

# REGISTRO LOMBARDO MALATTIE RARE

## ALLEGATO 3

RAPPORTO  
AL 31 DICEMBRE 2019

A cura del  
**Centro di Coordinamento**

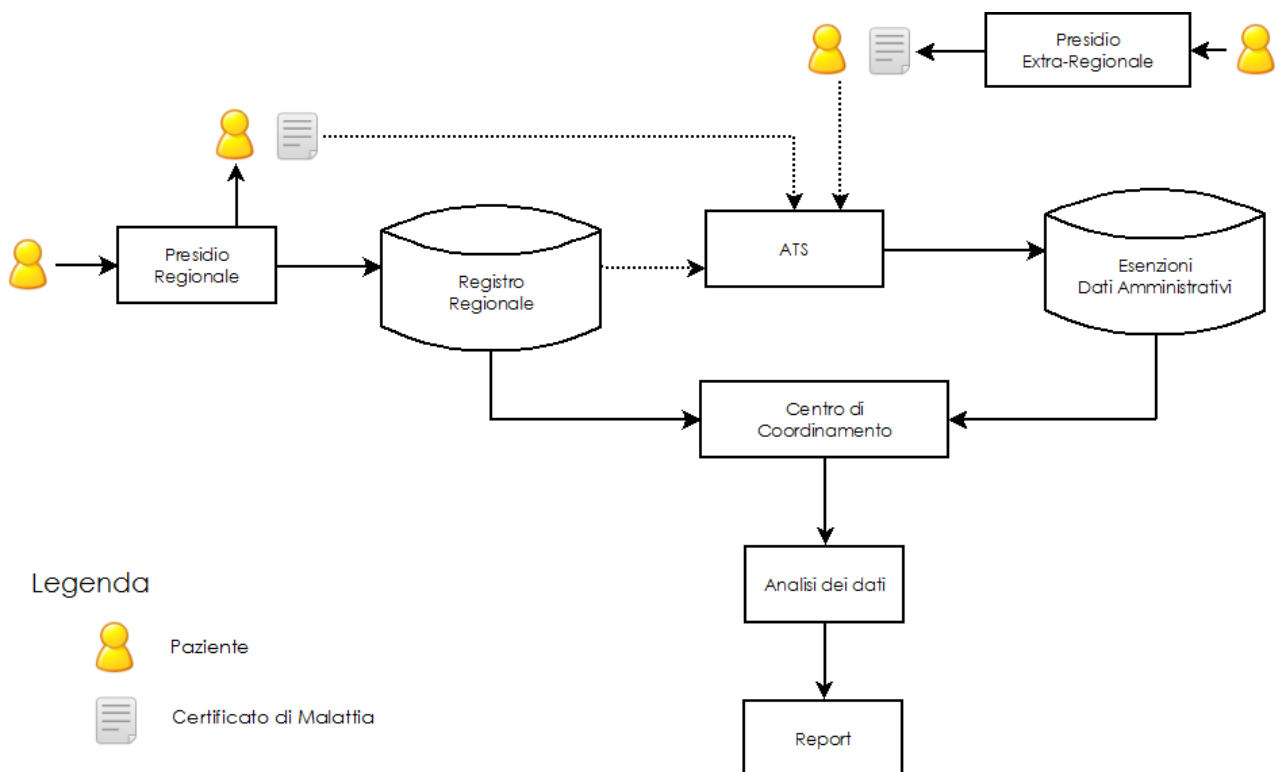
## ALLEGATO 3 - PROTOCOLLO DI VALIDAZIONE DEI DATI E METODO DI RECORD LINKAGE

Per essere utili i dati in un registro devono essere di buona qualità (1). Per assicurarla è stato redatto un protocollo di validazione, con l'obiettivo di evitare la presenza di errori durante le successive fasi di analisi. Il protocollo ha l'obiettivo di identificare i dati incompleti e inaccurati, migliorando la qualità delle analisi.

Al fine di potenziare il Registro Lombardo delle Malattie Rare quale strumento per l'individuazione dei casi di malattia rara, si è ritenuto necessario integrarne i dati con un altro flusso informativo sanitario regionale e, in particolare, con i dati Amministrativi di Esenzione, per ottenere una base statistica informativa unica, sulla quale effettuare le analisi per lo studio della distribuzione delle malattie rare in Lombardia.

La Figura 1 riassume il flusso dei dati regionali utilizzati per la stesura del report.

**Figura 1.** Flusso regionale dei dati.



Per la combinazione di dati registrati in due o più flussi informativi, viene utilizzato il metodo del Record Linkage (RL), uno strumento che permette l'integrazione delle informazioni provenienti da diverse sorgenti di dati (2). Tale strumento permette di ampliare la capacità informativa dei dati, considerando tuttavia che la qualità dell'esito è direttamente correlata ai dati immessi nelle diverse fonti. È per tanto necessaria una valutazione/validazione dei dati provenienti dalle diverse fonti oggetto di RL, al fine di evidenziare, prima della procedura stessa, incongruenze e segnalazioni inadeguate e allestire basi di dati il più possibile accurate.

Per la tipologia delle fonti di dati a nostra disposizione, si è optato per il RL deterministico, ovvero, per l'utilizzo di una serie di regole basate sull'accordo esatto dell'insieme di caratteristiche che costituiscono la chiave identificativa di un individuo. La tecnica prevede che due record provenienti da diverse sorgenti di dati si riferiscano allo stesso individuo solo se l'intera chiave identificativa coincide perfettamente (3).

Prima dell'esecuzione del RL tra le due fonti di dati, è stato creato un protocollo di validazione per il Registro e uno per i dati Amministrativi di Esenzione.

Il protocollo di validazione dei dati del Registro può essere riassunto in tre punti:

1. Validazione per l'analisi dell'attività dei Presidi: controllo della completezza, della congruenza e della validità dei dati presenti (esempio: schede di diagnosi annullate, malattie non più nell'elenco delle malattie rare, congruenza tra le diverse date presenti, valori mancanti);
2. Validazione per l'analisi dei casi di malattia rara: controllo e validazione dei pazienti con più di una malattia rara (esempio: casi di affinamento diagnostico);
3. Validazione per l'analisi delle modalità di diagnosi: controllo della completezza dei dati della modalità di diagnosi della patologia (esempio: presenza di almeno una modalità di diagnosi nella scheda di diagnosi).

Il protocollo di validazione dei dati Amministrativi di Esenzione si può riassumere in due passi:

1. Validazione dei dati anagrafici (esempio: provincia di residenza non valida);
2. Validazione dei dati di patologia (esempio: record con inizio esenzione precedente al 2002, esenzioni per patologie non più presenti nell'elenco delle malattie rare, esenzioni scadute).

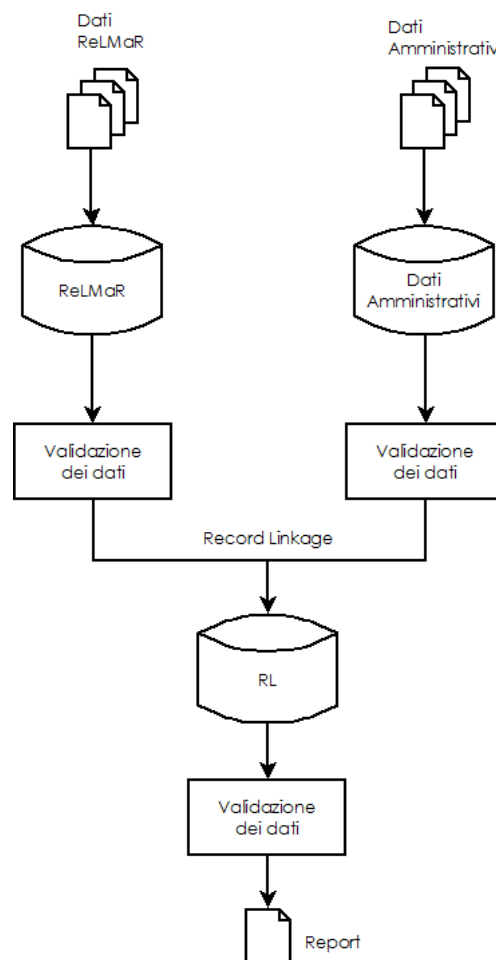
Al termine dei processi di validazione dei due flussi, viene eseguito il record linkage, per integrare le due fonti di dati e ampliarne la capacità informativa. In sintesi: vengono ricavati i casi di malattia rara dalle due fonti di dati ed eseguita su di essi la procedura di record linkage deterministico, al termine della quale si individuano eventuali casi di patologia scaduta e successivamente si esegue l'analisi dei casi di pazienti affetti da più di una patologia rara.

Al termine di tutte le procedure appena descritte si ottiene una base di dati unica, sulla quale poter eseguire le analisi e redigere il report.

In Figura 2 è riportato il procedimento complessivo di validazione dei dati.

---

**Figura 2.** Procedimento complessivo di validazione.



## **BIBLIOGRAFIA**

1. Arts DGT, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study and genetic framework. s.l. : J Am Med Inform Assoc. 2002; 9(6): 600-11.
2. HL., Dunn. Record Linkage. s.l. : Am J Public Health 1946; 36: 1412-16.
3. F., Scheuren. Linking health records: human right concerns. In: Proceeding of an international workshop and exposition: record linkage techniques. s.l. : 20-21 March 1997; Arlington, USA. National Academic Press, Washington DC 1999.