La DPIA (Data Protection Impact Assesment) – 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[1] e oggetto di parere da parte del Responsabile della protezione dei dati.

DESCRIZIONE DEL TRATTAMENTO DEI DATI
Indicare la denominazione del trattamento[2] Reperti morfologici, fenotipici e genetico-molecolari ricorrenti nella Leucemia/Linfoma a cellule T dell'Adulto (ATLL): l'esperienza di un Centro Italiano in area non endemica per HTLV-1.
Indicare la finalità del trattamento[3] Verificare se l'impiego di analisi genetico-molecolari, incluse la ricerca di aberrazioni cromosomiche, mutazioni delle regioni codificanti e studi di metilazione, possa migliorare l'accuratezza diagnostica nei casi di sospetta Leucemia/Linfoma a cellule T dell'Adulto (ATLL).
Indicare le tipologie di dati oggetto del trattamento, specificando ogni tipologia di dato[4] • Dati anagrafici: data di nascita, sesso • Dati istopatologici: tipologia di biopsia, sede di prelievo, eventuali altre sedi di malattia, data della diagnosi istologica • Dati laboratoristici: conta dei globuli bianchi, livelli di LDH, presenza di ipercalcemia • Dati clinico-strumentali di stadiazione: stadiazione radiologica di malattia • Dati relativi al trattamento: prima linea di chemioterapia, profilassi del sistema nervoso centrale, risposta al termine del trattamento (EOT), eventuale recidiva, seconda linea di trattamento chemioterapico • Dati di follow-up: stato in vita, decesso ed eventuale causa di morte
Indicare le tipologie di interessati al trattamento[5] Tre pazienti con sospetta Leucemia/Linfoma a cellule T dell'Adulto (ATLL) valutati presso la SOD di Istologia Patologica e Diagnostica Molecolare dell'AOU Careggi
Indicare i soggetti interni che partecipano al trattamento quali persone espressamente designate o autorizzate[6] Oltre al Principal Investigator (PI) (Prof. Raffaella Santi), il gruppo di sperimentazione è composto da un medico della SOD di Istologia Patologica e Diagnostica Molecolare dell'AOU Careggi (Dott. Gioia Di Stefano), un medico dell'Unità di Ematologia dell'AOU Careggi (Dott. Benedetta Sordi).
Indicare eventuali soggetti esterni che partecipano al trattamento quali titolari, responsabili o persone designate/autorizzate al trattamento L'Istituto di Genetica dell'Università di Ulm (Ulm, Germania) partecipa allo studio esclusivamente all'esecuzione delle analisi molecolari sui campioni biologici pseudonimizzati. L'Università di Ulm non ha accesso a dati anagrafici né a informazioni cliniche identificative e non dispone della chiave di ricodifica. Inoltre, non determina le finalità e i mezzi del trattamento, operando esclusivamente secondo le istruzioni del Titolare del trattamento. Pertanto, non riveste il ruolo di Titolare del trattamento ai sensi del Regolamento (UE) 2016/679.
Descrivere il flusso dati (cioè, come i dati sono spostati o elaborati). Occorre descrivere il flusso analiticamente nei suoi vari passaggi, operazioni, attori[7]. I dati clinici (ottenuti dalla piattaforma Archimed) e i dati biologici relativi ai campioni tissutali (ottenuti dalla piattaforma Winsap 3.0) vengono raccolti presso il centro coordinatore italiano. A ciascun partecipante viene assegnato un codice identificativo univoco (pseudonimo). La chiave di ricodifica è conservata esclusivamente presso il centro coordinatore. Prima del trasferimento, i campioni biologici vengono etichettati esclusivamente con il codice pseudonimo. Nessun dato anagrafico o clinico identificativo viene trasferito al partner tedesco. I campioni pseudonimizzati vengono spediti

all'Istituto di Genetica dell'Università di Ulm per le sole analisi genetiche, nel rispetto della normativa vigente e del Regolamento (UE) 2016/679. I dati genetici prodotti, associati esclusivamente al codice pseudonimo, vengono trasmessi come dati crittografati usando l'estensione FlowCrypt: Encryption per Gmail al centro Promotore che provvederà al loro inserimento sul sistema Redcap (accessibile esclusivamente a personale autorizzato mediante credenziali individuali).e alla loro rimozione al termine dello studio. I dati genetici saranno oggetto di analisi descrittiva, finalizzata alla caratterizzazione dei profili molecolari in relazione agli obiettivi dello studio.

Indicare dove vengono archiviati e conservati i dati[8]
I dati saranno archiviati sulla piattaforma Redcap accessibile esclusivamente a peronale autorizzato mediante credenziali individuali.

PRINCIPI FONDAMENTALI[9]

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)[10]
La base giuridica del trattamento è il consenso. Per gli interessati che non sarà possibile informare e per i quali non sarà possibile ottenere il consenso, è rappresentata, dal parere positivo del competente comitato etico a livello territoriale (e a successiva autorizzazione del Direttore Generale AOUC), alla luce della nuova formulazione dell'art. 110del D.Lgs. 30 giugno 2003 n. 196 Codice in materia di protezione dei dati personali, conseguente alle modifiche apportate dalla Legge 56 del29 aprile 2024.

Minimizzazione dei dati: indicare i criteri utilizzati per garantire l'adeguatezza, la pertinenza e la non eccedenza dei dati utilizzati[11]
I dati raccolti sono esclusivamente quelli indispensabili al raggiungimento degli obiettivi dello studio.

Limitazione della conservazione: indicare per quanto tempo sono conservati i dati e i criteri per la conservazione dei dati[12]
I dati saranno conservati per 7 anni. Considerato che, per gli studi osservazionali, la valenza normativa dei termini di conservazione previsti dalle disposizioni vigenti, sempre orientate a regolare gli studi interventistici, è, se certo non assente, comunque alquanto indiretta e non immediata, si è ritenuto opportuno applicare a questo studio osservazionale il termine dei 7 anni già previsto dal D.Lgs. 6 novembre 2007, n. 200.

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[13]
Il database utilizzato consente di effettuare un controllo dei dati inseriti, sia nella fase stessa dell'inserimento specificando la lunghezza della stringa ed il tipo di variabile, sia nella fase preliminare dell'elaborazione, in cui è possibile evidenziare valori outlyers e correggerli.

Integrità e riservatezza dei dati[14]: 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:
Per l'integrità e riservatezza dei dati ci si affida a RedCap l'applicativo dove vengono caricati i dati relativi al progetto. Il personale coinvolto esegue un processo di pseudonimizzazione che fa in modo di impedire l'associazione dei dati con ciascun paziente.

Indicare se nel trattamento o in qualche sua fase (specificare) i dati sono pseudonimizzati, e secondo quali modalità[15]
I dati saranno pseudonimizzati mediante l'attribuzione di un codice identificativo, da cui non sarà possibile in alcun modo risalire alla identità degli interessati. Lo Sperimentatore principale prof.ssa Santi sarà in possesso del documento dove sarà riportata la corrispondenza fra il codice ed i dati anagrafici e sarà responsabile della riservatezza di tali dati.

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) [16]
I dati saranno crittografati al momento del trasferimento presso altre strutture al di fuori del centro. I dati saranno crittografati usando l'estensione FlowCrypt: Encryption per Gmail.

Indicare se nel trattamento o in qualche sua fase (specificare) i dati sono anonimizzati, e secondo quali modalità[17]
Nella fase di pubblicazione i dati saranno riportati in modo totalmente anonimo, senza alcuna possibilità di risalire all'identità dei partecipanti.

Indicare i criteri di profilazione per l'accesso ai dati[18]
Nessun differente profili per tutti i medici appartenenti al gruppo di sperimentazione.

Indicare se gli accessi sono tracciati[19]
Gli accessi alla piattaforma Redcap sono tracciati.

Indicare con quale frequenza viene effettuato il backup dei dati[20]
Il backup dei dati viene effettuato con una cadenza settimanale.

Indicare se il sistema prevede misure contro virus e malware[21]
Tutti i computer sono aggiornati all'ultima versione del sistema operativo e sono dotati di efficaci software antivirus aggiornati volti a contrastare eventuali attacchi da parte di virus e malware.

Indicare se i dati sono trattati anche su supporti cartacei, e come questi sono gestiti[22]
I dati non saranno trattati su supporti cartacei.

DIRITTI DEGLI INTERESSATI

Ove applicabile: indicare come sono informati gli interessati al trattamento[23]
Gli interessati saranno informati mediante la messa a disposizione dell'informativa redatta ai sensi dell'Art. 13 del Regolamento UE 2016/679

Ove applicabile: indicare le ragioni per cui non è possibile informare gli interessati[24]
In considerazione della prognosi frequentemente infausta dell'Adult T-cell leukemia/lymphoma (ATLL) e dell'elevata prevalenza di pazienti provenienti da Paesi esteri, può risultare impossibile ottenere il consenso informato, sebbene si tratti di un numero limitato di soggetti. Le principali motivazioni includono il possibile decesso dei pazienti, la loro irreperibilità e/o l'oggettiva impossibilità organizzativa dovuta alla limitata disponibilità di recapiti completi e aggiornati, nonché alla distanza geografica che ostacola un contatto diretto ed efficace.

Ove applicabile: indicare come è acquisito il consenso degli interessati[25]
Dopo che lo studio sarà stato approvato da parte del Comitato Etico, il consenso informato sarà acquisito durante la prima visita ambulatoriale. Il processo di acquisizione del consenso prevederà la consegna dell'apposito modulo di consenso informato, accompagnata da una spiegazione chiara, completa ed esaustiva del progetto di ricerca, delle sue finalità, delle procedure previste, nonché dei potenziali benefici e rischi. Una volta che l'interessato si ritiene soddisfatto delle informazioni date si procede alla firma del documento.

Ove applicabile: indicare se il trattamento coinvolge soggetti qualificati come responsabili del trattamento[26]
Il trattamento non coinvolge soggetti qualificati come responsabili del trattamento.

GESTIONE DEI RISCHI[27]

ACCESSO ILLEGITTIMO AI DATI

Indicare una stima della probabilità e gravità del rischio (indefinita, trascurabile, limitata, importante, massima), anche alla luce delle misure pianificate, specificando, se possibile, le principali minacce che potrebbero concretizzare il rischio e le principali fonti di rischio.

Sebbene la gravità del rischio possa essere considerata di medio 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 il protocollo https (TLS) con esclusione di ogni accesso di altro tipo (SMB, FTP o altri). Gli accessi sistemistici di servizio (per manutenzione o aggiornamenti software) sono consentiti solo attraverso protocolli criptati (ssh o simili) e soltanto da rete intranet AOUC. Eventuali necessità di accessi da internet sono veicolate attraverso VPN. Le credenziali amministrative sono in possesso del solo personale interno autorizzato. Le credenziali di gestione dell'applicativo sono personali e rilasciate ai soli dipendenti autorizzati che sono stati istruiti riguardo la loro corretta custodia

MODIFICHE INDESIDERATE DEI DATI

Indicare una stima della probabilità e gravità del rischio (indefinita, trascurabile, limitata, importante, massima), anche alla luce delle misure pianificate, specificando, se possibile, le principali minacce che potrebbero concretizzare il rischio e le principali fonti di rischio.

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 giornaliero, con possibilità di rapido ripristino in caso si verifichi una modifica indesiderata. L'accesso in scrittura ai dati è riservato a selezionati utenti, ed avviene attraverso interfacce che minimizzano la probabilità di errore.

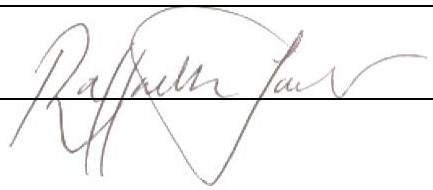
PERDITA DEI DATI

Indicare una stima della probabilità e gravità del rischio (indefinita, trascurabile, limitata, importante, massima), anche alla luce delle misure pianificate, specificando, se possibile, le principali minacce che potrebbero concretizzare il rischio e le principali fonti di rischio.

La probabilità di perdita dei dati è estremamente bassa, mentre l'eventuale danno sarebbe molto elevato. La stima considera le strutture hardware ridondate sulle quali si fonda il sistema, le procedure di backup sistematico e la resilienza intrinseca del data center che ospita l'applicativo. Per le eventuali perdite di dati causate da operatori infedeli, valgono le considerazioni dei punti precedenti.

IL PREPOSTO AL TRATTAMENTO (vedi nota 1)
(nome/cognome)

Prof.ssa Raffaella Santi

FIRMA 	*Data 26/10/2026*

[1] Il Preposto al trattamento, in Azienda, è, per quanto riguarda gli studi, il PI.
L'art. 2-quaterdecies del D.Lgs. 30 giugno 2003 n. 196 prevede che, tra le persone autorizzate al trattamento (cfr. nota 6), il titolare possa individuare, per lo svolgimento di specifiche funzioni di coordinamento e orientamento nel trattamento dei dati, persone fisiche, che operano sotto la loro autorità, a ciò "espressamente designate". La persona fisica "espressamente designata", che coincide con la vecchia nozione di "responsabile interno del trattamento" (oggi il responsabile del trattamento è solo un soggetto esterno), è stata sinteticamente ridenominata dai diversi titolari, utilizzando varie espressioni (delegato, referente ecc.): in Azienda la si è definita *Preposto*, con termine derivato dalla normativa in materia di sicurezza del lavoro, e che indica appunto un soggetto che sovraintende ad una data attività (a far intendere che il trattamento dei dati non è mai una attività sganciata da un concreto operare).
[2] Inserire titolo e codice dello studio.
[3] Finalità del trattamento vale il suo scopo pratico. Occorre dunque indicare, posto che il trattamento è ovviamente funzionale alla esecuzione dello studio, oltre allo scopo di ricerca in senso lato (es. "I dati sono trattati per scopo di ricerca in campo medico …"), quali sono gli scopi che si intendono raggiungere con lo studio medesimo (es. scopo dello studio è verificare …"). Qualora i dati vengano raccolti per una finalità ulteriore (es. di cura, il che significa che saranno trattati anche con modalità identificativa), occorre integrare tale specifico elemento nell'informativa sul trattamento dei dati.
[4] In via generale si tratta di dati afferenti alle categorie particolari, es. relativi alla salute o genetici, e di dati comuni (es. dati anagrafici e di contatto). Oltre a questa indicazione più generica, categorica, occorre esplicitare i dati che vengono effettivamente raccolti; ciò può essere fatto con un grado maggiore (es. esiti di questo o quell'esame di laboratorio) o minore (es. esiti esami di laboratorio) di analiticità: è comunque preferibile essere più analitici possibile –questi elementi più puntuali sono normalmente già elencati nel protocollo - anche per motivare, se necessario, tali scelte in una prospettiva di minimizzazione (cfr. nota 12), cioè di una loro stretta funzionalità/indispensabilità rispetto allo studio.
[5] L'interessato è la persona fisica cui si riferiscono i dati personali trattati: in uno studio, sono ad esempio i pazienti in esso arruolati, descritti attraverso le caratteristiche (es. di patologia, esiti, età) che li rendono in esso eleggibili. Occorre qui Indicare anche il range temporale entro il quale si vanno a identificare i pazienti eleggibili allo studio (es. pazienti diabetici trattati dal 1995 al 2020).
[6] È sufficiente indicare il numero dei componenti del gruppo di sperimentazione e le relative professionalità, senza indicazioni nominative. Per la persona espressamente designata, cfr. nota 1. La persona autorizzata al trattamento è la persona fisica – dipendente o collaboratore - sottoposta, per quanto concerne il trattamento dei dati, al Titolare (cioè, l'Azienda), e che tratta dati personali solo nella misura in cui sia stata a ciò autorizzata e istruita: le istruzioni delimitano l'ambito di trattamento autorizzato, e precisano le modalità secondo le quali il trattamento deve essere effettuato. Nessun incaricato può trattare dati senza adeguate istruzioni (che sono un suo diritto), e nessun incaricato, ricevutele, può effettuare operazioni di trattamento ulteriori rispetto a quelle da esse consentite. Tali istruzioni, nell'ottica della responsabilizzazione del titolare (che consiste nell'applicare i principi previsti all'art. 5 del regolamento UE 2016/679, documentandone le modalità di applicazione), devono essere raccolte in un atto di nomina a firma del P.I. (atto che potrà essere anche riferito al gruppo di sperimentazione nel suo complesso, oppure, qualora i compiti, all'interno del gruppo di sperimentazione siano significativamente differenziati, essere più personalizzato e quindi nominativo).
[7] Un trattamento di dati personali si traduce in un flusso di informazioni, che può coinvolgere vari spazi (es. banche dati), soggetti ecc., e che può sostanziarsi in una serie di operazioni (es. la raccolta dei dati, per la quale occorre indicare come essi vengono selezionati e archiviati, ad es. in un foglio di raccolta o in un database; o la loro comunicazione, tra due o più titolari; le modalità di elaborazione ecc.). È necessario indicare anche se i dati sono meramente trasferiti all'interno di uno stesso ambito di titolarità: il trasferimento del dato è nozione più ampia, e talvolta diversa, da quella della sua comunicazione, cioè della trasmissione del dato ad altro titolare; questa comporta spesso un trasferimento di dati (ma i dati possono essere comunicati anche mettendoli semplicemente a disposizione, senza trasmetterli); si ha trasferimento di dati sia se i dati sono trasferiti al di fuori dei sistemi aziendali, sia se sono meramente spostati anche all'interno del medesimo ambito di titolarità (cioè ad es. da un server all'altro dell'Azienda, o verso un server di un soggetto che agisce per il titolare, quale responsabile del trattamento). Occorre precisare se i dati sono eventualmente trasferiti:

- nell'ambito dell'Azienda
- fuori dall'Azienda
- fuori dall'Italia
- fuori dall'Unione Europea

Il trasferimento del dato, soprattutto se effettuato al di fuori del proprio ambito di titolarità (che normalmente corrisponde ad un perimetro presidiato), può rappresentare un momento critico, che necessita l'adozione di idonee misure di sicurezza tanto tecniche che organizzative: di quelle appunto specificamente riferibili al trasferimento del dato si richiede una breve descrizione. Il trasferimento dei dati (esclusi i dati genetici) deve essere effettato con modalità sicura, anche con strumenti di cooperazione applicativa oppure utilizzando strumenti di messaggistica che utilizzino canali di comunicazione protetti (ivi compresa la PEC), oppure, se si utilizzano sistemi di posta elettronica ordinaria, proteggendo l'allegato con tecniche di cifratura e rendendolo accessibile tramite una password per l'apertura del file trasmessa separatamente; qualora lo studio ricomprenda dati genetici, non sarà possibile utilizzare la mail ordinaria ma solo strumenti di cooperazione applicativa o di messaggistica che utilizzino canali di comunicazione protetti (ivi compresa la PEC), cifrando i dati e fornendo la chiave di

decifrazione attraverso canali di comunicazione differenti da quelli utilizzati per la trasmissione dei dati. È necessario richiamare eventuali agreement redatti per il trasferimento dei dati, e comunque documentare la valutazione della necessità e proporzionalità del trattamento che è stata effettuata.

[8] Si distingue qui tra archiviazione e conservazione, indicando con la prima voce la temporanea allocazione dei dati nel corso dello studio, con l'altra quella effettuata nel periodo successivo al termine dello studio, prima della definitiva cancellazione o anonimizzazione dei dati (sono comunque operazioni che possono essere effettuate con continuità sul medesimo sistema), è necessario individuare specificamente dove i dati vengono allocati, indicando anche il sistema o il data base utilizzato. Se per la loro successiva conservazione si utilizza, appunto, una banca dati diversa, occorrerà indicarla.
In ordine ai profili di sicurezza, anche in relazione alla esattezza ed integrità dei dati, è inutile precisare che un foglio Excel su un pc in locale non soddisfa i requisiti minimi (la DPIA non otterrà il parere positivo del Responsabile della Protezione dei dati aziendale, e quand'anche venisse trasmessa, è certo che non sarà possibile ottenere la autorizzazione del Garante).
Il sistema di archiviazione e conservazione dei dati di studio messo a disposizione dall'Azienda è RedCap; possono essere utilizzati strumenti diversi, ma che garantiscano, allo stesso modo, un tracciamento degli accessi e delle operazioni effettuate e garanzie contro virus, malware ecc.
Qualora venga utilizzata una piattaforma esterna, occorrerà procurarsi le relative informazioni tecnico informatiche, da mettere agli atti della documentazione di studio (di tale documentazione si potrà offrire evidenza, allegandola o meno, nel presente documento); non è necessario che tale documentazione sia esaustiva da un punto di vista tecnico, ma deve essere tale da fornire informazioni sufficienti ad effettuare una minima valutazione di adeguatezza, anche con il supporto della componente tecnico-informatica aziendale.

[9] L'art. 5 (*Principi applicabili al trattamento di dati personali*) par. 1 del Regolamento prescrive analiticamente alcuni principi che assicurano l'adeguatezza del trattamento (cd. *principi base del trattamento*); la *responsabilizzazione* del Titolare consiste appunto nel rispettare tali principi e nell'essere in grado di dimostrare, con idonea documentazione (redatta prima dell'inizio del trattamento, nell'ottica della privacy by design e by default) di averli rispettati. Dunque, il titolare del trattamento è responsabile del rispetto dei seguenti principi:

- limitazione della finalità del trattamento;
- limitazione della conservazione dei dati,
- minimizzazione dei dati;
- esattezza dei dati;
- sicurezza dei dati (integrità e riservatezza).
- trasparenza del trattamento (riguarda anzitutto le informazioni sul trattamento messe a disposizione degli interessati, se ne parla alla sezione successiva relativa ai Diritti degli interessati)

[10] La base giuridica ordinaria del trattamento dei dati a scopo di ricerca clinica è il consenso degli interessati, a seguito di idonee informazioni. Il consenso non è necessario se l'interessato non è contattabile, o se si tratta di studio previsto da una norma di legge/da una disposizione regolamentare/ dal diritto UE/dal programma di ricerca sanitaria di cui all'art. 12 bis del D.Lgs. 502/92.
Nel caso che non sia possibile informare l'interessato ed acquisirne il consenso, e si tratti di studio previsto da una norma di legge/da una disposizione regolamentare/ dal diritto UE/dal programma di ricerca sanitaria di cui all'art. 12 bis del D.Lgs. 502/92, occorre riportare quanto segue (scegliendo il caso d'interesse):

- La base giuridica del trattamento è rappresentata dalla legge (specificare), che ha previsto lo studio.
- La base giuridica del trattamento è rappresentata dalla disposizione regolamentare (specificare), che ha previsto lo studio.
- La base giuridica del trattamento è rappresentata dalla normativa UE
- La base giuridica del trattamento è rappresentata dal programma di ricerca sanitaria di cui all'art. 12 bis del D.Lgs. 502/92 (specificare l'anno), che ha previsto lo studio.

Se il paziente non è contattabile - perché i dati di contatto sono stati perduti o non sono aggiornati, oppure il paziente è deceduto, o è preferibile non informarlo per motivi etici (es. il paziente non è informato sulla patologia di cui è affetto) – oppure se i contatti non sono gestibili per oggettiva impossibilità di carattere organizzativo (contattare i pazienti comporterebbe un impegno sproporzionato rispetto alle risorse disponibili), la base giuridica del trattamento, è rappresentata dal parere positivo del comitato etico competente a livello territoriale, nonché dalla applicazione di misure di garanzia sulla sicurezza del trattamento (che qui stiamo appunto specificando).

[11] La minimizzazione dei dati si traduce appunto nella garanzia che i dati siano "adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati", art. 5 paragrafo 1 c del Regolamento). Ovvio che tali requisiti non possano essere assolutizzabili, in quanto strettamente funzionali allo scopo di un dato studio: sarà comunque lecito utilizzare, tanto da un punto di vista qualitativo che quantitativo, soltanto le informazioni indispensabili per quel determinato studio. Chi valuta quali dati sono o meno necessari? Ovviamente il Titolare (e per esso, in un progetto di ricerca, il P.I.) che, nell'ottica della responsabilizzazione, dovrà argomentare e sostenere tale valutazione. Nel nostro caso occorre dunque dimostrare che i dati trattati, e già sopra elencati, sono soltanto quelli necessari alla realizzazione dello studio, e non altri. È di tale necessità – strettamente correlata alla razionalità dello studio da un punto di vista eminentemente scientifico - che deve essere data brevemente evidenza, anche soltanto indicando in sintesi che "i dati raccolti sono quelli indispensabili alla

esecuzione dello studio". In relazione a certe tipologie particolari di informazioni, ad es. quelle relative alle origini razziali o alla appartenenza etnica, può essere opportuno offrire una motivazione più puntuale ed articolata.

[12] Un termine puntuale per la conservazione dei dati utilizzati per gli studi osservazionali non è previsto e dunque quello scelto deve essere motivato. Il termine deve essere commisurato allo scopo principale della conservazione dei dati, che è anzitutto quello di rendere possibili verifiche o controlli della base dati dello studio successivamente alla pubblicazione. Si consiglia di scrivere qualcosa di analogo a quanto segue:

> Il termine di conservazione dei dati è fissato a … (inserire il numero di anni ritenuto necessario) anni; si evidenzia la consapevolezza che la valenza normativa dei termini di conservazione previsti dalle disposizioni vigenti, sempre orientate a regolare gli studi interventistici, non è direttamente ed immediatamente prescrittiva per gli studi osservazionali, così che viene comunque chiamata in causa la responsabilizzazione del Titolare.
>
> Si è considerato opportuno applicare a questo studio osservazionale il termine di … anni in quanto …

Se si utilizza il termine di prassi di 7 anni, la motivazione può essere resa come segue, sostituendo l'ultima frase:

> Si è considerato opportuno applicare a questo studio osservazionale il termine di conservazione di 7 anni già previsto dal D.Lgs. 6 novembre 2007, n. 200, riferibile ad una prassi consolidata e soprattutto ritenuto sufficiente e non eccedente in relazione allo scopo di consentire eventuali controlli successivi sulla correttezza delle inferenze effettuate nella valutazione dei dati raccolti nel corso della ricerca. Il termine settennale è commisurato alla opportunità di conservare una base dati statistica per successive verifiche o richieste di precisazioni circa i risultati pubblicati.

Si ricorda che mediante l'informativa ex art. 13 o ex art. 14 del Regolamento occorre indicare e comunicare ai soggetti interessati, che:

- sono raccolti solo i dati strettamente necessari per il perseguimento delle finalità;
- decorsi i termini di conservazione, i dati personali saranno distrutti, cancellati o resi anonimi (descrivendo i meccanismi per la cancellazione o anonimizzazione dei dati).

Se i dati sono conservati a tempo indeterminato a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici è necessario indicarlo e motivarlo, anche in riferimento a specifiche prescrizioni normative.

[13] In questo caso l'esattezza del dato non si intende riferita al suo aggiornamento, ma alle modalità con le quali i dati sono raccolti dalla documentazione originale e dunque duplicati, garantendone appunto l'esattezza rispetto a quella, per le finalità dello studio. Ovvio che misure di controllo sono meno necessarie quando l'estrazione da un data base informatico avviene quasi automaticamente a seguito dell'inserimento di dati parametri, rispetto alla copia manuale, per la quale occorre individuare una procedura di verifica e controllo.

[14] Ai sensi dell'art. 5 par. 2 del Regolamento, i dati devono essere "trattati in maniera da garantire un'adeguata sicurezza dei dati personali, compresa la protezione, mediante misure tecniche e organizzative adeguate, da trattamenti non autorizzati o illeciti e dalla perdita, dalla distruzione o dal danno accidentali («integrità e riservatezza»)". Le politiche di sicurezza sono necessarie per evitare:

- la divulgazione di dati personali o l'accesso agli stessi non autorizzati o accidentali;
- la modifica non autorizzata o accidentale di dati personali (così che i dati sono modificati o incompleti);
- la perdita della possibilità di accesso o distruzione accidentale o non autorizzata di dati personali.

Occorre indicare, sinteticamente, le misure adottate da un punto di vista organizzativo, nonché quelle informatiche assicurate dal sistema sul quale i dati sono archiviati, anche attraverso il rimando alla relativa documentazione tecnica.

È ovvio che la modifica, la perdita o la non accessibilità ai dati sono questioni che non attengono esclusivamente alla privacy, ma direttamente alla qualità del dato di ricerca.

[15] La pseudonimizzazione (non *pseudo-anonimizzazione*, come si trova in qualche protocollo) consiste nell'associare dei dati (es. quelli relativi alla salute del partecipante allo studio) ad una informazione di carattere non identificativo (ad es. un codice), sostituendo con essa quella di carattere identificativo, ad es. il nome/cognome dell'interessato, e mantenendo riservata, con specifiche misure di sicurezza, la correzione tra dato identificativo e dato non identificativo (tra anagrafica e codice). Essa è una operazione di trattamento che si traduce in una misura di sicurezza e di minimizzazione dei dati. Non ogni codificazione può tradursi in una pseudonimizzazione dei dati: non lo è certo l'utilizzo del codice fiscale (ben più identificativo del mero nome giuridico), ma neppure un codice che sia conosciuto al di fuori del gruppo di sperimentazione (es. il numero nosologico o simile, anche a livello di singolo reparto).

Occorre descrivere come è costruito il codice, e come è strutturato e gestito il processo di pseudonimizzazione dei dati, cioè in quale fase dello studio si attua.

Comunque, se si crea un elenco, e questo ha una sua logica (ad es. alfabetica o cronologica), non è sufficiente togliere l'anagrafica ed inserire ad es. dei codici progressivi, occorre che siano non sequenziali e randomizzati (almeno se l'estrazione dei dati è eseguibile una seconda volta con identici risultati). Insomma, il codice di pseudonimizzazione non può contenere elementi oggettivi – informativi o di carattere procedurale – che rendano possibile una identificazione dell'interessato a prescindere dalla chiave di pseudonimizzazione. Si può scrivere quanto segue:

> La pseudonimizzazione dei dati avverrà attraverso l'assegnazione di un codice. 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 coinvolto nello studio,

assicurandone, con idonee misure di sicurezza tecniche ed organizzative, la riservatezza. I codici di pseudonimizzazione sono costruiti secondo la seguente modalità: ……. I dati sono pseudonomizzati …. (*indicare in quale fase avvien la pseudonimizzazione dei dati*)

[16] Occorre precisare se i dati, in qualche momento del processo (es. trasferimento o comunicazione, oppure archiviazione, sono cifrati, e con quale tecnica.

[17] Si ricorda che, per anonimizzazione ci si riferisce ad una tecnica che si applica ai dati personali al fine di ottenere una loro deidentificazione assoluta e irreversibile. In pratica, il dato anonimizzato non potrà più essere, in nessun contesto di trattamento, neppure in quello originario, ricollegato all'interessato. In pratica, un set di dati privato dell'anagrafica non è, come secondo la nozione etimologica o di senso comune, un dato anonimizzato: è, piuttosto, un dato personale non immediatamente identificativo. Un set di dati è anonimizzato solo quando è definitivamente e irreversibilmente privato, anche prospetticamente, di una possibilità di raccordo con la relativa anagrafica, nel senso che questa non è più recuperabile (e non è dunque più possibile una reidentificazione, cioè la eventualità che, partendo da dati erroneamente ritenuti anonimi, si riesca a recuperare informazioni identificative degli interessati, sia direttamente, sia tramite metodi di correlazione e deduzione).

Con questi presupposti, il dato anonimo/anonimizzato ben raramente può essere presente in uno studio se non nella fase conclusiva, quando si aggregano i dati in vista della pubblicazione degli esiti. La procedura con cui si anonimizzano i dati in vista della pubblicazione deve essere descritta; ordinariamente, non essendo auspicabile, in uno studio clinico il ricorso a tecniche di *randomizzazione*, che consistono nella modifica della veridicità dei dati, si ricorrerà a tecniche di *generalizzazione*, consistono nel generalizzare gli attributi delle persone interessate, diluendo i livelli di dettaglio. Si utilizzerà di solito, tra queste, il K.-Anonimato, tecnica volta ad impedire l'individuazione di persone interessate mediante il loro raggruppamento con almeno K altre persone (K=valore di soglia). Secondo la regola della soglia, le persone cui si riferiscono i dati si considerano non identificabili se il loro numero è superiore ad un certo valore prestabilito (valore di soglia). Il valore minimo ordinariamente attribuibile alla soglia è pari a tre (ma nel valutare il valore della soglia si deve tenere conto del livello di sensibilità delle informazioni, e dell'effettivo rischio di danno ad esse correlato). La regola della soglia sottende che il valore originale X possa essere riferito non al solo Caio, ma anche a Tizio, Tazio e Sempronio. La relazione biunivoca tra il valore X ed una (una sola) persona fisica viene così meno. Occorre indicare come si procede quando una tipologia di informazione resta sotto la soglia minima.

[18] La profondità di accesso indica il *quantum* di accessibilità ai dati che è riconosciuto ad una determinata persona autorizzata al trattamento (cfr. nota 6); essa deve riguarda tanto la quantità e la tipologia di informazioni accessibili, che le operazioni (lettura, scrittura, cancellazione, elaborazione ecc.) eseguibili sui dati. Tutte queste prerogative sono connesse ad uno o più profili di autorizzazione (e, correlativamente e simmetricamente, di protezione dei dati), che si chiede – qualora plurali - di elencare e descrivere nei loro contenuti.

[19] Il tracciamento degli accessi, con finalità di sicurezza e controllo, può riguardare tanto operazioni che modificano la consistenza dei dati che la loro mera consultazione. Tale tracciamento si traduce nella conservazione, per un certo periodo di tempo, di file di log (il log file è appunto un file che contiene un elenco cronologico delle attività svolte da un sistema operativo, da un database o da altri sistemi, per permettere una verifica successiva). È richiesto di specificare, appunto, se sono tracciati gli accessi degli utenti e degli amministratori, se sono tracciati anche gli accessi in consultazione, se sono tracciati i riferimenti temporali degli accessi, per quanto tempo gli eventuali file di log sono conservati. Il tracciamento degli accessi, con la registrazione delle operazioni effettuate, in particolare di modifica dei dati, è una misura essenziale per garantire la sicurezza dei dati, in particolare la loro esattezza ed integrità.

Per quanto riguarda la documentazione cartacea, si deve indicare se si procede o meno ad un controllo degli accessi fisici.

[20] Tra le misure che ostano alla perdita, totale o parziale, dei dati, vi è il backup, che può essere svolto con una diversa frequenza. Si chiede di precisare se il backup dei dati è assicurato, e con quale tempistica.

[21] Il termine malware indica un programma che è stato progettato per danneggiare un computer; è una sorta di genere ampio, rispetto alle specie quale troian, virus ecc. Un virus è un malware che tende a danneggiare file e dati.

[22] La gestione dei supporti cartacei, in questo caso, riguarda la loro archiviazione sicura e la loro accessibilità. Si ricorda che, anche se il trattamento è solitamente effettuato con strumenti elettronici, laddove presente l'acquisizione del consenso è quasi sempre effettuata utilizzando supporti cartacei.

[23] La modalità ordinaria è la messa a disposizione dell'interessato dell'informativa redatta ai sensi dell'art, 13 del Regolamento.

[24] Qualora non sia possibile o opportuno informare gli interessati ed acquisirne il consenso occorre non solo attestarne ma documentarne e comprovarne i motivi tra i seguenti:

- ✓ motivi etici riconducibili alla circostanza che l'interessato ignora la propria condizione e l'informativa comporterebbe la rivelazione di notizie la cui conoscenza potrebbe arrecare un danno materiale o psicologico agli interessati stessi;
- ✓ motivi di impossibilità organizzativa, nel senso che gli interessati, all'esito di ogni ragionevole sforzo compiuto per contattarli (anche attraverso la verifica dello stato in vita, la consultazione dei dati riportati nella documentazione clinica, l'impiego dei recapiti telefonici eventualmente forniti, nonché l'acquisizione dei dati di contatto presso l'anagrafe degli assistiti o della popolazione residente) risultino essere al momento dell'arruolamento nello studio deceduti o

comunque non contattabili, e la mancata considerazione dei dati riferiti a questi, rispetto al numero complessivo dei soggetti che si intende coinvolgere nella ricerca, produrrebbe conseguenze significative per lo studio in termini di alterazione dei relativi risultati (avuto riguardo ai criteri di inclusione previsti dallo studio, alle modalità di arruolamento, alla numerosità statistica del campione prescelto, nonché al periodo di tempo trascorso dal momento in cui i dati riferiti agli interessati sono stati originariamente raccolti).

Alcuni esempi:

- irreperibilità e/o oggettiva impossibilità organizzativa dovuta alla limitata disponibilità di indirizzi completi ed aggiornati dei pazienti;
- irreperibilità e/o oggettiva impossibilità organizzativa dovuta all'elevata percentuale di pazienti non più seguiti dal centro (di sperimentazione coinvolto);
- irreperibilità e/o oggettiva impossibilità organizzativa dovuta all'elevato intervallo di tempo tra il primo accesso del paziente al centro (di sperimentazione coinvolto) ed il data entry dello Studio;
- impossibilità organizzativa e/o di fatto dovuta alla lontananza geografica dei pazienti che rende eccessivamente difficoltoso e costoso il loro ritorno al centro (di sperimentazione coinvolto) per le procedure di consenso, unitamente alla difficoltà di interagire con l'ausilio di strumenti elettronici da parte di pazienti anziani o aventi poca dimestichezza con le attrezzature elettroniche/informatiche;
- decesso del paziente;
- intervenuta incapacità di intendere e/di volere dovuta all'aggravarsi dello stato clinico;
- sforzo oggettivamente sproporzionato rispetto agli obiettivi dello Studio che rischia di rendere impossibile o di pregiudicare gravemente il conseguimento delle finalità della ricerca.

Comunque, nel caso in cui informare gli interessati risulti impossibile o implichi uno sforzo sproporzionato, oppure rischi di rendere impossibile o di pregiudicare gravemente il conseguimento delle finalità della ricerca, occorre documentare le valutazioni effettuate e le evidenze raccolte per sostenere ciò, anche con riferimento a dati statistici (ad es. circa la mortalità della patologia oggetto dello studio) e, se del caso, i tentativi di contatto effettuati ed i loro esiti percentuali sul totale dei pazienti arruolabili, oppure l'impegno di risorse materiali ed umane che, in riferimento al numero dei pazienti da contattare, rende l'operazione non sostenibile dal punto di vista organizzativo.

Occorre inoltre predisporre una informativa ex art. 14 del Regolamento, articolo che riguarda appunto le informazioni da mettere a disposizione dei pazienti non contattabili (nel caso dei defunti, dei loro aventi causa) come previsto dall'art. 6 delle *Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica*; l'informativa sarà pubblicata in una sezione dedicata del sito istituzionale per tutta la durata dello studio stesso (nel caso di pazienti defunti, a beneficio di familiari ecc.).

Nell'informativa occorre indicare il soggetto cui sarà possibile rivolgersi, nel Centro di sperimentazione, per far valere i diritti degli interessati; si indica ordinariamente il responsabile aziendale della protezione dei dati, rpd@aou-careggi.toscana.it , 3666823917.

[25] Il «consenso al trattamento» è qualsiasi manifestazione di volontà libera, specifica, informata e inequivocabile, con la quale l'interessato manifesta il proprio assenso, mediante dichiarazione o azione positiva inequivocabile, che i dati personali che lo riguardano siano oggetto di trattamento. Il consenso, in quanto "manifestazione di volontà", deve appunto manifestarsi, ed è dunque prestato mediante un atto positivo inequivocabile, ad esempio mediante dichiarazione scritta, anche attraverso mezzi elettronici, o orale. Ciò può comprendere la selezione di un'apposita casella in un sito web o qualsiasi altra dichiarazione o qualsiasi altro comportamento che indichi chiaramente in tale contesto che l'interessato accetta il trattamento proposto. Non configura pertanto consenso il silenzio, l'inattività o la preselezione di caselle. Ad ogni modo, il titolare del trattamento deve essere in grado di dimostrare che l'interessato ha prestato il proprio consenso.

[26] È Responsabile del trattamento il soggetto esterno rispetto al titolare che tratta dati per conto – cioè, per le finalità – del titolare, secondo le modalità da questo indicate. Ai sensi dell'art. 28 paragrafo 3 del Regolamento tale incarico deve essere formalizzato in un contratto o altro atto giuridico, che precisi la materia disciplinata e la durata del trattamento, la natura e la finalità del trattamento, il tipo di dati personali e le categorie di interessati, gli obblighi e i diritti del titolare del trattamento; tale atto deve essere redatto in modo tale che il responsabile tratti i dati personali soltanto su istruzione documentata del titolare del trattamento.

[27] La parte conclusiva della DPIA, dopo la descrizione del trattamento e delle misure tecnico-organizzative individuate a garanzia della sua adeguatezza, è quella propriamente dedicata alla valutazione circa la sostenibilità dei rischi individuati. Tali rischi si articolano in riferimento alla perdita:

- di riservatezza dei dati
- di integrità dei dati
- di disponibilità dei dati

La stima conclusiva della probabilità e gravità di ogni tipologia di rischio è da indicarsi nei seguenti termini:

- indefinita
- trascurabile
- limitata
- importante

- massima.

Ogni valutazione sintetica deve essere adeguatamente motivata.

Qualora si utilizzi REDCAP; è possibile limitarsi indicare quanto segue:

Accesso illegittimo ai dati

Sebbene la gravità del rischio possa essere considerata di medio 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 il protocollo https (TLS) con esclusione di ogni accesso di altro tipo (SMB, FTP o altri).

Gli accessi sistemistici di servizio (per manutenzione o aggiornamenti software) sono consentiti solo attraverso protocolli criptati (ssh o simili) e soltanto da rete intranet AOUC. Eventuali necessità di accessi da internet sono veicolate attraverso VPN.

Le credenziali amministrative sono in possesso del solo personale interno autorizzato.

Le credenziali di gestione dell'applicativo sono personali e rilasciate ai soli dipendenti autorizzati che sono stati istruiti riguardo la loro corretta custodia

Modifiche indesiderate ai 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 giornaliero, con possibilità di rapido restore in caso si verifichi una modifica indesiderata.

L'accesso in scrittura ai dati è riservato a selezionati utenti, ed avviene attraverso interfacce che minimizzano la probabilità di errore.

Perdita dei dati

La probabilità di perdita dei dati è estremamente bassa, mentre l'eventuale danno sarebbe molto elevato.

La stima considera le strutture hardware ridondate sulle quali si fonda il sistema, le procedure di backup sistematico e la resilienza intrinseca del data center che ospita l'applicativo.

Per gli eventuali data loss causati da operatori infedeli, valgono le considerazioni dei punti precedenti.