



## IL GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

NELLA riunione odierna, alla quale hanno preso parte il prof. Pasquale Stanzione, presidente, la prof.ssa Ginevra Cerrina Feroni, vicepresidente, il dott. Agostino Ghiglia e l'avv. Guido Scorza, componenti, e il cons. Fabio Mattei, segretario generale;

VISTO il Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE-Regolamento generale sulla protezione dei dati (di seguito "Regolamento");

VISTI, in particolare, gli artt. 35 e 36 del Regolamento relativi, rispettivamente, alla valutazione d'impatto sulla protezione dei dati e alla consultazione preventiva dell'Autorità;

VISTO il d.lgs. 30 giugno 2003, n. 196 recante "*Codice in materia di protezione dei dati personali*" (di seguito "Codice");

VISTO l'art. 110 comma 1, secondo periodo del Codice relativo al trattamento di dati personali per ricerca medica, biomedica e epidemiologica;

VISTE le *Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica* adottate dal Garante, ai sensi dell'art. 20, comma 4, del d.lgs. 10 agosto 2018, n. 101, con provvedimento n. 515, del 19 dicembre 2018 (doc. web n. 9069637, di seguito "Regole deontologiche");

VISTE le *Prescrizioni relative al trattamento dei dati personali effettuato per scopi di ricerca scientifica*, allegato n. 5 al Provvedimento che individua le prescrizioni contenute nelle Autorizzazioni generali che risultano compatibili con il Regolamento e con il d.lgs. n. 101/2018 di adeguamento del Codice, del 5 giugno 2019 (doc. web n. 9124510);

VISTA l'istanza di consultazione preventiva presentata dall'Azienda Socio Sanitaria Territoriale degli Spedali Civili di Brescia, ai sensi degli artt. 110 del Codice e 36 del Regolamento per la realizzazione dello studio retrospettivo osservazionale denominato "*Intelligenza artificiale per la definizione di nuove signature e modelli per la personalizzazione delle strategie di preservazione d'organo del cancro laringeo e ipofaringeo - PRESERVE*" (nota del 22 novembre 2022);

VISTA la documentazione in atti;

VISTE le osservazioni formulate dal Segretario generale ai sensi dell'art. 15 del Regolamento del Garante n. 1/2000 sull'organizzazione e il funzionamento dell'ufficio del Garante per la protezione dei dati personali, in [www.gpdp.it](http://www.gpdp.it), doc. *web* n. 1098801;

Relatore l'avv. Guido Scorza;

## **PREMESSO**

### **1. L'istanza di consultazione preventiva e l'attività istruttoria svolta**

Con nota del 22 novembre 2022, l'Azienda Socio Sanitaria Territoriale degli Spedali Civili di Brescia (di seguito "Azienda") ha presentato un'istanza di consultazione preventiva, ai sensi dell'art. 110, comma 1, ultimo capoverso del Codice e dell'art. 36 del Regolamento, in qualità di promotore dello studio retrospettivo osservazionale denominato *"Intelligenza artificiale per la definizione di nuove signature e modelli per la personalizzazione delle strategie di preservazione d'organo del cancro laringeo e ipofaringeo - PRESERVE"* (di seguito "Studio"), trasmettendo il protocollo, la valutazione di impatto sul trattamento dei dati svolta ai sensi dell'art. 35 del Regolamento (Vip), il parere favorevole del Comitato etico territorialmente competente e l'informativa *privacy* da rendere agli interessati.

Lo scopo dello Studio è quello di *"creare un modello multimodale in grado di predire la risposta a chemioterapia di induzione (IC) come strategia per la preservazione d'organo nel cancro localmente avanzato della laringe e dell'ipofaringe"*, ciò in quanto attraverso l'*"analisi e l'integrazione delle caratteristiche cliniche, radiologiche, genetiche e biomolecolari della malattia e del paziente"* è possibile *"identificare preventivamente i soggetti rispondenti a strategie di preservazione d'organo comprendente la chemioterapia di induzione (IC), aumentando così il tasso di conservazione della laringe e incidendo positivamente sulla qualità di vita del paziente e sul rapporto costo-beneficio della strategia terapeutica"*.

Lo Studio inoltre *"porrà le basi per un futuro studio prospettico di fase II, in cui verrà studiata l'applicabilità e l'efficacia di un trattamento di induzione personalizzato"*. [...] *"Lo studio prospettico prevedrà inoltre, l'analisi approfondita della qualità di vita dei pazienti arruolati, outcome fondamentale della strategia di preservazione"*.

Gli obiettivi secondari dello Studio sono: *"1) Definire percorsi terapeutici alternativi da sfruttare nei pazienti non responsivi alla chemioterapia tradizionale. 2) Sviluppare strumenti di integrazione dei dati e algoritmi predittivi di risposta a[lla] chemioterapia di induzione (IC), definendo così un CDSS (sistema di supporto alla decisione clinica)"*.

Lo Studio, che è parte del più ampio progetto Europeo dal titolo: *“Multidisciplinary Research Projects on Personalised Medicine – Pre-/Clinical Research, Big Data and ict, Implementation and Users Perspective”*, prevede l’arruolamento di circa n. 300 pazienti sul territorio nazionale, trattati con chemioterapia di induzione (IC) nel periodo compreso tra il 2008 e il 2020, e si articola nelle seguenti in cinque fasi:

1. *“Raccolta dei dati clinici, biomolecolari e radiologico da una coorte retrospettiva di pazienti trattati con IC;*
2. *Integrazione del profilo trascrittomico con il profilo genetico somatico (ovvero profilo genetico della cellula tumorale) e con quello molecolare, evidenziando eventuali differenze intercorrenti fra il gruppo dei rispondenti e quello dei non rispondenti a IC;*
3. *Analisi radiomica;*
4. *Progettazione di un modello predittivo di risposta alla IC;*
5. *Integrazione dei modelli sviluppati al fine di sostenere un sistema di supporto alla decisione clinica”.*

Trattandosi di un tumore molto aggressivo, la maggior parte dei pazienti risultano deceduti e l’esclusione di tali pazienti *“non consentirebbe di avere una casistica rappresentativa dei fallimenti terapeutici per lo svolgimento della fase retrospettiva”*. Inoltre, poiché il progetto *“prevede che con l’analisi dei dati retrospettivi sia generato un algoritmo di predittività della risposta alla terapia [...], che sarà poi applicato ad uno studio prospettico randomizzato (fase 2 del progetto)”*, la mancanza dei dati *“ottenuti da pazienti deceduti, genererebbe un bias di selezione portando ad un’errata generazione dell’algoritmo”*.

Lo Studio prevede la partecipazione di venti centri in Italia e di alcuni centri europei localizzati in Spagna, Germania e Grecia. È prevista altresì la partecipazione dell’Università di Oslo, *Department of Informatics P.O*, che, in qualità di *“partner del progetto, contribuisce attraverso la fornitura e la gestione della piattaforma elettronica per la raccolta del dato necessario per l’analisi statistica”*, all’uopo nominata responsabile del trattamento ai sensi dell’art. 28 del Regolamento. È altresì rappresentato che l’Azienda assumerà il ruolo di promotore e di *“capofila del progetto europeo”*.

Nel protocollo dello Studio è evidenziato che i *“Clinical data will be retrieved from the Genomic Data Commons portal from TCGA Research Network”* e che *“1. We will perform bioinformatics analysis on the biomolecular data resulting from the predictive patients’ samples processing from WP2. 2. Starting from our retrospective pool of treated patients, where the efficacy of the treatment has been recorded together with a rich set of omics and clinical data from WP1, WP2 and WP3, we will design predictive models, using a similar approach as described in [paragraph 1]”*.

Con riferimento al principio di trasparenza e all'obbligo di fornire le informative agli interessati, l'Azienda ha inviato l'informativa sul trattamento dei dati personali che verrà fornita ai pazienti. Con riguardo ai pazienti deceduti o non raggiungibili l'Azienda ha dichiarato che *"in ottemperanza del disposto di cui all'art. 14, par. 5, lett. b del Regolamento UE 2016/679, il Promotore, prima dell'avvio dello Studio, renderà pubblica la relativa informativa attraverso una specifica inserzione sui propri siti internet aziendali"*.

In relazione alla durata dello Studio, nell'istanza è rappresentato che esso avrà una durata di 36 mesi e che i dati raccolti verranno conservati per 10 anni, fermo restando che *"Tale termine potrà estendersi al fine del rispetto dei termini di conservazione stabiliti nel Massimario di Scarto approvato dalla Regione Lombardia attualmente in vigore e ss.mm.ii. comunque non superiori a quelli necessari per la gestione dei possibili ricorsi/contenziosi"*; nella valutazione d'impatto è indicato un periodo di conservazione di 25 anni dalla conclusione dello Studio (cfr. 1\_Scheda trattamento *Preserve*).

È rappresentato, inoltre che, a garanzia della minimizzazione, i dati, in forma pseudonimizzata, *"saranno inseriti nella e-cfr TSD piattaforma web based"* gestita dall'Università di Oslo e saranno conservati solo in formato elettronico.

L'istanza di consultazione preventiva è accompagnata dalla valutazione d'impatto, svolta ai sensi dell'art. 35 del Regolamento, che si compone di una *"scheda dei trattamenti"* e di 4 allegati recanti rispettivamente una descrizione della metodologia adottata, l'analisi dei rischi generali e specifici dei trattamenti relativa allo Studio e la descrizione delle misure tecniche e organizzative che si intendono adottare per garantire un livello di sicurezza adeguato al rischio, prevedendo altresì specifiche tutele per il trasporto dei campioni biologici contenenti dati genetici.

Con nota del 16 dicembre 2022 (prot. n. 82086), l'Ufficio ha formulato una richiesta di informazioni al promotore al fine di ottenere specifici chiarimenti in ordine a:

- la correlazione tra lo Studio oggetto della presente istanza di consultazione preventiva e il richiamato Progetto europeo, i ruoli dei partecipanti localizzati in Europa e, qualora questi abbiano accesso ai dati personali -seppure pseudonimizzati- dei pazienti dello Studio che confluiranno nella piattaforma gestita dall'Università di Oslo, il presupposto giuridico sulla base del quale sarebbe consentito tale accesso;
- la modalità di assolvimento degli oneri informativi, secondo quanto indicato all'art. 12 del Regolamento, atteso che risultavano riportate in un unico documento sia le informazioni da fornire direttamente degli interessati, ai sensi dell'art. 13 del Regolamento, che quelle da rendere ai soggetti non contattabili, ai sensi dell'art. 14 del

Regolamento medesimo, con ciò pervenendosi ad un documento molto lungo e di non facile lettura;

- i tempi di conservazione, atteso il titolare ne ha indicati diversi (10 e 25 anni), facendo altresì riferimento a una ipotetica ulteriore conservazione dei dati, nel rispetto dei termini di stabiliti nel *“Massimario di Scarto approvato dalla Regione Lombardia attualmente in vigore e ss.mm.ii”*, senza per altro indicare le ragioni tecnico scientifiche per cui tale termine sia ritenuto necessario per il conseguimento delle finalità per le quali i dati sono raccolti;
- le specifiche garanzie che si intendono implementare a tutela dei diritti degli interessati, considerato che i dati saranno oggetto di trattamento attraverso strumenti di intelligenza artificiale e con logiche predittive basate su sistemi di *machine learning*.

Con nota de 27 gennaio 2023, lo *sponsor* ha fornito i chiarimenti richiesti, rappresentando quanto segue.

- *“Multidisciplinary Research Projects on Personalised Medicine – Pre-/Clinical Research, Big Data and ICT, Implementation and Users Persepective”* fa riferimento al titolo del Bando pubblico competitivo ERAPerMed Joint Translational Call 2020, cui ASST Spedali Civili ha partecipato in qualità di *“Coordinatore”* insieme ad altri enti Partner;
- il Progetto Europeo prevede la realizzazione di due studi clinici. Lo studio oggetto dell’istanza di consultazione preventiva fa esclusivamente riferimento ad una specifica fase del predetto progetto, ossia allo studio retrospettivo corrispondente ai *“Work Package/WP-1, WP-2, WP-3, WP-4, WP-5 del richiamato Progetto Europeo”*;
- il modulo di informativa e consenso allo Studio sono quindi relativi al predetto studio osservazionale retrospettivo, promosso da ASST Spedali Civili di Brescia, e sono rivolti ai potenziali soggetti arruolabili, afferenti ai Centri Italiani che non sono *“Partner”* del Progetto Europeo, ma sono stati identificati da ASST Spedali Civili al fine di contribuire alla selezione della casistica;
- i due moduli informativi ai sensi degli art. 13 e 14 del Regolamento nei quali è stato specificato il ruolo dell’ASST quale promotore dello Studio con l’*“eliminazione dell’indicazione della funzione di Centro di Sperimentazione”* e nelle quali sono stati indicati solo *“due distinti trattamenti che il Titolare andrà a svolgere nel corso della Sperimentazione clinica”*.

In relazione ai tempi di conservazione dei dati, l'Azienda ha chiarito di aver indicato 10 anni in quanto *"Alla conclusione del Progetto Europeo, potrebbe essere necessario avere a disposizione i dati che hanno determinato il calcolo matematico iniziale (raccolti nel corso dello studio in esame) per un congruo lasso di tempo (ragionevolmente individuato in 10 anni), per una eventuale revisione del dato fonte. Inoltre, anche qualora il Promotore si rendesse conto dell'impossibilità di raggiungere i risultati sperati, potrebbe risultare necessario verificare il dato fonte e rianalizzare i dati raccolti per capire le cause della determinazione del fallimento. Dette operazioni potrebbero ragionevolmente richiedere un lungo lasso di tempo al Promotore, individuato - appunto - in 10 anni"*. Inoltre, l'Azienda ha rappresentato che in quanto Pubblica Amministrazione – è anche *"obbligata a rispettare le tempistiche di conservazione stabilite dal Massimario di scarto predisposto dal gruppo di lavoro coordinato da Regione Lombardia per tutte le aziende sanitarie e socio-sanitarie presenti sul territorio regionali"*, che in ordine ai tempi di conservazione dei documenti relativi alla sperimentazione clinica dei farmaci e dei dispositivi per uso umano, indicano un periodo di 25 anni dalla conclusione della sperimentazione.

Per tale ragione, l'Azienda ha inteso individuare come termine di conservazione quello di 25 anni dalla conclusione dello Studio. Fermo restando che *"qualora Codesta Autorità ritenesse che detto termine venticinquennale non risulti adeguato ai sensi della normativa in materia di protezione dei dati personali, l'ASST procederà a ridurre la data retention dei dati personali dello Studio a 10 anni, ossia l'arco temporale necessario per il raggiungimento delle finalità del Progetto"*.

Da ultimo, l'Azienda ha precisato, che per la conduzione dello Studio (di natura prettamente retrospettiva), i dati personali dei pazienti arruolati saranno trattati esclusivamente per creare l'algoritmo necessario per pianificare la costruzione dello *"strumento"* di intelligenza artificiale, che troverà invece applicazione solo nella fase 2 del progetto (studio clinico farmacologico), dichiarando pertanto che *"la sperimentazione in esame non prevede l'utilizzo di strumenti di intelligenza artificiale, né di logiche predittive basate su sistemi di machine learning"*.

Persistendo specifici profili di criticità, con nota del 3 marzo 2023 (prot. n. 38163), è stato formulato un supplemento istruttorio con particolare riferimento alle basi giuridiche del trattamento, ai tempi di conservazione e al rapporto tra l'algoritmo che si intende sviluppare nella prima fase dello Studio oggetto dell'istanza di consultazione preventiva in esame per la creazione di un modello matematico predittivo e lo strumento di intelligenza artificiale che si intenderebbe utilizzare nella fase 2 dello Studio.

In relazione alle basi giuridiche del trattamento, l'Azienda ha confermato come il *"Bando pubblico competitivo "ERAPerMed Joint Translational Call 2020" [...]"non possa assurgere a presupposto di liceità dei dati personali dei pazienti che saranno arruolati allo Studio in parola"* ciò anche in quanto *"il suddetto bando non*

*contiene indicazioni di misure appropriate e specifiche per la tutela dei diritti fondamentali degli interessati". La condizione di liceità del trattamento va quindi rinvenuta nell'art. 9, par. 2, lett. j) del Regolamento e nell'art. 110 del Codice (nota del 27 marzo 2023).*

Con riferimento all'impossibilità di informare gli interessati e di acquisirne un valido consenso, l'Azienda ha evidenziato che lo Studio ha ad oggetto l'analisi di un tumore raro, il carcinoma laringeo e che a causa dell'aggressività della malattia, la maggior parte dei pazienti che si intenderebbero arruolare per la raccolta retrospettiva dei dati, risultano essere deceduti; in particolare, presso il centro coordinatore il 46,67% di detti pazienti risultano deceduti.

Inoltre, il centro coordinatore (rispetto ai propri potenziali soggetti arruolabili allo Studio) ha rappresentato di aver *"tentato di contattare alcuni pazienti che sono risultati non rintracciabili; in particolare, il centro ha tentato di raggiungere detti pazienti tramite i recapiti che gli stessi avevano comunicato all'ASST nell'ambito del percorso di cura a cui erano stati sottoposti, senza successo [...]"*.

In merito alle *"connessioni tra l'algoritmo e lo strumento di intelligenza artificiale"*, l'Azienda ha confermato che *"la Fase 1 dello Studio non prevede l'applicazione e/o l'utilizzo di sistemi atti a trattare dati personali secondo logiche di intelligenza artificiale"* che saranno utilizzati esclusivamente della Fase 2 dello Studio. Allo stato, quindi, l'Azienda non è in grado di dettagliare le caratteristiche che i sistemi di IA potranno assumere nella successiva fase del Progetto. L'Azienda ha in ogni caso rappresentato *"le logiche e i presupposti che condurranno all'individuazione e definizione del modello predittivo (algoritmo) come esito della conclusione"* della fase 1 dello Studio.

Nello specifico, *"i dati raccolti saranno analizzati con modelli statistici [...] volti ad analizzare "le caratteristiche radiomiche e genomiche con metodi univariati"*. Ciò consentirà un'analisi a *"cluster"* che identificherà gruppi di caratteristiche radiomiche o di espressione genica con similitudini. *"Successivamente, si applicheranno modelli di multivariata per considerare associazioni tra variabili e potenziali effetti confondenti"*. L'Azienda ha evidenziato inoltre che molteplici analisi saranno effettuate sui dati: *"[...] analisi multivariate (MANOVA), analisi di regressione multivariate (stepwise feature selection), Principal Component Analysis (PCA), Canonical Correspondence Analysis (CCA) e Partial Least Square Discriminant Analysis (PLS-DA). Sarà quindi misurata la performance dell'algoritmo in termini di area al di sotto la "receiver operating characteristic" (ROC), con relative sensibilità, specificità e precisione. A seguito del completamento dell'analisi statistica, impiegando i metodi statistici di cui sopra, sarà possibile definire un modello predittivo (algoritmo) in grado di differenziare, in base a tutte le variabili analizzate ed oggetto di analisi statistica, quei pazienti che risponderanno o non al trattamento. Tale algoritmo, se si dimostrerà statisticamente significativo, sarà implementato/utilizzato per la realizzazione della Fase 2 del Progetto. In ordine alle logiche di funzionamento dell'algoritmo è stato precisato che*

*“l’obiettivo è identificare in quali pazienti arruolati allo Studio sono presenti variabili (edite e visibili dall’analisi) che fanno ipotizzare quale trattamento non funziona/ha più possibilità di funzionare rispetto alla patologia cui sono affetti”. Inoltre, con particolare riferimento alla qualità dei dati e alla loro trasparenza, l’Azienda ha rappresentato che “la qualità del dato viene confermata con l’elaborazione statistica” e che “esistono metodi di valutazione della qualità del dato che consistono – ad esempio – nella valutazione del numero di c.d. “data missing”, ovvero nel valutare la presenza di intervalli numerici “inverosimili”; generalmente a seguito di questa prima verifica, vengono generate query ai Centri partecipanti per la verifica dal dato fonte (cartella clinica). Alla luce di tali argomentazioni l’Azienda ha da ultimo escluso che nella Fase 1 del Progetto, si rientri nelle ipotesi previste nel “documento del Comitato Nazionale per la biosicurezza [...] che sarà invece tenuto in debita considerazione nella stesura dei processi del trial clinico, quale Fase 2 del Progetto”.*

Sui tempi di conservazione dei dati, l’Azienda ha dichiarato che procederà *“a ridurre la data retention dei dati personali dello Studio a 10 (dieci) anni, ossia l’arco temporale necessario per il raggiungimento delle finalità del Progetto”.*

L’Azienda ha infine trasmesso i nuovi testi delle informative redatte ai sensi degli artt. 13 e 14 del Regolamento *“ulteriormente revisionati al fine di renderli maggiormente intelleggibili e coerenti con il dettato normativo”* nell’ambito dei quali è stato indicato in particolare il ruolo del Centro partecipante in qualità di autonomo titolare del trattamento ed è stato indicato, come sopra anticipato, in 10 anni il periodo di conservazione dei dati.

## **2. La normativa applicabile**

Il trattamento di dati personali deve avvenire nel rispetto della normativa applicabile in materia di protezione dei dati personali.

In base al Regolamento, i dati personali devono essere trattati *“in modo lecito, corretto e trasparente nei confronti dell’interessato”* (principio di «liceità, correttezza e trasparenza») (art. 5, par. 1, lett. a) del Regolamento).

Con specifico riferimento alle particolari categorie di dati, tra cui rientrano i dati sulla salute, l’art. 9 del Regolamento sancisce un generale divieto al trattamento di tali dati a meno che non ricorra una delle specifiche esenzioni a tale divieto (art. 9, par. 2).

In tale quadro, il trattamento di dati personali per scopi di ricerca scientifica deve essere effettuato altresì nel rispetto delle Prescrizioni e delle Regole deontologiche che costituiscono condizione essenziale di liceità e correttezza dei trattamenti (art. 2-*quater* del Codice e art. 21, comma 5, del d.lgs. 10 agosto 2018, n. 101).

Nello specifico rileva, in particolare, l’art. 110 del Codice che riguarda la ricerca medica, biomedica ed epidemiologica e dispone che *“Il consenso dell’interessato per il trattamento dei dati relativi alla salute, a fini di ricerca scientifica*



*in campo medico, biomedico o epidemiologico, non è necessario quando la ricerca è effettuata in base a disposizioni di legge o di regolamento o al diritto dell'Unione europea in conformità all'articolo 9, paragrafo 2, lettera j), del Regolamento, ivi incluso il caso in cui la ricerca rientra in un programma di ricerca biomedica o sanitaria previsto ai sensi dell'articolo 12-bis del decreto legislativo 30 dicembre 1992, n. 502, ed è condotta e resa pubblica una valutazione d'impatto ai sensi degli articoli 35 e 36 del Regolamento. Il consenso non è inoltre necessario quando, a causa di particolari ragioni, informare gli interessati risulta impossibile o implica uno sforzo sproporzionato, oppure rischia di rendere impossibile o di pregiudicare gravemente il conseguimento delle finalità della ricerca. In tali casi, il titolare del trattamento adotta misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, il programma di ricerca è oggetto di motivato parere favorevole del competente comitato etico a livello territoriale e deve essere sottoposto a preventiva consultazione del Garante ai sensi dell'articolo 36 del Regolamento".*

In relazione alla valutazione di impatto si segnalano le *"Linee guida concernenti la valutazione di impatto sulla protezione dei dati nonché i criteri per stabilire se un trattamento "possa presentare un rischio elevato" ai sensi del regolamento 2016/679"* – adottate dal Gruppo art. 29 il 4 aprile 2017 come modificate e adottate da ultimo il 4 ottobre 2017.

Il principio di limitazione della conservazione dei dati impone che essi siano *"conservati in una forma che consenta l'identificazione degli interessati per un arco di tempo non superiore al conseguimento delle finalità per le quali sono trattati; i dati personali possono essere conservati per periodi più lunghi a condizione che siano trattati esclusivamente a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici, conformemente all'articolo 89, paragrafo 1, fatta salva l'attuazione di misure tecniche e organizzative adeguate richieste dal presente regolamento a tutela dei diritti e delle libertà dell'interessato"* (art. 5, par. 1, lett. e) del Regolamento).

I dati personali, inoltre, devono essere trattati nel rispetto del principio di trasparenza (art. 5, par. 1, lett. a) del Regolamento), fornendo preventivamente agli interessati le informazioni di cui all'art. 13 del Regolamento, in caso di dati raccolti direttamente presso di essi, ovvero ai sensi dell'art. 14, in caso di dati raccolti presso soggetti terzi.

I dati trattati per scopi di ricerca non possono essere utilizzati per prendere decisioni o provvedimenti relativamente all'interessato, né per trattamenti di dati per scopi di altra natura (*cfr.* art. 105 del Codice e Cons. 162 del Regolamento; Cons. 27 del Regolamento CE N. 2009/223 sulle statistiche europee e, a livello internazionale, l'art. 4 della Raccomandazione del Consiglio d'Europa n. R (97) relativa alla protezione dei personali raccolti e trattati per scopi statistici).

Ciò premesso, il titolare del trattamento in omaggio al principio di responsabilizzazione, deve conformarsi ed essere in grado di comprovare sia il

rispetto dei principi che degli adempimenti previsti dal Regolamento, sia di avere effettivamente tutelato il diritto alla protezione dei dati personali degli interessati fin dalla progettazione e per impostazione predefinita (artt. 5, par. 2, 24 e 25 del Regolamento).

In base al rinnovato quadro normativo in materia di protezione dei dati personali, si richiede, infatti, una valutazione ponderata di tutte le scelte connesse ai trattamenti di dati personali, dimostrabile sul piano logico attraverso specifiche motivazioni, volte all'individuazione di misure necessarie e proporzionate rispetto alla concreta efficacia del principio di volta in volta tutelato. In ossequio all'obbligo della protezione dei dati sin dalla progettazione, i titolari devono, inoltre, assumere una condotta attiva nell'applicazione dei principi, ponendosi l'obiettivo di ottenere un reale effetto di tutela (*cfr. Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, adottate il 13 novembre 2019 dal Comitato europeo per la protezione dei dati; provv. del Garante del 23 gennaio 2020, doc. web 9261093).

Con riferimento alla circostanza che i dati raccolti presso la piattaforma dovrebbero essere successivamente trattati attraverso logiche algoritmiche, con strumenti di intelligenza artificiale e con logiche predittive basate su sistemi di *machine learning*, si evidenziano gli specifici vincoli, in termini di protezione dei dati e di trasparenza, che dovranno essere rispettati in tal caso.

In via preliminare, si evidenzia che in base all'art. 22 del Regolamento *"L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona"*, ciò a meno che, con riferimento al trattamento delle particolari categorie di dati, esso si basi sul consenso degli interessati ovvero sia svolto sulla base di uno specifico presupposto normativo per motivi di interesse pubblico rilevante e *"siano in vigore misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato"* (art. 22, par. 4 del Regolamento; *cfr. anche Cons. 71*).

Al riguardo, si richiama la sentenza del Consiglio di stato (Cons. di St., sez. VI, 13 dicembre 2019, n. 8472) secondo la quale *"dal diritto sovranazionale emergono tre principi, da tenere in debita considerazione nell'esame e nell'utilizzo degli strumenti informatici. In primo luogo, il principio di conoscibilità, per cui ognuno ha diritto a conoscere l'esistenza di processi decisionali automatizzati che lo riguardino ed in questo caso a ricevere informazioni significative sulla logica utilizzata [...] il principio di non esclusività della decisione algoritmica [...]. In terzo luogo, dal considerando n. 71 del Regolamento 679/2016 il diritto europeo trae un ulteriore principio fondamentale, di non discriminazione algoritmica, secondo cui è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, mettendo in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori*

*e al fine di garantire la sicurezza dei dati personali, secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche [...]*" (cfr. da ultimo sent. Tar Campania, sez. III, n. 05119 del 11 novembre 2022).

Merita di essere richiamato anche quanto espresso nel parere congiunto del Comitato europeo per la protezione dei dati e il Garante europeo, n. 5/2021 sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale, del 21 aprile 2021), nel quale viene accolto con favore l'approccio basato sul rischio su cui si fonda la proposta.

Nel suddetto parere congiunto n. 5/2021 viene infatti evidenziato che i rischi per i diritti e le libertà fondamentali degli interessati derivanti dall'utilizzo di tali strumenti e le implicazioni per la protezione dei dati personali, sono molto rilevanti e viene sottolineato come tra i sistemi di IA ad alto rischio dovrebbero essere contemplati anche quelli utilizzati nell'ambito della ricerca medica.

Sotto altro profilo, merita evidenziarsi l'emendamento avanzato dal Parlamento europeo in relazione alla proposta di legge sull'intelligenza artificiale volta a precisare che *"È opportuno che il presente regolamento contribuisca a sostenere la ricerca e l'innovazione, non comprometta le attività di ricerca e sviluppo e rispetti la libertà della ricerca scientifica. È pertanto necessario escludere dal suo ambito di applicazione i sistemi di IA specificamente sviluppati al solo scopo di ricerca e sviluppo in ambito scientifico e garantire che il regolamento non incida altrimenti sulle attività scientifiche di ricerca e sviluppo relative ai sistemi di IA. In ogni circostanza, qualsiasi attività di ricerca e sviluppo dovrebbe essere svolta conformemente alla Carta, al diritto dell'Unione nonché al diritto nazionale"*<sup>1</sup>.

Generare contenuti, fare previsioni o adottare decisioni in maniera automatica, come fanno i sistemi di IA, per mezzo di tecniche di apprendimento automatico o regole di inferenza logica e probabilistica è cosa ben diversa rispetto alle modalità con cui queste stesse attività sono svolte dagli esseri umani attraverso il ragionamento creativo o teorico, nella piena consapevolezza della responsabilità delle relative conseguenze. Se da una parte, quindi, l'intelligenza artificiale amplierà significativamente la quantità di previsioni che si possono fare in molti ambiti – a cominciare dalle correlazioni quantificabili tra i dati-, dall'altra affidare solo alle macchine il compito di prendere decisioni sulla base di dati comporterà rischi per i diritti e le libertà delle persone che incideranno sulla loro vita privata e potrebbero nuocere a categorie sociali o persino a intere società. A tale riguardo, il Comitato e il Garante europeo sottolineano la centralità del concetto di sorveglianza (o supervisione) umana contenuto nella proposta di Regolamento (cfr. articolo 14) evidenziando a tale proposito che l'effettiva centralità degli esseri umani dovrebbe fondarsi su una supervisione altamente

---

1 Cfr. [https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236\\_IT.html](https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_IT.html)

qualificata e sulla liceità del trattamento, al fine di assicurare il rispetto del diritto di non essere assoggettato a una decisione basata esclusivamente su un trattamento automatizzato.

I trattamenti in esame, oggetto di analisi alla luce dei principi del Regolamento e del Codice, dovranno altresì conformarsi, una volta definita, alla cornice normativa in via di formazione nel contesto dell'Unione europea – sulla scia degli orientamenti etici (non di competenza specifica di questa Autorità) per un'IA affidabile del Gruppo indipendente di esperti istituito dalla Commissione europea – come pure ai principi oggetto di elaborazione in seno al Consiglio d'Europa improntati al *design*, sviluppo e applicazione di sistemi di intelligenza artificiale affidabile (*trustworthy AI*), rispettosi dei diritti fondamentali e delle libertà degli interessati.

A tale riguardo, si evidenzia altresì il recente Decalogo per la realizzazione di servizi sanitari nazionali attraverso sistemi di Intelligenza Artificiale adottato dal Garante il 10 ottobre 2023, dal quale possono essere comunque tratti utili spunti in relazione al trattamento dei dati personali effettuati attraverso strumenti di Intelligenza artificiale (doc. *web* n.9938038).

Sotto altro profilo, occorre evidenziare che la disciplina in materia di protezione dei dati personali non trova invece applicazione in relazione ai dati anonimi. A tale riguardo, si considerano anonime le “(...) informazioni che non si riferiscono a una persona fisica identificata o identificabile o a dati personali resi sufficientemente anonimi da impedire o da non consentire più l'identificazione dell'interessato”, ciò anche per i trattamenti svolti per finalità statistiche o di ricerca (cfr. considerando n. 26 del Regolamento). Il rischio di reidentificazione dell'interessato va, tuttavia, accuratamente valutato tenendo conto di “tutti i mezzi, [...], di cui il titolare del trattamento o un terzo può ragionevolmente avvalersi per identificare detta persona fisica direttamente o indirettamente. Per accertare la ragionevole probabilità di utilizzo dei mezzi per identificare la persona fisica, si dovrebbe prendere in considerazione l'insieme dei fattori obiettivi, tra cui i costi e il tempo necessario per l'identificazione, tenendo conto sia delle tecnologie disponibili al momento del trattamento, sia degli sviluppi tecnologici” (cfr. considerando n. 26 del Regolamento e WP29 *Opinion 05/2014 on Anonymisation techniques*, adottato il 10 aprile 2014). Un processo di anonimizzazione non può definirsi effettivamente tale qualora non risulti idoneo ad impedire che chiunque utilizzi tali dati, in combinazione con i mezzi “ragionevolmente disponibili”, possa:

1. isolare una persona in un gruppo (*single-out*);
2. collegare un dato anonimizzato a dati riferibili a una persona presenti in un distinto insieme di dati (*linkability*);
3. dedurre nuove informazioni riferibili a una persona da un dato anonimizzato (*inference*).

### **3. Le valutazioni dell'Autorità**

#### **3.1. Liceità e correttezza del trattamento dei dati personali**

L'Azienda promotrice dello Studio e titolare del trattamento, come previsto dall'art. 110 del Codice e 36 del Regolamento, ha presentato al Garante il protocollo dello Studio e la valutazione di impatto sulla protezione dei dati personali connessa ai trattamenti necessari per la realizzazione dello stesso, svolta ai sensi dell'art. 35 del Regolamento.

Al riguardo, il Garante ritiene che essa abbia chiarito l'ambito oggettivo dell'istanza di consultazione preventiva che si riferisce esclusivamente alla raccolta retrospettiva di dati necessari, per la Fase 1 del progetto europeo, presentato in risposta al bando intitolato "*A.I. for new signatures and models for tailored organ preservation approaches in laryngeal and hypopharyngeal cancer*" (Intelligenza artificiale per la definizione di nuove *signature* e modelli per la personalizzazione delle strategie di preservazione d'organo del cancro laringeo e ipofaringeo - PRESERVE").

Il Garante ritiene che l'Azienda abbia correttamente individuato le basi giuridiche del trattamento dei dati personali riferiti ai soggetti deceduti o non contattabili nella procedura di consultazione preventiva, ai sensi dell'art. 110 del Codice, e per i pazienti contattabili nel relativo consenso (cfr. anche artt. 9, par. 2, lett. a) e 22, par. 4) del Regolamento).

L'Azienda ha inoltre rappresentato in maniera esaustiva i motivi che giustificano l'impossibilità di riuscire ad informare gli interessati e ad acquisirne il relativo consenso, come previsto al punto 5.3 delle Prescrizioni.

L'Azienda, infatti, con dichiarazioni della cui veridicità risponde penalmente ai sensi dell'art 168 del Codice, ha evidenziato che l'impossibilità di acquisire il consenso è correlata all'aggressività della malattia oggetto di indagine e all'elevata incidenza di mortalità dei pazienti che si intendono arruolare.

L'Azienda ha infatti condotto una verifica sui pazienti potenzialmente eleggibili per lo Studio al fine di valutare la percentuale dei pazienti vivi, deceduti e non eleggibili, al fine di poter ipotizzare una stima applicabile a tutti i centri partecipanti. Nello specifico, sono state analizzate n. 60 cartelle cliniche di pazienti affetti da carcinoma a cellule squamose localmente avanzato della laringe e ipofaringe. Dalla verifica condotta è emerso che il 46,67% dei pazienti era deceduto, il 23,33% dei pazienti avrebbe dovuto essere escluso in quanto non eleggibili e il 30% di pazienti era ancora in vita.

Inoltre, per quanto riguarda i pazienti in vita, sono state rappresentate difficoltà oggettive in ordine al reperimento dei recapiti attuali di contatto di tali pazienti, visto anche il lasso di tempo trascorso dal momento in cui i dati riferiti agli interessati sono stati originariamente raccolti. Al riguardo, l'Azienda ha dichiarato che "*(rispetto ai propri potenziali soggetti arruolabili allo studio) ha tentato*

*di contattare alcuni pazienti che sono risultati non rintracciabili; in particolare, il centro ha tentato di raggiungere detti pazienti tramite i recapiti che gli stessi avevano comunicato all'ASST nell'ambito del percorso di cura a cui erano stati sottoposti, senza successo.*

Sul punto si ritiene necessario che l'Azienda tenga traccia dei tentativi di contatto effettuati attraverso uno dei canali indicati al punto 5.3 delle prescrizioni (anche attraverso la verifica dello stato in vita, la consultazione dei dati riportati nella documentazione clinica, l'impiego dei recapiti telefonici eventualmente forniti, l'acquisizione dei dati di contatto presso l'anagrafe degli assistiti o tramite l'Ufficio di Anagrafe del Comune.

A tale riguardo, si ritiene inoltre necessario che l'Azienda svolga almeno tre tentativi di contatto in quanto, tenuto conto della natura giuridica del titolare del trattamento e dell'ampiezza del campione, costituiscono uno sforzo ragionevole e proporzionato.

Inoltre, in conformità alla disposizione di cui al secondo periodo del comma 1, dell'art. 110 del Codice, in base alla quale il programma di ricerca deve essere oggetto di motivato parere favorevole del competente comitato etico a livello territoriale, tenuto conto che l'Azienda ha prodotto in atti il parere del comitato etico di Brescia, nella sua versione definitiva del 5 marzo 2021, resta fermo che i Centri partecipanti coinvolti nello Studio potranno dare inizio ai trattamenti solo dopo l'ottenimento dei pareri favorevoli dei rispettivi comitati etici territorialmente competenti, in quanto condizione di liceità e correttezza del trattamento dei dati personali per le finalità in esame, laddove non sia possibile acquisire il consenso degli interessati (art. 110 del Codice; provv. n. 202 del 29 ottobre 2020, doc. *web* 9517401 e provv. n. 406 del 1° novembre 2021, doc. *web* 9731827).

Tenuto conto che il presente provvedimento riguarda solo la fase 1 del progetto, volta esclusivamente allo "sviluppo di un algoritmo predittivo" il Garante ritiene che l'Azienda abbia individuato idonee garanzie al fine di assicurare la qualità del dato per l'effettiva applicazione del principio di correttezza.

Sotto tale profilo, si valuta quindi positivamente, da una parte, la motivata indispensabilità anche di informazioni riferite a soggetti deceduti e non contattabili in assenza delle quali il campione selezionato sarebbe incompleto creando di conseguenza, come rappresentato, possibili *bias* nello sviluppo dell'algoritmo. Spetta infatti all'intervento umano intervenire nella selezione delle informazioni necessarie ad "addestrare" l'algoritmo, ciò tenuto anche conto che la qualità delle capacità predittive di un algoritmo varia in funzione di una pluralità di fattori, a partire dal numero, dalla qualità e dall'accuratezza dei dati di addestramento (profili oggetto della prima fase dello Studio).

Dall'altro canto, si apprezza la puntuale elencazione dei modelli matematici statistici utilizzati per elaborare i dati che, se resa conoscibile, favorisce altresì, in un'ottica di trasparenza dell'attività di ricerca, l'apertura del progetto ad eventuali osservazioni da parte della comunità scientifica, superando criticità connesse ad eventuali opacità delle logiche utilizzate (cfr. *A Preliminary Opinion on data protection and scientific research* del 6 gennaio 2020 e cfr. *infra* par. 3.2).

### **3.2 Le misure volte a garantire l'effettività del principio di trasparenza nei confronti degli arruolati allo Studio**

In via generale, si prende favorevolmente atto delle misure intraprese dall'Azienda per garantire l'effettività del principio di trasparenza nei confronti degli interessati e, in particolare visti i rilievi formulati dall'Ufficio, che quest'ultima abbia predisposto due distinti moduli di informativa agli interessati, ai sensi degli artt. 13 e 14 del Regolamento, quest'ultimo da rendere pubblico attraverso una specifica inserzione sul sito *internet* aziendale.

Con specifico riguardo alle modalità per fornire le informazioni agli interessati non contattabili o deceduti (in questo ultimo a caso beneficio dei soggetti di cui all'art. 2-*terdecies* del Codice), tenuto conto che lo Studio coinvolge n. 20 Centri partecipanti italiani, al fine di assicurare l'effettiva applicazione dei richiamati principi di correttezza e trasparenza, si ritiene necessario, che l'Azienda renda pubbliche le informazioni, ai sensi dell'art. 14 del Regolamento, per tutta la durata dello Studio anche attraverso una specifica inserzione sui siti *internet* istituzionali dei centri di sperimentazione coinvolti nello Studio, in una sezione facilmente accessibile.

Nel merito, si precisa in primo luogo che, diversamente da quanto indicato nelle predette informative, in relazione ai trattamenti in esame, gli interessati non godono del diritto di opposizione, ma dei più ampi diritti di revoca e cancellazione. Il diritto di opposizione infatti spetta agli interessati in relazione ai trattamenti necessari per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri, ovvero per il perseguimento di un legittimo interesse del titolare o di terzi, ai sensi dell'art. 6, par. 1, lett. e) e f), del Regolamento.

Si rende pertanto necessario che l'Azienda modifichi il testo delle informative sul trattamento dei dati personali, predisposte ai sensi degli artt. 13 e 14 del Regolamento, espungendo il riferimento al diritto di opposizione.

Sotto altro profilo, e in relazione a quanto osservato al precedente paragrafo 3.1, si evidenzia come l'implementazione di efficaci misure volte a garantire una sostanziale trasparenza dei trattamenti di dati personali effettuati attraverso logiche algoritmiche costituisca un elemento essenziale affinché lo sviluppo e l'applicazione di tali sistemi avvenga in maniera rispettosa dei diritti fondamentali e delle libertà degli interessati, contribuendo alla creazione di quel

clima di fiducia auspicato dal Regolamento nello sviluppo della economia digitale (cons. 7).

A tal fine, si ritiene necessario, in primo luogo, che nelle informative venga rappresentato con un linguaggio semplice e chiaro che il progetto di ricerca in è volto all'implementazione di un algoritmo.

Attesa tuttavia l'indiscutibile difficoltà per la maggior parte degli interessati di comprendere nel dettaglio le modalità di funzionamento di tali innovativi mezzi di elaborazione dei dati, la correttezza e trasparenza degli stessi passa anche per la trasparenza delle metodologie e logiche algoritmiche applicate che ne consenta un controllo diffuso e la verifica anche e soprattutto da parte di soggetti esperti (autorità di controllo, comunità scientifica).

Si ritiene quindi necessario che l'Azienda provveda alla pubblicazione sul proprio sito *internet* istituzionale della valutazione di impatto anche solo per estratto, integrata con la puntuale elencazione dei modelli matematici statistici utilizzati per elaborare i dati, di cui sopra, e con l'indicazione delle logiche algoritmiche utilizzate.

### **3.3. I ruoli dei soggetti coinvolti nello Studio e trasferimento dei dati in Paesi terzi**

L'Azienda ha dichiarato che partecipano allo Studio, in qualità di Centri partecipanti, i soggetti indicati nella documentazione in atti quali autonomi titolari del trattamento, così indicati nelle informative previste ai sensi degli artt. 13 e 14 del Regolamento. È stato inoltre dichiarato che parteciperà al progetto l'Università di Oslo, *Department of Informatics P.O*, che fornirà la piattaforma elettronica per la raccolta dei dati, all'uopo nominata responsabile del trattamento ai sensi dell'art. 28 del Regolamento e che il trasferimento dei dati alla predetta Università avverrà sulla base di adeguate garanzie quali l'adozione delle *Standard Contractual Clauses* (c.d. "SCC"), ai sensi dell'art. 46, par.2, lett. c) del Regolamento).

A tale riguardo, si rileva che la struttura organizzativa implementata per la realizzazione dello Studio appare tale da escludere, in astratto, che soggetti terzi non autorizzati possano essere coinvolti nelle operazioni di trattamento dei dati sulla salute dei pazienti arruolati nel richiamato Studio e che il trasferimento dei dati all'Università di Oslo sia conforme alle richiamate disposizioni in materia di protezione dei dati personali.

### **3.4 I termini di conservazione e anonimizzazione dei dati personali**

Con riferimento ai tempi di conservazione dei dati si prende favorevolmente atto che l'Azienda li abbia da ultimo individuati in 10 anni dalla conclusione dello Studio, motivando in maniera esaustiva le ragioni per cui è necessario conservare i dati per tale periodo, ossia *l'arco temporale necessario per il*



*raggiungimento delle finalità del Progetto* che comprendono anche le successive fasi di analisi volte alla verifica e controllo dei dati e dei risultati della ricerca.

A tale riguardo, si ribadisce che il riferimento alla conservazione dei dati personali dello Studio per un periodo di 25 anni, in analogia a quanto previsto per le sperimentazioni cliniche dei farmaci, deve considerarsi non solo inconferente, non trattandosi nel caso di specie di tale tipologia di studio, ma anche erroneo.

Né rileva, sul punto, la circostanza che tale termine sarebbe previsto nel “Massimario di scarto” della Regione Lombardia in relazione alle sperimentazioni cliniche dei farmaci, al quale l’Azienda sarebbe soggetta in quanto pubblica amministrazione.

Sul punto, sotto un primo profilo, si osserva che il “Massimario di scarto” non potrebbe infatti indicare, per i richiamati studi, tempi di conservazione differenti da 25 anni, in quanto essi sono indicati nella specifica normativa di settore (art. 58 del Regolamento (UE) n. 536/2014 del Parlamento europeo e del Consiglio, del 16 aprile 2014, sulla sperimentazione clinica di medicinali per uso umano e che abroga la direttiva 2001/20/CE).

In secondo luogo, si ricorda che il Regolamento pone un generico divieto al trattamento -ivi inclusa pertanto la conservazione (art. 4, n. 2)- delle particolari categorie di dati, salve le deroghe di cui all’art. 9, par. 2 del Regolamento stesso.

Ciò rende la normativa che legittima l’uso di tale tipologia di informazioni eccezionale e come tale non passibile di interpretazioni analogiche o estensive.

Sotto altro profilo, si segnala che la documentazione raccolta presso i centri partecipanti è già conservata e archiviata da tali enti in base alla normativa di settore (*cfr.* Circolare del Ministero Sanità 19 dicembre 1986, n. 61; D.P.R. 30 settembre 1963, n. 1409; d.lgs. 22/01/2004, n. 42 recante Codice dei beni culturali e del paesaggio).

L’Azienda ha da ultimo dichiarato che i dati personali oggetto della sperimentazione clinica non vengono in alcun caso diffusi. Essi potranno essere eventualmente diffusi in forma rigorosamente anonima ed aggregata, ad esempio attraverso pubblicazioni scientifiche, statistiche e convegni scientifici. È stato chiarito in ogni caso che al termine del trattamento i dati verranno resi anonimi.

In relazione a tale profilo, tenuto conto del generico riferimento alla sola tecnica di aggregazione dei dati quale misura per l’anonimizzazione degli stessi al termine del trattamento o in caso di diffusione, il Garante ribadisce che un numero elevato di statistiche aggregate aumenta il potere identificativo di ciascuna di esse, fino alla possibile completa ricostruzione di un dataset (cd “*reconstruction attack*”). Pertanto, al fine di evitare tale rischio, è necessario che il numero delle statistiche oggetto di diffusione sia significativamente inferiore

rispetto al numero delle variabili che si intendono divulgare. In altri termini, assicurando la diffusione di un numero contenuto di statistiche, si evita che attraverso calcoli matematici, si possa pervenire all'identificazione dei singoli soggetti facenti parte del campione.

Tutto ciò premesso, si ritiene necessario che l'Azienda, al termine del periodo di conservazione dei dati per lo svolgimento dello Studio, e comunque in ipotesi di condivisione dei dati con soggetti terzi, assicuri che il numero di statistiche aggregate da rendere conoscibili sia significativamente inferiore rispetto al numero delle variabili considerate; ciò, al fine di scongiurare il rischio di ricostruzione di dati riferibili a singoli individui.

Inoltre, nell'ambito delle verifiche periodiche che il titolare del trattamento è comunque tenuto a svolgere anche in riferimento alla persistenza dell'efficacia delle misure di anonimizzazione dei dati e all'evoluzione tecnologica, si ritiene necessario che l'Azienda si impegni altresì a rimuovere ogni singolarità, qualora, con qualsiasi mezzo, ne venga a conoscenza in una fase successiva all'applicazione delle predette tecniche di anonimizzazione, e a tenere traccia di tali eventi, e a ripetere in tale circostanza la valutazione del rischio di re-identificazione alla luce delle cause che hanno determinato l'insorgenza di tale singolarità (cfr. sul punto, il Parere del 30 giugno 2022, doc. *web* 9791886 e del 7 marzo 2023, doc. *web* n. 9875254).

### **3.5 Le misure di sicurezza implementate**

Fermo quanto sopra osservato in ordine alle carenze concernenti il trattamento di dati personali attraverso logiche algoritmiche, si osserva che nella valutazione d'impatto è stata in particolare condotta un'analisi dei rischi connessi ai trattamenti di dati personali necessari al perseguimento dello scopo della ricerca.

Con specifico riferimento alla piattaforma *web based TSD*, utilizzata per il caricamento dei dati e gestita dall'Università di Oslo, è previsto che l'accesso sia riservato esclusivamente al personale autorizzato cui vengono rilasciate le credenziali personali inviate a mezzo di due distinte *email* e che i dati siano inseriti nella *e-cfr* in forma pseudonimizzata, ciò al fine di garantire il rispetto del principio di minimizzazione.

L'Azienda dispone anche di misure organizzative e di sicurezza informatica quali: di controllo accessi e di profilazione e autorizzazione degli utenti; il piano di *disaster recovery*, *antivirus/firewall*; *policy* di sicurezza anche dei documenti cartacei; *policy* di gestione di incidenti relativi alla sicurezza dei dati (*data breach*); istruzioni operative per le persone autorizzate; che sono informate sulla possibilità di incorrere in un processo disciplinare in caso di mancato rispetto delle *policy* in tema di trattamento dei dati personali.

L'Azienda inoltre ha previsto che sia *“richiesta la compliance di privacy by design e by default per l'acquisto e/o per lo sviluppo in outsourcing di sistemi informativi o di servizi affidati all'esterno e per i relativi aggiornamenti, compresi i rilasci di nuove versioni”*.

### **TUTTO CIO' PREMESSO IL GARANTE**

ai sensi dell'art. 110 del Codice e dell'art. 36 del Regolamento, esprime alla Azienda Socio Sanitaria Territoriale degli Spedali Civili di Brescia, Piazzale Spedali Civili, 1 25123 (Brescia), P.IVA 03775110988 C.F. 03775110988, parere favorevole in ordine al trattamento dei dati personali per finalità di ricerca medica, biomedica ed epidemiologica, riferiti alla coorte di pazienti deceduti o non contattabili nello studio retrospettivo osservazionale denominato *“Intelligenza artificiale per la definizione di nuove signature e modelli per la personalizzazione delle strategie di preservazione d'organo del cancro laringeo e ipofaringeo - PRESERVE”* a condizione che :

- a) svolga almeno tre tentativi prima di attestare la non contattabilità dei pazienti e ne tenga traccia nella documentazione dello Studio (punto 3.1);
- b) modifichi il testo delle informative sul trattamento dei dati personali, predisposte ai sensi degli artt. 13 e 14 del Regolamento, espungendo il riferimento al diritto di opposizione (par. 3.2);
- c) nelle informative venga rappresentato con un linguaggio semplice e chiaro che il progetto di ricerca è volto all'implementazione di un algoritmo (par. 3.2);
- d) pubblichi sul proprio sito *internet* istituzionale la valutazione di impatto, anche solo per estratto, integrata con la puntuale elencazione dei modelli matematici statistici utilizzati per elaborare i dati e con l'indicazione delle logiche algoritmiche utilizzate (par. 3.2.);
- e) al termine del periodo di conservazione dei dati per lo svolgimento dello Studio, e comunque in ipotesi di condivisione dei dati con soggetti terzi, assicuri che il numero di statistiche aggregate da rendere conoscibili sia significativamente inferiore rispetto al numero delle variabili considerate (par 3.4.);
- f) rimuova ogni singolarità, qualora, con qualsiasi mezzo, ne venga a conoscenza in una fase successiva all'applicazione delle predette tecniche di anonimizzazione, e a tenere traccia di tali eventi, e a ripetere in tale circostanza la valutazione del rischio di re-

identificazione alla luce delle cause che hanno determinato l'insorgenza di tale singolarità (par. 3.4).

Ai sensi dell'art. 78 del Regolamento, degli artt. 152 del Codice e 10 del d.lgs. n. 150/2011, avverso il presente provvedimento è possibile proporre ricorso dinnanzi all'Autorità giudiziaria ordinaria, a pena di inammissibilità, entro trenta giorni dalla data di comunicazione del provvedimento stesso ovvero entro sessanta giorni se il ricorrente risiede all'estero.

Roma, 24 gennaio 2024

IL PRESIDENTE

IL RELATORE

IL SEGRETARIO GENERALE