENTO DEI DATI

Indicare la denominazione del trattamento

STRAP 2: Profilo di salute delle frazioni dell'area pistoiese in cui è stato identificato un cluster di sarcomi - approfondimento nell'ambito del progetto STRAP

Indicare la finalità del trattamento

Il trattamento è funzionale alla conduzione di uno studio osservazionale retrospettivo il cui obiettivo principale è quello di valutare il profilo di salute delle frazioni di Casalguidi e Cantagrillo del comune di Serravalle Pistoiese e le frazioni limitrofe dei Comuni di Pistoia e Quarrata (Masiano e Santonuovo, rispettivamente, e altre che verranno eventualmente individuate) entro un raggio di 3,5 km da Casalguidi/Cantagrillo attraverso uno studio di coorte residenziale retrospettivo, che consideri anche il possibile ruolo di determinanti ambientali, con particolare riferimento ai casi di sarcoma.

Mentre l'obiettivo secondario è la valutazione del grado di mobilità/stanzialità delle famiglie del cluster (oggetto di uno studio di epidemiologia genetica, sottoposto alla valutazione del Comitato Etico Regione Toscana - Pediatrico) e della popolazione residente nelle frazioni in esame rispetto al resto del territorio compreso nei Comuni interessati per acquisire informazioni sul comportamento migratorio e sulla stabilità della popolazione, contribuendo a una migliore comprensione delle dinamiche demografiche a livello locale. Si fa presente che le modalità di trattamento dei dati nell'ambito dello studio STRAP2 saranno le stesse eseguite come da normale pratica della UFC, che routinariamente, per fini di programmazione e monitoraggio salute, tratta dati nominativi sensibili di residenti in AUSL TC o che hanno usato servizi della stessa AUSL.

Indicare le tipologie di dati oggetto del trattamento, specificando ogni tipologia di dato

dati demografici, dati sanitari ovvero basso peso alla nascita, nati pretermine, prevalenza dei difetti congeniti; decessi, ricoveri, e dati di mortalità delle patologie del sistema circolatorio, dell'apparato respiratorio, gastrointestinali, dell'apparato nefrologico e urinario, del sistema nervoso; dati incidenza delle malattie oncologiche con particolare focus sui sarcomi; dati relativi alla mobilità residenziale, ai cambi di residenza, inclusa la frequenza e, se disponibile, alla distanza degli spostamenti.

Indicare le tipologie di interessati al trattamento

La coorte sarà costituita dalle persone che hanno risieduto nel Comune di Serravalle Pistoiese, nelle frazioni di Masiano di Pistoia e di Santonuovo di Quarrata e in altre frazioni che eventualmente verranno individuate entro un raggio di 3,5 km da Casalguidi/Cantagrillo, per almeno un giorno, dal gennaio 1999, oppure prima data disponibile delle anagrafi comunali informatizzate, al 30 giugno 2024. Si evidenzia che i record linkage dei dati anagrafici e sanitari coinvolgono sia la coorte dei residenti nelle frazioni menzionate che i residenti nel resto del territorio comunale, che costituiscono la popolazione di confronto.

Indicare i soggetti interni che partecipano al trattamento quali persone espressamente designate o autorizzate

Personale medico e statistico della UFC Epidemiologia opportunamente istruito e formato nella gestione dei dati.

Indicare eventuali soggetti esterni che partecipano al trattamento quali titolari, responsabili o persone designate/autorizzate al trattamento

L'ISPRO (dott. Calogero Saieva) ha il compito di fornire i dati di incidenza dei tumori elencati in Tabella 1 del protocollo e i dati di mortalità delle cause di morte indagate nella stessa tabella.

L'Università di Parma (Prof. Matteo Manfredini) sviluppa e implementa le metodologie statistiche per il calcolo dell'indice di mobilità.

Valutazione d’Impatto sulla Protezione dei dati (Data Protection Impact Assessment)



M/903/150-C
Rev. 4

Descrivere il flusso dati (cioè come i dati sono spostati o elaborati). Occorre descrivere il flusso analiticamente nei suoi vari passaggi, operazioni, attori

Le fonti dei dati degli esiti sanitari sono quelle del Registro di Mortalità Regionale (RMR) di ISPRO per i decessi, e delle schede di dimissione ospedaliera (SDO) dell’Azienda USL Toscana Centro per i ricoveri e di Azienda Regionale di Sanità (ARS) per i riferimenti provinciale e regionale.

Per gli esiti avversi della riproduzione, in particolare il basso peso alla nascita e la prevalenza di nati pretermine, verrà utilizzata la banca dati dei Certificati di Assistenza al Parto (CedAP), disponibili a livello regionale presso l’ARS.

Accanto a queste banche dati verranno analizzati i dati dei Registri di patologia disponibili a livello regionale, che già hanno fornito il loro supporto e analisi durante gli accertamenti precedenti.

A tal fine, per l’incidenza delle malattie oncologiche verrà analizzato il Registro Tumori Regionale (RTT) gestito da ISPRO, i cui dati attualmente disponibili coprono il periodo 2013-2019. Verrà in particolare fatto un focus specifico sui “sarcomi”, come definiti nella classificazione utilizzata in oncologia (ICD O3), dato che le codifiche attualmente utilizzate per le cause di morte e le schede di dimissione ospedaliera, ICD 9 e ICD 10-CM rispettivamente, non includono codici specifici per tutti i sarcomi e quindi non permettono un’analisi adeguata di queste patologie.

Per l’analisi della prevalenza delle malformazioni congenite verranno utilizzati i dati del Registro Toscano dei Difetti Congeniti (RTDC), gestito dal CNR di Pisa.

Gli eventi presi in considerazione saranno quelli occorsi dal 1999, oppure per i dati di incidenza delle malattie oncologiche, dal 2013. Il termine del periodo di osservazione sarà esito-specifico in base alla disponibilità dei dati (e.g. 31-12-2021 per le cause di morte, 30-06-2024 per le cause di ricovero e CedAP, 31-12-2020 per incidenza oncologica, 31-12-2023 per i dati sulle anomalie congenite). Per la definizione dello stato in vita saranno utilizzati i dati sui decessi registrati nelle anagrafi comunali e nell’Anagrafe degli Assistibili dell’AUSL TC, più aggiornati rispetto al RMR.

Per la valutazione della durata della residenza e della mobilità residenziale saranno raccolte dall’anagrafe comunale, e, se disponibili, dai dati dei censimenti della popolazione, le informazioni su età, genere, stato socio-economico e i dettagli sui cambi di residenza, inclusa la frequenza e, se disponibile, la distanza degli spostamenti.

Indicare dove vengono archiviati e conservati i dati

I dati verranno archiviati e conservati in un database (Excel) in forma pseudonimizzata per cui ciascun paziente avrà un codice personale (ID paziente). I codici ID saranno assegnati dal PI e potranno essere ricollegati alle identità dei pazienti solo dal team di Sperimentatori coinvolti presso il centro (lista di decodifica interna).

L’accesso al database presso il centro sarà vincolato da password temporanea in possesso al solo team di sperimentatori coinvolti nello studio presso lo stesso centro.

Il database Excel sarà salvato su una cartella protetta da password all’interno e localizzata su server aziendale.

Anche la lista di decodifica sarà salvata su una cartella protetta da password all’interno e localizzata su server aziendale.

PRINCIPI FONDAMENTALI

Limitazione delle finalità: indicare la base giuridica del trattamento, cioè la sua finalità lecita, quale prevista ex artt. 6 e 9 del Regolamento UE 2016/679 (d’ora in poi Regolamento)

In questo studio retrospettivo, non sarà possibile informare ottenere il consenso degli utenti che hanno risieduto nel Comune di Serravalle Pistoiese, nelle frazioni di Masiano di Pistoia e di Santonuovo di Quarrata e in altre frazioni per almeno un giorno, dal gennaio 1999, oppure prima data disponibile delle anagrafi comunali informatizzate, al 30 giugno 2024.

Per cui la base giuridica è rappresentata dal parere positivo del competente comitato etico a livello territoriale (e la successiva autorizzazione del Direttore Generale dell’Azienda USL Toscana Centro), alla luce della nuova formulazione dell’art. 110 del D.Lgs. 30 giugno 2003 n. 196 Codice in materia di protezione dei dati personali, conseguente alle modifiche apportate dalla Legge 56 del 29 aprile 2024.

Minimizzazione dei dati: indicare i criteri utilizzati per garantire l’adeguatezza, la pertinenza e la non eccedenza dei dati utilizzati
I dati estratti dagli applicativi sono solamente quelli ritenuti indispensabili alla ricerca.

Si fa presente che le modalità di trattamento dei dati nell’ambito dello studio STRAP2 saranno le stesse eseguite come da normale pratica della UFC, che routinariamente, per fini di programmazione e monitoraggio salute, tratta dati nominativi sensibili di residenti in AUSL TC o che hanno usato servizi della stessa AUSL.

Limitazione della conservazione: indicare per quanto tempo sono conservati i dati e i criteri per la conservazione dei dati

Il termine di conservazione dei dati è fissato a 5 anni, dopo i quali i dati saranno cancellati; si è considerato opportuno applicare a questo studio osservazionale tale termine ritenuto sufficiente per la tipologia di studio (osservazionale retrospettivo) e di pubblicazione. I dati

Valutazione d'Impatto sulla Protezione dei dati (Data Protection Impact Assessment)



M/903/150-C
Rev. 4

aggregati (non personali) generati dall'analisi statistica saranno conservati su file csv/dta per la durata di 5 anni dal momento della pubblicazione dello studio e poi, a loro volta, cancellati.

Esattezza dei dati: indicare le misure individuate per aggiornare, correggere o cancellare i dati che risultano non esatti in riferimento alla finalità per la quale sono trattati
I dati estrapolati e raccolti sono ricontrollati dai source documents originali dal Medico prima della loro pseudonimizzazione e del trasferimento dei dati su database.

Integrità e riservatezza dei dati: indicare le misure tecniche ed organizzative adottate per garantire la sicurezza dei dati rispetto a trattamenti non autorizzati o illeciti, perdita, distruzione o danni accidentali, precisando quanto segue:
L'accesso al database sarà vincolato al solo personale nominato, istruito e formato presso il centro.
Si fa presente che le modalità di trattamento dei dati nell'ambito dello studio STRAP2 saranno le stesse eseguite come da normale pratica della UFC, che routinariamente, per fini di programmazione e monitoraggio salute, tratta dati nominativi sensibili di residenti in AUSL TC o che hanno usato servizi della stessa AUSL.

Indicare se nel trattamento o in qualche sua fase (specificare) i dati sono pseudonimizzati, e secondo quali modalità
I dati personali, estrapolati dalle cartelle cliniche del centro, saranno immediatamente pseudonimizzati con codice alfa numerico al momento della raccolta e saranno riportati su database.
La modalità di pseudonimizzazione dei dati avverrà attraverso l'assegnazione di un codice numerico (ID paziente). I dati personali sono trattati in associazione con questa informazione non direttamente identificativa, e la correlazione tra questa e i dati identificativi dell'interessato è conservata separatamente, accessibile al solo personale del centro coinvolto nello studio, assicurandone la riservatezza con idonee misure di sicurezza, sia tecniche sia organizzative.
Il Subject ID consisterà in un codice numerico progressivo, generato ogni qual volta un nuovo paziente viene arruolato nello studio.

Indicare se nel trattamento o in qualche sua fase (specificare) i dati sono crittografati, e secondo quali modalità (ovvero quale sistema di crittografia è utilizzato)
Viene utilizzato il sistema di crittografia AES256 del database pseudonimizzato nel momento di inserimento dei dati del primo utente.

Indicare se nel trattamento o in qualche sua fase (specificare) i dati sono anonimizzati, e secondo quali modalità
I dati sono resi anonimi una volta distrutta la lista di decodifica, al termine dell'attività di elaborazione ovvero entro un massimo di 12 mesi dall'inizio dello studio, e comunque prima di ogni eventuale pubblicazione.

Indicare i criteri di profilazione per l'accesso ai dati
Gli accessi al database Excel, per le caratteristiche stesse del programma, non sono profilati. Solo lo staff del centro opportunamente formato ed autorizzato a lavorare sullo studio sarà in possesso della password per l'accesso al database.

Indicare se gli accessi sono tracciati
Gli accessi al database Excel, per le caratteristiche stesse del programma, non sono tracciati. L'accesso sarà consentito al solo personale del centro opportunamente formato e in possesso della password di accesso alla cartella contenente il database.

Indicare con quale frequenza viene effettuato il backup dei dati
Backup dei dati viene effettuato regolarmente come avviene per le cartelle su server aziendale (settimanalmente).

Indicare se il sistema prevede misure contro virus e malware
Tutti i PC aziendali sono aggiornati e dotati di efficaci software antivirus aggiornati e volti a contrastare eventuali malware

Indicare se i dati sono trattati anche su supporti cartacei, e come questi sono gestiti
Non è previsto l'utilizzo di supporti cartacei. Questi saranno conservati presso il centro nell'Investigator Site File in una stanza con accesso limitato e sotto la responsabilità del PI e del gruppo di lavoro.

DIRITTI DEGLI INTERESSATI

Ove applicabile: indicare come sono informati gli interessati al trattamento
Non sarà possibile rintracciare alcun paziente. Pertanto, si prevede, nel rispetto dell'art. 6 delle Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica, di procedere a rendere disponibili le informazioni, di cui si allega copia, redatte ai sensi all'art. 14 del Regolamento UE 2016/679, mediante pubblicazione sul sito istituzionale del Centro di sperimentazione per la durata dello studio

**Valutazione d'Impatto sulla Protezione dei dati
(Data Protection Impact Assessment)**



M/903/150-C
Rev. 4

stesso.

Ove applicabile: indicare le ragioni per cui non è possibile informare gli interessati

La raccolta dati avverrà in assenza del consenso scritto per motivi di impossibilità organizzativa per il fatto che il profilo di salute coinvolge migliaia di individui: contattare ciascuno di essi per ottenere il consenso sarebbe un compito enorme e dispendioso in termini di tempo e risorse. Inoltre, sarebbe impossibile rintracciare tutti gli individui i cui dati sono inclusi nello studio, considerando che si utilizzano dati storici o di persone che si sono trasferite.

Ove applicabile: indicare come è acquisito il consenso degli interessati

Non applicabile

Ove applicabile: indicare se il trattamento coinvolge soggetti qualificati come responsabili del trattamento

Non applicabile

GESTIONE DEI RISCHI

ACCESSO ILLEGITTIMO AI DATI

La gravità del rischio può essere considerata di basso livello, vista la specificità e le caratteristiche dei dati sensibili trattati, la probabilità del rischio si ritiene trascurabile. I dati sono infatti pseudonimizzati e separati dalle informazioni anagrafiche dei pazienti; il server che ospita il database è accessibile esclusivamente attraverso la rete Aziendale. Le password sono in possesso del solo personale interno autorizzato.

MODIFICHE INDESIDERATE DEI DATI

La probabilità del rischio di modifica indesiderata dei dati può essere ritenuta trascurabile, anche alla luce delle misure pianificate. La gravità del rischio è stimata di medio livello. I dati vengono sottoposti a backup settimanale, con possibilità di rapido restore in caso si verifichi una modifica indesiderata. L'accesso in scrittura ai dati è riservato a selezionati utenti in possesso di password di accesso, ed avviene attraverso controlli interni che minimizzano la probabilità di errore.

PERDITA DEI DATI

La probabilità di perdita dei dati è estremamente bassa visti i backup settimanali previsti dal sistema.

IL PREPOSTO AL TRATTAMENTO
(nome/cognome)

Miriam Levi

FIRMA

Data 12/02/2025