VALUTAZIONE DI IMPATTO ai sensi dell'art. 35 del GDPR per lo studio

"La rete italiana OMOP per la ricerca clinica con Real World Data sanitari: uno studio di fattibilità" - OMOP_RWD_IT"

STORIA DELLE MODIFICHE AL DOCUMENTO					
Versione	**Data**	**Descrizione Modifica**	**Redatto da**	**Rivisto da**	**Approvato da**
1.0	03/10/24		Lucia Sacchi	Ufficio Legale e DPO UNIPV	Lucia Sacchi

INDICE

1 DEFINIZIONI 4
2 CONTESTO 5
2.1 PANORAMICA DEL TRATTAMENTO 5
2.2 DATI, PROCESSI E RISORSE DI SUPPPORTO 7
3 PRINCIPI FONDAMENTALI 8
3.1 PROPORZIONALITA' E NECESSITA' 8
3.2 MISURA A TUTELA DEI DIRITTI DEGLI INTERESSATI 9
4 MISURE ESISTENTI E PIANIFICATE 11
4.1 Crittografia 11
4.2 Anonimizzazione 11
4.3 Controllo degli accessi logici 11
4.4 Partizionamento 12
4.5 Tracciabilità 12
4.6 Archiviazione 13
4.7 Minimizzazione dei dati 13
4.8 Lotta contro il malware 13
4.9 Backup 13
4.10 Gestione postazioni 14
4.11 Contratto con il responsabile del trattamento 14
4.12 Sicurezza dei canali informativi 14
4.13 Gestione del personale 14
5 RISCHI 15
5.1 ACCESSO ILLEGITTIMO AI DATI 15
5.2 MODIFICHE INDESIDERATE DEI DATI 16
5.3 PERDITA DI DATI 17
5.4 PANORAMICA DEI RISCHI 18
6 CONVALIDA 19
6.1 MAPPATURA DEL RISCHIO 19
6.2 PIANO D'AZIONE 19

1 DEFINIZIONI

- **Studio**: protocollo “La rete italiana OMOP per la ricerca clinica con Real World Data sanitari: uno studio di fattibilità” - OMOP_RWD_IT”
- **OMOP**: Observational Medical Outcomes Partnership è una collaborazione pubblico-privata, presieduta dal FDA (Food and Drug Administration), che ha progettato l’OMOP Common Data Model (CDM), un modello standardizzato (basato su SNOMED, ICD9-10, RxNORM, LOINC, ecc.) per la memorizzazione di Real World Data e sviluppato per facilitare la generazione di evidenze scientifiche tramite studi osservazionali.
- **OHDSI**: Observational Health Data Sciences and Informatics è una collaborazione internazionale interdisciplinare nata nel 2014 per aggiornare l’OMOP-CDM e sviluppare tool di supporto (https://www.ohdsi.org/). Con centinaia di ricercatori provenienti da svariati paesi e database OMOP per oltre 950 milioni di pazienti in tutto il mondo, OHDSI consente alla comunità dei ricercatori di generare in modo collaborativo studi osservazionali per promuovere decisioni sanitarie e cure migliori.
- **OHDSI ITALIA:** sottogruppo di dell’organizzazione OHDSI Europe, formato da ricercatori iscritti ad OHDSI afferenti ad enti italiani
- **Promotore**: Ente promotore dello studio (Università degli Studi di Pavia (UNIPV))
- **Centri partecipanti:** Centri clinici coinvolti nello Studio
- **Database OMOP**: Repository di dati basato sul modello OMOP e presente all’interno di ciascun Centro partecipante allo Studio

2 CONTESTO

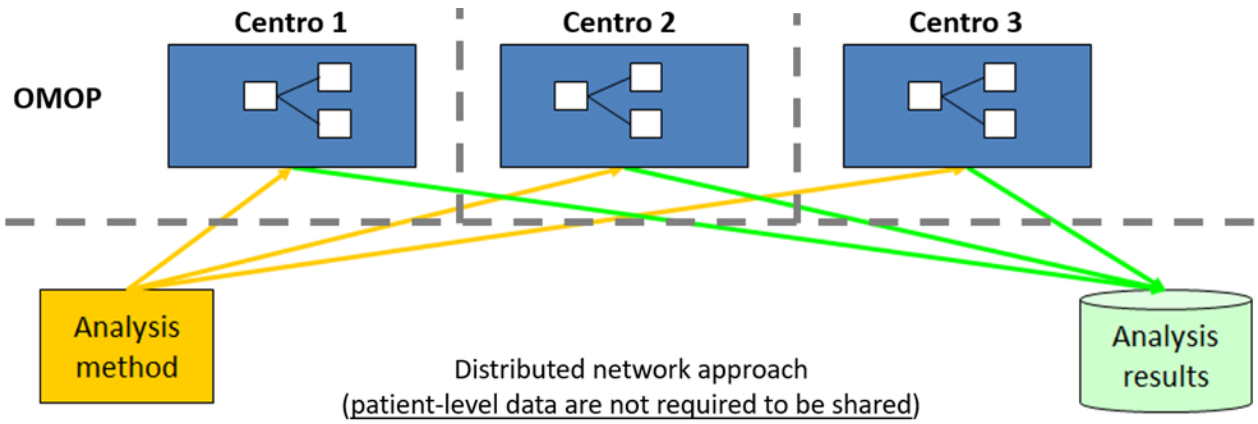
2.1 PANORAMICA DEL TRATTAMENTO

2.1.1 Quale è il trattamento in considerazione?

Oggetto del presente atto di DPIA è il trattamento dei dati che avverrà nell'ambito dello Studio.

Lo Studio verrà condotto seguendo il paradigma OHDSI mostrato nella figura sottostante che prevede studi osservazionali di tipo federato/distribuito. Il paradigma prevede che ogni ente partecipante allo studio abbiamo creato un database OMOP con dati di pazienti pseudonomizzati a partire dal proprio sistema informativo ospedaliero (SIO) e che tale infrastruttura sia manutenuta. Tale trattamento non è oggetto della presente DPIA, che ogni ente approva per quanto di competenza, ma viene gestito dal singolo ente in conformità alla norma vigente in materia di trattamento dati personali per finalità di ricerca scientifica. Nello specifico, ogni ente partecipante si impegna a pubblicare sul proprio sito web la Valutazione d'impatto, comprensiva di tutte le misure necessarie e propedeutiche al trattamento dei dati per la ricerca volte ad assicurare adeguate garanzie a tutela degli interessati. Una volta che ogni ente partecipante ha al proprio interno un database OMOP, lo studio procede come segue.

1. L'avvio dello studio implica l'implementazione di uno script/metodo di analisi (solitamente tramite il package R) da parte del promotore. Il promotore dovrà rendere pubblico o condividere il metodo e il relativo protocollo con gli enti partecipanti. L'utilizzo dello standard OMOP garantirà la completa compatibilità dello script con tutti gli enti della rete.
2. Un ente partecipante dovrà eseguire il metodo/script sul proprio database OMOP
3. L'ente partecipante dovrà quindi condividere al promotore solo i **risultati composti da dati aggregati** e non individuali relativi ai pazienti.
4. Il Promotore si occuperà di analizzare i risultati di tutti gli enti partecipanti e produrre i risultati finali dello studio.



Paradigma per gli studi OHDSI

Nello specifico l'obiettivo dello Studio è quello di descrivere e verificare l'operatività degli attuali *database OMOP* dei centri coinvolti nel nodo italiano OHDSI. Questo permetterà di avere una misura delle potenzialità di sviluppo di studi epidemiologici nazionali e internazionali utilizzando dati standardizzati. Una volta testate le operatività e descritto il set di dati standardizzati sarà possibile utilizzare tale framework come strumento di ricerca epidemiologica su cui basare le future strategie di Sanità pubblica. La descrizione dettagliata dei

dati disponibili nel nodo italiano OHDSI permetterà anche di individuare possibili aree di ricerca da presentare a Network europei come EHDEN (European Health Data & Evidence Network) e DARWIN EU (Data Analysis and Real World Interrogation Network), estendendo la ricerca a livello europeo. Nel caso dei centri finanziati dal progetto EHDEN, la descrizione dettagliata dei database diverrà parte integrante del report pubblicato sul portale del progetto EHDEN, anch'esso finalizzato a caratterizzare di dati inseriti nel database OMOP di ciascun centro con l'obiettivo di potenziare la ricerca epidemiologica collaborativa su dati standardizzati a livello europeo.

In particolare, seguendo il paradigma OHDSI, il trattamento seguirà questi passi

1. UNIPV, in qualità di promotore svilupperà, uno script R che eseguirà le estrazioni e aggregazioni previste dallo Studio in ciascun centro partecipante, e pubblicherà lo script dal repository GitHub dello Studio, inviando via email ai referenti dei Centri partecipanti le istruzioni per partecipare allo Studio.
2. Il referente operativo di ogni centro partecipante scaricherà dal repository GitHub lo script R e lo eseguirà sul proprio OMOP. Lo script produrrà per ogni centro alcuni file di risultati descrittivi dei dati contenuti nel database OMOP
3. Il referente operativo di ogni centro partecipante dovrà quindi caricare i risultati per il proprio centro su un'apposita cartella OneDrive condivisa al centro da UNIPV.
4. UNIPV si occuperà di analizzare i risultati di tutti i Centri partecipanti e produrre i risultati finali dello studio

Un dettaglio dei passaggi è descritto nell'Allegato 1 - OMOP_RWD_IT-flusso.

2.1.2 Quali sono le responsabilità connesse al trattamento?

- Principal Investigator: Ing. Lucia Sacchi
- Titolare del trattamento: Università degli Studi di Pavia (UNIPV)
- Responsabile esterno del trattamento:
    - Microsoft OneDrive
    - BIOMERIS SRL (supporto a creazione dello script e all'analisi dei risultati)
- Soggetti interessati: Pazienti
- Centri coinvolti:
    - Fondazione IRCCS Policlinico San Matteo, Pavia
    - Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico
    - IRCCS Policlinico San Donato
    - Istituto Maugeri IRCCS
    - Istituto Tumori della Romagna IRST IRCCS
    - Fondazione Istituto Nazionale dei Tumori
    - Fondazione Toscana Gabriele Monasterio per la Ricerca Medica e di Sanità Pubblica
    - Ospedale Pediatrico Bambino Gesù, Roma
    - Azienda Ospedaliero-Universitaria di Parma
    - Policlinico S.Orsola-Malpighi, Bologna
    - Azienda Ospedaliera Universitaria Integrata Verona
    - ASST Papa Giovanni XXIII di Bergamo
    - Casa di Cura Privata del Policlinico, Milano
    - Fondazione IRCCS Istituto Neurologico Carlo Besta

2.1.3 Ci sono standard applicabili al trattamento?

I dati verranno utilizzati nel rispetto delle Good Clinical Practice ("GCP") ossia lo standard internazionale di etica e qualità scientifica per progettare, condurre, registrare e relazionare gli studi clinici che coinvolgano soggetti umani (D.M. 15 luglio del 1997 e ss.mm.ii.). L'aderenza alle GCP garantisce pubblicamente non solo

la tutela dei diritti, della sicurezza e del benessere dei soggetti che partecipano allo studio, in conformità con i princìpi stabiliti dalla Dichiarazione di Helsinki dell'Associazione medica mondiale del giugno 1964, ma anche l'attendibilità dei dati relativi allo studio clinico.

Nella conduzione dello Studio sono, inoltre, rispettate le prescrizioni delle Linee guida sulla sperimentazione clinica adottate dal Garante per la protezione dei dati personali con Provvedimento 24 luglio 2008 n. 52 e delle Regole Deontologiche (Provvedimento 19 dicembre 2018 n. 515 e 9 maggio 2024 n. 298), nonché il Provvedimento n. 146 del 5 giugno 2019.

2.2 DATI, PROCESSI E RISORSE DI SUPPPORTO

2.2.1 Quali sono i dati trattati?

I dati personali oggetto del trattamento sono

- dati clinici dei pazienti arruolati per lo Studio, che equivalgono ai dati inseriti nei database OMOP dei centri partecipanti
- informazioni riguardanti il personale coinvolto nello Studio, sia relativo al Promotore sia ai centri partecipanti

Le categorie di Dati Personali oggetto della presente DPIA ricomprendono:

- sesso, data di nascita e luogo di residenza.
- dati antropometrici, dati relativi alla salute raccolti durante le visite, i ricoveri, gli accessi al pronto soccorso nonché gli esami e gli accertamenti effettuati presso la Struttura Sanitaria. In particolare, i dati riguardano diagnosi, interventi, terapie somministrate, esiti degli esami di laboratorio, dispositivi e campioni biologici.

Il Promotore tratterà tuttavia tali dati in forma aggregata.

2.2.2 Qual è il ciclo di vita del trattamento dei dati (descrizione funzionale)?

Il ciclo vita del trattamento è descritto in dettaglio in Allegato 1 - OMOP_RWD_IT-flusso

2.2.3 Quali sono le risorse di supporto ai dati?

- Database OMOP presso i centri partecipanti
- Tool OHDSI da utilizzare presso i centri partecipanti
- Script R per estrazione data aggregati
- GitHub per condivisione Script R ai centri partecipanti
- OneDrive di UNIPV dove i centri partecipanti condividono i dati aggregati con il Promotore
- Script R per l'analisi finale dei risultati

3 PRINCIPI FONDAMENTALI

3.1 PROPORZIONALITA' E NECESSITA'

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

Si. Si riportano gli obbiettivi dello Studio descritti nel Protocollo.

- Obbiettivo primario sarà valutare una serie di indicatori di performance, specialmente in termini di tempi di esecuzione, delle Strutture sanitarie afferenti al nodo OHDSI Italia nell'eseguire le attività propedeutiche a studi epidemiologici tramite il proprio data warehouse OMOP, in particolare valutando la capacità di utilizzare alcuni strumenti OHDSI descritti nella sezione "Metodologia e procedure".
  Per l'obiettivo primario, endpoint principale è il tempo di esecuzione dello script in R e del tempo di risposta da parte dei centri coinvolti. Saranno riportati i tempi mediani, con range interquartile, necessari a ciascun centro per svolgere le attività richieste.

- Obiettivo secondario è descrivere dettagliatamente il set di dati inclusi nel campione della popolazione italiana afferente agli enti sanitari del nodo in termini di numero di osservazioni, informazioni demografiche, patologie, esami erogati e prescrizioni effettuate.
  Per l'obiettivo secondario, la caratterizzazione della popolazione sarà eseguita sui seguenti endpoints:
  - sesso, età, anno nascita;
  - numero di osservazioni per dominio OMOP (condition, procedure, drug, measurement).

  In particolare, verranno calcolati i seguenti indicatori:
  - Malattie prevalenti;
  - Esami più erogati;
  - Farmaci più somministrati.

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

Prima di tutto è da sottolineare che il trattamento eseguito dal Promotore riguarda dati in forma aggregata. Sono messe in atto misure di sicurezza, descritte in seguito, che garantiscono che tali dati aggregati siano anonimizzati e di conseguenza non soggetti alle restrizioni previste dal GDPR Considerando 26.

3.1.3 I dati raccolti sono adeguati, pertinenti e limitati a quanto è necessario in relazione alle finalità per cui sono trattati (minimizzazione dei dati)?

I dati raccolti da Promotore soddisfano il criterio di minimizzazione dei dati in quanto sono raccolti solo dati in forma aggregata. Inoltre, per assicurarsi che tale dato sia anonimo saranno considerati conteggi non inferiori a 30 pazienti.

I dati raccolti sono infatti limitati agli scopi dello Studio. Infatti, per l'obbiettivo primario di verificare l'operatività dei Centri nell'utilizzo dei tool OHDSI si raccolgono i dati relativi al tempo di esecuzione dello script R e del tempo di risposta da parte dei centri coinvolti. Per l'obbiettivo secondario di descrivere i dati inclusi nei database OMOP del nodo italiano OHDSI, si raccolgono dati descrittivi aggregati in grado di profilare tali database prodotti da alcuni tool OHDSI (DQD, Achilles, CdmInspection, CatalogueExport)

3.1.4 I dati sono esatti e aggiornati?

L'esattezza dei dati raccolti dipenderà dall'attività di trasferimento e armonizzazione verso il database OMOP eseguita presso ogni Centro partecipante.

Il report di qualità generato tramite il tool DQD ed eseguito sul database OMOP nell'ambito dello Studio garantirà la qualità dei dati.

Nell'ambito delle analisi richieste dallo Studio non sarà strettamente necessario un aggiornamento dei dati alla situazione reale più recente. Lo Studio prevede di eseguire una e una sola "fotografia" del database OMOP indipendente dal suo stato di aggiornamento e dal peridio di osservazione dei dati presenti in ciascun database OMOP.

L'esattezza delle analisi eseguite dallo script R sviluppato dal Promotore è garantita anche dal fatto che tale script è basato su funzioni "standard" fornite e validate dalla community di OHDSI.

3.1.5 Qual è il periodo di conservazione dei dati?

I dati aggregati dei singoli Centri e i risultati finali saranno conservati dal Promotore per un periodo di 5 anni per permettere anche il loro sfruttamento per pianificare possibili studi futuri nell'ambito del nodo OHDSI Italia.

3.2 MISURA A TUTELA DEI DIRITTI DEGLI INTERESSATI

3.2.1 Come sono informati del trattamento gli interessati?

Ogni Centro partecipante ha gestito internamente e secondo la normativa vigente in materia di trattamento dati personali la modalità di informativa ai pazienti riguardante il riutilizzo dei loro dati a fini di ricerca oppure, nello specifico, di creazione del database OMOP finalizzato a facilitare futuri studi.

Nello specifico dello Studio in oggetto, verrà pubblicato dal Promotore sul proprio sito web un estratto della presente DPIA.

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

Utilizzando le informazioni presenti nei database di 5 dei centri partecipanti, è stato stimato che il numero medio sarà di circa 70000 pazienti all'anno per ciascun database OMOP e che la maggior parte dei database contengono 3 anni di dati. Mediamente ogni Centro partecipante "arruolerà" per lo Studio oltre 200 mila pazienti. Tali numeri implicano un'impossibilità organizzativa di sottoporre ad ogni paziente un consenso specifico per lo Studio (provvedimento recante Prescrizioni relative al trattamento dei dati personali effettuato per scopi di ricerca scientifica, ai sensi dell'art. 21, comma 1 del d.lgs. 10 agosto 2018, n. 101)

Tale impossibilità è facilmente dimostrabile per vari fattori: primo fra tutti l'impossibilità di avere a disposizione dati di contatto di tutti i pazienti che hanno avuto in passato accesso ad un Centro. L'impossibilità è dimostrabile anche ipotizzando un tempo di circa 15 min necessario per contattare un paziente e richiedere il consenso per lo Studio e considerando che il referente operativo per ogni Centro per lo Studio è una sola persona. Per tale attività il referente operativo di un Centro necessiterebbe mediamente oltre 1000 giornate di lavoro full-time incidendo in maniera sostanziale sull'obbiettivo dello Studio di avere una "fotografia" attuale dei database OMOP in Italia.

3.2.3 Come fanno gli interessati a esercitare i loro diritti di accesso e di portabilità dei dati?

Sui siti web dei centri partecipanti allo Studio verrà indicato un riferimento del Promotore da contattare da parte dei pazienti interessati ad esercitare i propri diritti.

3.2.4 Come fanno gli interessati a esercitare i loro diritti di rettifica e di cancellazione (diritto all'oblio)?

Ogni centro partecipante all'interno della propria informativa avrà specificato i diritti di rettifica e di cancellazione. Ogni centro partecipante sarà responsabile di aggiornare i dati aggregati risultanti e notificare il Promotore.

3.2.5 Come fanno gli interessati a esercitare i loro diritti di limitazione e di opposizione?

Ogni centro partecipante all'interno della propria informativa avrà specificato i diritti di limitazione e opposizione. Ogni centro partecipante sarà responsabile di aggiornare i dati aggregati risultanti e notificare il Promotore.

3.2.6 Gli obblighi dei responsabili del trattamento sono definiti con chiarezza e disciplinati da un contratto?

Il Protocollo dello Studio descrive nel dettaglio il trattamento eseguito dal Promotore.

In alcuni casi il Promotore stipula un accordo con il Centro partecipante per definire le modalità di conduzione dello studio osservazionale.

Per utilizzare SharePoint (e più in generale i servizi cloud di Microsoft, come Microsoft 365) UNIPV ha accettato una Data Processing Agreement (DPA) o Contratto di trattamento dei dati. Questa DPA è un accordo legale che specifica come Microsoft, in qualità di responsabile del trattamento (data processor), gestisce e protegge i dati personali per conto del cliente, che è considerato il titolare del trattamento (data controller), in conformità al GDPR e ad altre normative sulla protezione dei dati. L'ultima versione di tale DPA è accessibile a questo link:

https://www.microsoft.com/licensing/docs/documents/download/MicrosoftProductandServicesDPA(WW)(Italian)(Jan022024)(CR).docx

Il Dipartimento di Ingegneria Industriale e dell'Informazione dell'Università di Pavia e BIOMERIS sottoscrivono una DPA per definire il ruolo di BIOMERIS in qualità di responsabile esterno del trattamento.

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

Come già descritto, tutti i dati trattati al di fuori dei Centri Partecipanti sono aggregati

Il flusso di dati previsto non implica nessun trasferimento di dati al di fuori dell'Unione europea; tuttavia, l'utilizzo di SharePoint da parte di organizzazioni italiane come UNIPV può implicare una trasmissione di dati al di fuori dell'Unione Europea, ma Microsoft adotta misure per garantire che tali trasferimenti siano conformi al GDPR e alle normative europee sulla protezione dei dati.

4 MISURE ESISTENTI E PIANIFICATE

4.1 Crittografia

Database OMOP

Crittografia dei dati in transito da/verso database OMOP: i dati vengono crittografati durante la trasmissione tra le diverse componenti del sistema, ad esempio quando vengono inviati dati da un'applicazione web a un server di database OMOP. L'utilizzo di protocolli di crittografia come HTTPS (TLS > 1.2) per le comunicazioni web è essenziale per proteggere i dati durante il transito

Crittografia delle password per accesso a database OMOP tramite interfaccia web ATLAS:

Nel caso di Centri che sfruttano lo strato di autenticazione integrato in ATLAS, le password degli utenti vengono memorizzate sul database "Security" in formato criptato, utilizzando l'algoritmo di hash "bcrypt" noto per avere un'elevata resistenza agli attacchi di forza bruta e di dizionario. Le password degli utenti di PostgreSQL vengono memorizzate in tabelle di sistema, inaccessibili agli utenti creati per la piattaforma, utilizzando l'algoritmo di hash SHA256 con salt.

Nel caso invece di Centri che sfruttano lo strato di autenticazione fornito da Keycloak, oltre ad avere accesso a funzionalità più avanzate (come 2FA), le password vengono memorizzate tramite algoritmo PBKDF2 con utilizzo di salt e 27.500 iterazioni.

SharePoint:

La crittografia dei dati in SharePoint è gestita con:

- Crittografia a riposo (at rest): Dati archiviati crittografati con AES-256 e ulteriormente protetti da BitLocker.
- Crittografia in transito (in transit): Dati trasferiti protetti tramite TLS e connessioni HTTPS.
- Gestione delle chiavi: Microsoft gestisce le chiavi crittografiche, ma le aziende possono usare le proprie con Azure Key Vault.
- Customer Lockbox: Controllo sui permessi di accesso ai dati da parte di tecnici Microsoft.

4.2 Anonimizzazione

I dati sono raccolti in forma aggregata e per assicurarsi che tale dato sia anonimo saranno considerati conteggi non inferiori a 30 pazienti

4.3 Controllo degli accessi logici

Accesso al **database OMOP** da parte del referente operativo del Centro

- Tramite l'interfaccia web ATLAS con
  - l'autenticazione, che può essere gestita utilizzando l'interfaccia web ATLAS oppure tramite l'applicazione per autenticazioni Keycloak che può essere configurata con una two-factor authentication;
  - l'autorizzazione che viene gestita sempre all'interno della Piattaforma utilizzando l'interfaccia web ATLAS. Viene implementato un modello RBAC (Role-Based Access Control) dove i privilegi, relativi alle singole funzioni erogabili dal sistema rispetto ad ogni dataset in esso contenuto, vengono collegati a specifici ruoli che, in una fase successiva, sono assegnati agli utenti.
  - accesso diretto tramite utenza database specifica per l'accesso al CDM per poter svolgere analisi con i package R.

Il controllo degli accessi logici in **SharePoint** è gestito attraverso una combinazione di meccanismi di autenticazione, autorizzazione e gestione dei ruoli:

- Autenticazione: Gli utenti devono autenticarsi tramite Azure Active Directory (AAD), che supporta autenticazione multifattore (MFA) e Single Sign-On (SSO) per garantire l'accesso sicuro.
- Autorizzazione basata sui ruoli: Gli utenti ottengono permessi in base ai loro ruoli, con un modello di controllo degli accessi basato sui ruoli (RBAC) che limita ciò che possono visualizzare o modificare all'interno di SharePoint.
- Accesso condizionato: Integrato con AAD, SharePoint può limitare l'accesso in base a fattori come la posizione geografica, il dispositivo utilizzato e il rischio associato all'accesso.
- Condivisione esterna e link protetti: Gli utenti possono essere autorizzati a condividere file o documenti con utenti esterni in modo controllato e sicuro, richiedendo anche autenticazione aggiuntiva.

Questi meccanismi garantiscono che solo utenti autorizzati possano accedere e interagire con le risorse all'interno di SharePoint.

4.4 Partizionamento

Il referente operativo di ogni Centro ha accesso solo al CDM OMOP che è per sua natura pseudonimo ma non può accedere alla tabella di LOOKUP che permettono di re-identificare il paziente.

Il personale del Promotore non ha alcun accesso dal database OMOP dei singoli centri ma si dedica solo all'analisi dei risultati aggregati condivisi dai centri.

4.5 Tracciabilità

La tracciabilità nell'utilizzo dei database **OMOP** dei Centri viene garantita attraverso diversi meccanismi:

- L'utilizzo di chiavi primarie e chiavi esterne per collegare le tabelle e i record del database in modo univoco e coerente.
- L'utilizzo di tabelle di log per registrare la data e l'ora di creazione, modifica o cancellazione dei record del database.
- L'utilizzo di tabelle di corrispondenza per indicare la fonte originale dei dati e il sistema da cui sono stati estratti o trasformati.
- L'utilizzo di controlli di qualità per valutare la completezza, l'accuratezza, la consistenza e la validità dei dati.
- L'utilizzo di log e audit trail per registrare le operazioni effettuate sul database da parte degli utenti o dei processi autorizzati.
- Il salvataggio della history giornaliera dei comandi di tutti gli utenti compresi gli Amministratori di Sistema, utilizzabile come log delle attività effettuate da terminale
- Periodicamente viene controllata la sicurezza dei server, attraverso la ricezione di un report automatico, che riporta l'avvenuta raccolta di tutte le misurazioni e i log da parte dei server e segnala eventuali anomalie.

La tracciabilità delle operazioni eseguite sulle cartelle **SharePoint** utilizzate per la condivisione dei risultati è garantita dal fatto che sono tracciate tutte le operazioni degli utenti (consultabili dall'amministratore). Nello specifico la tracciabilità delle operazioni in SharePoint è gestita tramite il logging e le funzionalità di audit che monitorano le attività degli utenti. Ecco i principali strumenti utilizzati:

- Audit Log: Raccoglie informazioni dettagliate sulle attività, come visualizzazione, modifica, cancellazione o condivisione di file e documenti. Gli amministratori possono configurare e rivedere questi registri.
- Microsoft 365 Compliance Center: Consente di accedere ai log di audit per SharePoint e altri servizi, fornendo un monitoraggio centralizzato delle attività in tutta l'organizzazione.

- Monitoraggio delle modifiche: Le operazioni come il caricamento o la modifica di file vengono registrate, inclusi dettagli su chi ha effettuato l'azione e quando.
- Avvisi e notifiche: Gli amministratori possono impostare avvisi per attività sospette o specifiche azioni, come la condivisione esterna di documenti.

4.6 Archiviazione

I dati aggregati dei singoli Centri e i risultati finali saranno conservati dal Promotore per un periodo di 5 anni per permettere anche il loro sfruttamento per pianificare possibili studi futuri nell'ambito del nodo OHDSI Italia.

4.7 Minimizzazione dei dati

I dati raccolti da Promotore soddisfano il criterio di minimizzazione dei dati in quanto sono raccolti solo dati in forma aggregata. I dati raccolti sono infatti limitati agli scopi dello Studio. Infatti, per l'obbiettivo primario di verificare l'operatività dei Centri nell'utilizzo dei tool OHDSI si raccolgono i dati relativi al tempo di esecuzione dello script R e del tempo di risposta da parte dei centri coinvolti. Per l'obbiettivo secondario di descrivere i dati inclusi nei database OMOP del nodo italiano OHDSI, si raccolgono dati descrittivi aggregati in grado di profilare tali database prodotti da alcuni tool OHDSI (DQD, Achilles, CdmInspection, CatalogueExport)

4.8 Lotta contro il malware

Per i database **OMOP**, la misura più importante riguarda evitare accesso indiscriminato ad Internet per i server OMOP che accedono solamente ai repository per permettere gli aggiornamenti. Inoltre, periodicamente si controlla la sicurezza dei server attraverso la ricezione di un report automatico che riporta l'avvenuta raccolta di tutte le misurazioni e i log da parte dei server e segnala eventuali anomalie

La lotta contro il malware in **SharePoint** è gestita tramite una serie di misure di sicurezza integrate:

- Scansione antivirus integrata: SharePoint utilizza motori di antivirus per esaminare i file caricati sulla piattaforma, prevenendo la diffusione di malware.
- Protezione avanzata da minacce (ATP): Con Microsoft Defender for Office 365, SharePoint monitora e analizza i file alla ricerca di comportamenti sospetti e potenziali minacce, come ransomware o malware avanzati.
- Isolamento dei file: Se un file sospetto viene identificato, può essere automaticamente isolato per prevenire ulteriori danni o distribuzione all'interno dell'ambiente.
- Strumenti di rilevamento basati su cloud: Utilizzando tecnologie di machine learning e intelligenza artificiale, SharePoint può rilevare minacce emergenti in tempo reale.

4.9 Backup

I backup di SharePoint sono gestiti principalmente da Microsoft attraverso un'infrastruttura cloud robusta e automatizzata, garantendo il ripristino dei dati in caso di necessità. Ecco come viene gestito:

1. Backup automatici: Microsoft esegue backup regolari dei dati di SharePoint su Microsoft 365, conservando copie per un periodo di tempo prestabilito (generalmente 14 giorni). Questi backup vengono effettuati in modo trasparente all'utente.
2. Ripristino della versione precedente: SharePoint offre funzionalità di versioning, che permettono agli utenti di ripristinare versioni precedenti di documenti, utile in caso di errore o cancellazioni accidentali.

3. Cestino: I file cancellati dagli utenti sono recuperabili tramite il cestino di SharePoint per un periodo di tempo, generalmente 93 giorni.
4. Protezione contro il ransomware: SharePoint integra strumenti di difesa avanzati per rilevare e ripristinare file in caso di attacchi ransomware.

UNIPV ha inoltre attivato il servizio di Backup di Veem Data Cloud

4.10 Gestione postazioni

Le postazioni dei consulenti del Promotore, vale a dire il personale di Biomeris, che si occuperanno delle analisi dei risultati condivisi dai Centri, sottostanno allo standard ISO 27001 che per esempio determina che ogni postazione sia controllata da XDR. Tale sistema assicura per esempio l'aggiornamento del sistema operativo e dell'antivirus.

Similmente le postazioni dei referenti operativi dei Centri sono controllate per garantirne la sicurezza. Inoltre l'acceso al database OMOP è abilitata solamente all'interno della rete privata del Centro tramite accesso web criptato.

4.11 Contratto con il responsabile del trattamento

Il Protocollo dello Studio descrive nel dettaglio il trattamento eseguito dal Promotore.

In alcuni casi il Promotore stipula un accordo con il Centro partecipante per definire le modalità di conduzione dello studio osservazionale

Per utilizzare SharePoint (e più in generale i servizi cloud di Microsoft, come Microsoft 365) UNIPV ha accettato una Data Processing Agreement (DPA) o Contratto di trattamento dei dati. Questa DPA è un accordo legale che specifica come Microsoft, in qualità di responsabile del trattamento (data processor), gestisce e protegge i dati personali per conto del cliente, che è considerato il titolare del trattamento (data controller), in conformità al GDPR e ad altre normative sulla protezione dei dati. L'ultima versione di tale DPA è accessibile a questo link:

https://www.microsoft.com/licensing/docs/documents/download/MicrosoftProductandServicesDPA(WW)(Italian)(Jan022024)(CR).docx

Il Dipartimento di Ingegneria Industriale e dell'Informazione dell'Università di Pavia e BIOMERIS sottoscrivono una DPA per definire il ruolo di BIOMERIS in qualità di responsabile esterno del trattamento.

4.12 Sicurezza dei canali informativi

La sicurezza dei canali informativi in SharePoint è garantita dal fatto che tutti i dati in transito tra utenti e server sono protetti tramite TLS (Transport Layer Security), mentre i dati a riposo sono crittografati con AES-256.

4.13 Gestione del personale

Tutti i referenti operativi dei Centri hanno frequentato corsi di formazione all'utilizzo dei tool OHDSI e del CDM OMOP da parte di aziende certificate dal progetto IMI-EHDEN come esperte in materia di infrastruttura OMOP/OHDSI.

Il personale dell'azienda consulente del Promotore (Biomeris) è costantemente formato in materia di scurezza delle informazioni tramite corsi su ISO 27001 e Amministratori di Sistema.

5 RISCHI

Non considereremo i rischi legati all'esistenza dei database OMOP presso i singoli Centri, eventualmente affrontati dai singoli centri secondo modalità conformi alla normativa vigente in materia di trattamento dei dati personali per finalità di ricerca scientifica.

Si considerano in questo capitolo solo i rischi direttamente connessi all'esecuzione dello Studio.

5.1 ACCESSO ILLEGITTIMO AI DATI

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

Esposizione dei dati aggregati del singolo Centro

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

- Errore referente Centro: Errore da parte del Referente operativo di un centro partecipante durante l'attività di condivisione dati aggregati del singolo Centro.
- Errore personale Promotore: Errore da parte del personale coinvolto nella fase di analisi dei risultati o di condivisione dei dati da parte del Promotore che determina un'esposizione dei risultati.
- Attacco esterno a PC Centro: Attaccante esterno che riesce a infiltrarsi sui terminali del personale coinvolto nello studio.
- Attacco esterno a PC Promotore: Attaccante esterno che riesce a infiltrarsi nella cartella di condivisione OneDrive

5.1.3 Quali sono le fonti di rischio?

- Referente Centro: referenti operativi dei centri
- Personale/consulenti del Promotore
- Attaccante esterno

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

- Crittografia
- Anonimizzazione
- Controllo accessi logici
- Partizionamento
- Tracciabilità'
- Minimizzazione dati
- Lotta contro il malware
- Gestione delle postazioni

- Contratto con il responsabile del trattamento
- Sicurezza dei canali informativi
- Gestione del personale

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

Trascurabile

Motivazione: un eventuale accesso illegittimo sarebbe comunque un accesso a dati aggregati anonimizzati

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

Limitata

Motivazione: la minaccia più probabile può essere considerata un attacco esterno al terminale del personale coinvolto nello Studio ha comunque probabilità limitata in quanto dovrebbe superare le misure di sicurezza delle postazioni del personale coinvolto.

5.2 MODIFICHE INDESIDERATE DEI DATI

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

Cambiare risultati: Una modifica dei dati aggregati prodotti da un Centro o dei risultati finali prodotti dal Promotore potrebbe influire sulle considerazioni statistico/descrittive riguardati i dati inclusi nei database OMOP in Italia ed eventualmente creare problemi nel riutilizzo di tali dati come base per analisi di fattibilità per futuri studi.

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

Modifica referente Centro: Nel caso in cui il referente operativo di un Centro modifichi volutamente lo script R fornito dal Promotore per generare i risultati condivisi.

Errore personale Promotore: Nel caso di un errore del personale del Promotore nell'eseguire le analisi finali.

5.2.3 Quali sono le fonti di rischio?

- Personale coinvolto nello Studio: referenti operativi dei centri e personale del Promotore

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

- Tracciabilità'
- Archiviazione
- Backup

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

Trascurabile

Motivazione: L'impatto identificato risulta comunque avere conseguenze trascurabili sugli interessati in quanto l'eventuale riutilizzo dei risultati di questo Studio per l'arruolamento di pazienti per futuri studi sarà

comune incrociato di nuovo con i dati reali dei singoli database OMOP non impattati dalle minacce identificate.

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

Trascurabile

Motivazione: La modifica del referente del Centro deve essere voluta quindi molto improbabile, inoltre risulta improbabile anche l'errore del personale del Promotore dovrebbe passare le revisioni previste.

5.3 PERDITA DI DATI

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

Non esiste rischio per gli interessati dalla perdita di dati raccolti dallo Studio. In caso di perdita sia dei dati generati dal singolo Centro che dal Promotore, è possibile ricrearli eseguendo di nuovo il flusso operativo dello Studio quindi tale rischio avrebbe un impatto trascurabile

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

NA

5.3.3 Quali sono le fonti di rischio?

NA

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

NA

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

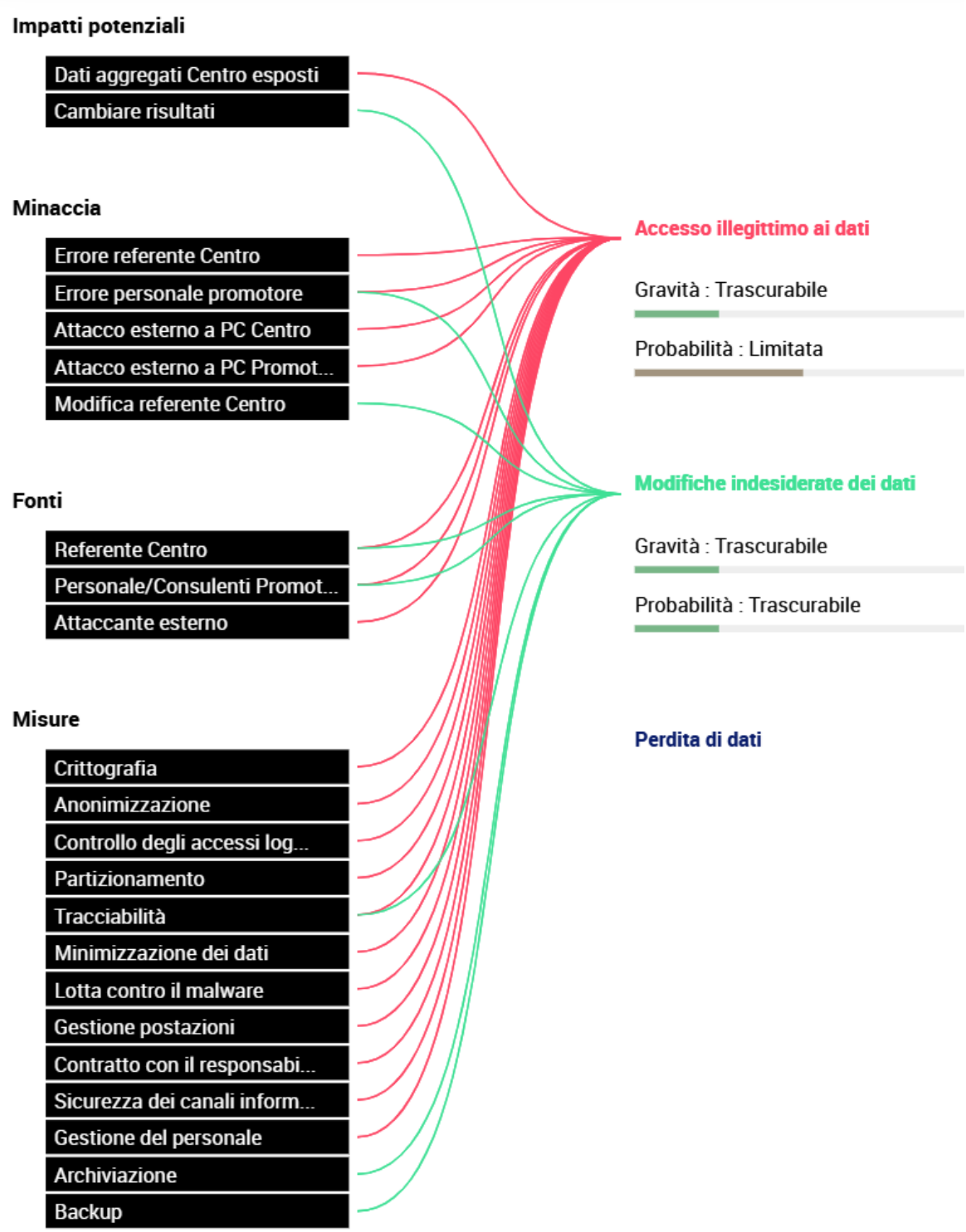
NA

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

NA

5.4 PANORAMICA DEI RISCHI

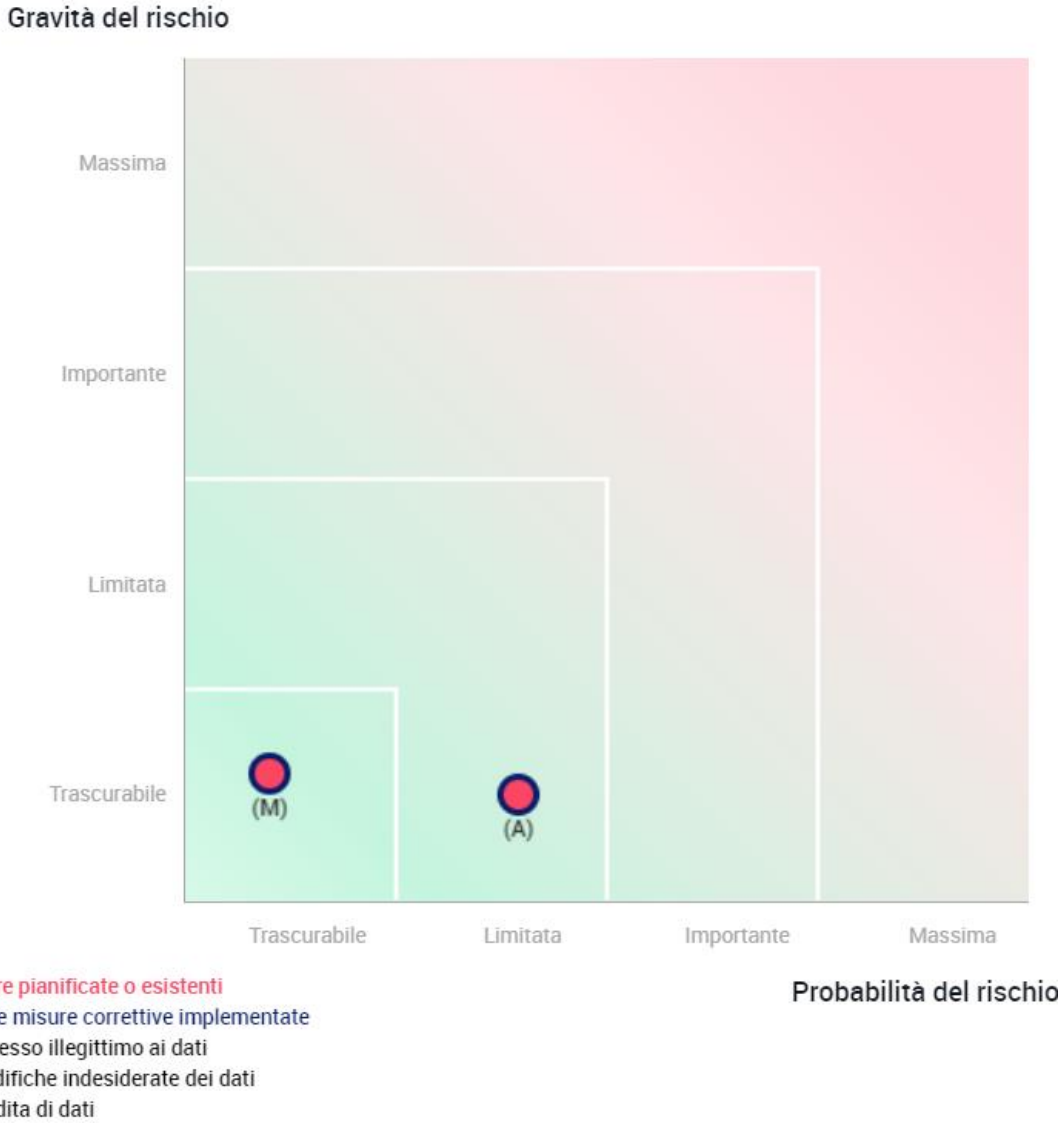
Questa visualizzazione permette una panoramica globale e sintetica degli effetti prodotti dalle misure sulle componenti di rischio che esse contribuiscono a mitigare.



6 CONVALIDA

6.1 MAPPATURA DEL RISCHIO

Questa visualizzazione permette di confrontare il posizionamento del rischio prima e dopo l'applicazione delle misure aggiuntive.



- Misure pianificate o esistenti
- Con le misure correttive implementate
- (A)ccesso illegittimo ai dati
- (M)odifiche indesiderate dei dati
- (P)erdita di dati

6.2 PIANO D'AZIONE

Piano d'azione ente specifico con azioni da eseguire tra le seguenti

- Pubblicazione estratto di DPIA e Informativa dello Studio