

**PROGETTO**

**“CREAZIONE DI UN DATA WAREHOUSE PER L’ANALISI DEI  
DATI CLINICI REGISTRATI NELLA PRATICA CLINICA”**

*RELAZIONE DEL TERZO QUADRIMENSTRE DI ATTIVITA’*

*(GENNAIO 2014 – APRILE 2014)*

Elaborazione a cura di  
Francesca Maria Morena

## Indice

OBIETTIVI .....	2
ATTIVITA' SVOLTE .....	2
1. DISTRIBUZIONE DEI PACCHETTI ETL .....	2
2. ISTRUZIONE AL PERSONALE SANITARIO .....	2
3. MESSA A PUNTO DELLA REPORTISTICA .....	3
4. ANALISYS SERVICES.....	3
SINTESI DEI REPORT ELABORATI RISPETTO ALLE DIMENSIONI ANALISI INDIVIDUATE.....	5
PROPOSTE DI SVILUPPO.....	8

## OBIETTIVI

I principali obiettivi del II quadrimestre di attività possono essere sintetizzati come segue:

- distribuzione dei pacchetti ETL implementati, al fine di ottenere l'aggiornamento periodico e automatico del contenuto delle tabelle create;
- istruzione del personale sanitario coinvolto sulle modalità di modifica dei report creati e di creazione di nuovi report;
- creazione di nuovi report e messa a punto della reportistica per analisi statistica dei dati;
- analisi dello strumento Microsoft analysis services per la creazione di cubi di dati.

## ATTIVITA' SVOLTE

### *1. DISTRIBUZIONE DEI PACCHETTI ETL*

Al fine di rendere automatica l'estrazione dei dati da Babele ed il popolamento e l'aggiornamento delle tabelle create sul data warehouse clinico, si è provveduto a creare appositi processi per l'allineamento dei fatti e delle dimensioni di DW\_CLI. All'interno dei processi "Allinea fatti in DW\_CLI" ed "Allinea dimensioni in DW\_CLI" sono stati creati opportuni passaggi da eseguire in sequenza per l'aggiornamento corretto dei dati.

L'aggiornamento automatizzato è stato programmato per tutti i giorni della settimana dalle ore 11.30 (per le dimensioni) e dalle 12 (per i fatti).

### *2. ISTRUZIONE AL PERSONALE SANITARIO*

Negli ultimi mesi di attività, al fine di rendere il personale sanitario coinvolto nel progetto il più possibile autonomo, è iniziata l'attività di illustrazione dei principali strumenti utilizzati e degli obiettivi raggiunti nel corso dei mesi precedenti. L'attività è stata rivolta al referente individuato dalla S.C. Nefrologia e Dialisi, che è stato il principale interfaccia sanitario anche per tutta l'attività di progettazione.

Per poter iniziare l'attività di affiancamento, è stata necessaria l'installazione da parte del servizio informatico aziendale dei programmi necessari sul computer dell'utente e la successiva verifica del corretto funzionamento degli stessi e della possibilità di accesso alle cartelle di rete da parte dell'utente.

Per favorire l'attività e la successiva autonomia dell'utente, è stato realizzato e consegnato un breve manuale che riepilogasse:

- le cartelle di rete nelle quali è stato salvato il materiale relativo al progetto e l'illustrazione del contenuto di ognuna di esse
- gli strumenti ed i software utilizzati con una breve spiegazione delle principali operazioni da eseguire per aprire correttamente i progetti realizzati

- per il software utilizzato per la realizzazione della reportistica, la spiegazione dei passi da seguire per modificare un report esistente e per creare un nuovo report.

### 3. *MESSA A PUNTO DELLA REPORTISTICA*

Durante l'attività di affiancamento del personale sanitario, sono emerse alcune necessità di modifica dei report fino a quel momento creati.

In particolare, è stata richiesta la possibilità di visualizzare il valore di  $kt/V$ , indicatore dell'efficacia dialitica, per paziente, in modo da poter valutare il suo andamento nel corso delle varie sedute di dialisi.

Per rispondere a tale necessità è stato creato il report: kt V per paziente:

Filtrando per cognome e nome del paziente visualizza i valori di  $Kt/V$  nel tempo, al fine di verificarne l'andamento.

L'utente può selezionare il cognome del paziente da ricercare dall'elenco di tutti i cognomi dei pazienti di dialisi. Il nome è da scegliere tra quelli associati al cognome precedentemente selezionato.

Come per il report  $Kt/V$ , per ogni paziente ed ogni seduta di dialisi vengono estratti, quando disponibili, i dati relativi ai valori di urea prima e dopo il trattamento dialitico e, quando possibile effettuare l'operazione ( $\log$  di numeri  $> 0$ ), viene calcolato il valore di  $kt/V$  secondo la formula di Daugirdas.

Il report contiene la tabella che riporta i valori di urea prima e dopo la dialisi,  $Kt/V$ , la data dell'esame ed il turno, l'anno, il mese, il giorno della settimana ed il turno.

### 4. *ANALISYS SERVICES*

Al fine di ricercare lo strumento più adatto e semplice per l'elaborazione di reportistica di analisi dei dati clinici da parte del personale sanitario, mi è stato richiesto di analizzare uno degli strumenti Microsoft, Analysis Services, attualmente non utilizzato in nessuna struttura dell'A.O. Ordine Mauriziano di Torino.

E' stata innanzitutto effettuata una ricognizione sullo strumento, che ha permesso di evidenziare quanto segue.

Analysis Services fornisce una gamma di soluzioni per la compilazione e la distribuzione di database analitici utilizzati per il supporto decisionale in Excel, PerformancePoint, Reporting Services e altre applicazioni di Business Intelligence.

Alla base di qualsiasi soluzione di Analysis Services c'è un modello di dati Business Intelligence e un'istanza del server che creano un'istanza, processi, query e gestiscono oggetti in quel modello.

I modelli sono compilati in dati storici che sono già raccolti in database transazionali e altri archivi dati, quindi vengono annotati con metadati che consentono di misurare, modificare e confrontare i dati

aziendali in query ad hoc o report personalizzati. Dopo la progettazione del modello, questo viene distribuito a un server Analysis Services come database, dove diventa disponibile a utenti autorizzati che si connettono tramite Excel e gli altri strumenti.

E' possibile generare database analitici a prestazioni elevate che si integrano facilmente con Excel, con i report di Reporting Services, con altre applicazioni di Business Intelligence.

La creazione di una cartella di lavoro di Excel che contiene dati PowerPivot richiede Excel 2010 e il componente aggiuntivo PowerPivot per Excel. Dopo aver creato la cartella di lavoro, è possibile pubblicarla in una farm di SharePoint Server 2010 in cui siano disponibili Excel Services e una distribuzione di SQL Server PowerPivot per SharePoint. È possibile aprire le cartelle di lavoro di PowerPivot in Excel 2007. Tuttavia, Excel 2007 non consente di creare o modificare dati PowerPivot o di interagire con tabelle o grafici pivot in cui vengono utilizzati dati PowerPivot. È necessario utilizzare Excel 2010 per ottenere accesso completo a tutte le funzionalità PowerPivot.

Lo strumento Microsoft è stato confrontato con un altro strumento di Business Intelligence, QlikView. I risultati del confronto sono riportati nella tabella sottostante.

MICROSOFT	QLIKVIEW
Più complesso da utilizzare Richiede una buona formazione agli utenti	Più semplice ed intuitivo Necessita di una formazione di base
Minor richiesta di memoria	Crea una copia dei dati all'interno del progetto File di grandi dimensioni
Tempi lunghi di implementazione	Tempi brevi di implementazione
Costo basso	Costo più elevato (circa 4000 € a licenza)
Uso di excel	Visualizzazione dei dati interattiva e drill down

Lo strumento Microsoft necessita di una progettazione complessa e di utenti esperti per il corretto utilizzo. Utilizza inoltre un linguaggio proprio dedicato, MDX (MultiDimensional eXpressions). D'altra parte è caratterizzato da costi inferiori e, successivamente alla fase di progettazione iniziale, consente l'accesso ai dati pre-elaborati su Excel, strumento conosciuto dalla maggior parte degli utenti, almeno per le sue funzionalità di base.

Dopo la fase di analisi dello strumento, è stato intrapreso un tentativo di utilizzo del programma Analysis Service, consultando le guide disponibili on-line. Come suggerito dai manuali, si è provveduto alla creazione di viste di dati, correlando tra loro le diverse tabelle di DW\_CLI. Non conoscendo lo strumento, si è partiti dalla creazione di una vista di dati semplice, contenente un numero limitato di tabelle, seppur già in grado di consentire le prime analisi dei dati. Sono stati poi creati le dimensioni ed i fatti ed il cubo di dati. Arrivati a questo punto sono stati riscontrati errori legati all'impossibilità di connessione con il server di Analysis Services. Nonostante vari tentativi e la richiesta di supporto ai sistemi informativi aziendali, non è stato possibile risalire alla causa dell'errore e risolvere quindi il problema. L'attività di progettazione attraverso il nuovo strumento è stata quindi interrotta.

## SINTESI DEI REPORT ELABORATI RISPETTO ALLE DIMENSIONI ANALISI INDIVIDUATE

Di seguito vengono riepilogati i report elaborati per rispondere alle dimensioni di analisi individuate nella prima fase del progetto, in relazione ai differenti analytic themes.

Elenco dei report realizzati:

1. Pazienti prevalenti
2. Andamento storico di numero ed età dei pazienti
3. Età dialitica nei pazienti prevalenti in dialisi
4. Accessi vascolari nei pazienti prevalenti in EC
5. Cause di uscita dalla dialisi
6. Mortalità dei pazienti in dialisi
7. Andamento storico mortalità pazienti in dialisi
8. Pazienti incidenti
9. Accessi vascolari nei pazienti incidenti
10. Andamento storico dei pazienti incidenti
11. Fasce di età all'ingresso in dialisi
12. Adeguatezza dialitica: Kt/V
13. Kt/V per paziente
14. Spesa annuale dei farmaci utilizzati in dialisi
15. Andamento spesa di farmaci utilizzati in dialisi
16. Spesa annuale per tipologia di farmaci

### *Analytic theme: analisi dei pazienti prevalenti in trattamento dialitico cronico*

<i>Requested analyses</i>	<i>n° report</i>	<i>Comments</i>
numero dei pazienti prevalenti nel 2013	1	
n° pazienti in dialisi suddivisi per modalità di trattamento e per fasce di età e sesso	1	
utilizzo della dialisi peritoneale: pazienti in dialisi peritoneale rispetto al totale dei pazienti in trattamento dialitico	1	
andamento negli ultimi 10 anni della numerosità di pazienti prevalenti	2	dati disponibili dal 2005
n° pazienti prevalenti suddiviso per modalità di trattamento (trapianto renale, dialisi EC, dialisi peritoneale)	2	non disponibili dati relativi ai trapianti
n° pazienti in dialisi per professione	non realizzabile	campo professione popolato solo per il 46% dei pazienti, non sufficiente ad effettuare analisi

n° pazienti portatori di trapianto	non realizzabile	non disponibili dati relativi ai trapianti
percentuale di pazienti prevalenti in dialisi suddivisi per fasce di età dialitica	3	
% di tipologie di accessi vascolari in dialisi extracorporea (fistola, cateteri estemporanei, cateteri a permanenza, protesi) nei pazienti prevalenti	4	
cause di uscita dalla dialisi (decesso, suddividendo per dialisi extracorporea e dialisi peritoneale; ripresa della funzione renale; trasferimento ad altro centro)	5	
mortalità nella popolazione prevalente in dialisi (numero e tasso grezzo %) per modalità di trattamento (dialisi extracorporea, dialisi peritoneale, trapianto)	6	
decessi in dialisi (%) dei pazienti nelle due modalità di trattamento in rapporto all'età del paziente, all'età dialitica (anni)	6	
cause di morte del paziente in dialisi (% per cause cardiovascolari, % per cause infettive, % per neoplasie, % per denutrizione con cachessia, % per altro) suddivise per modalità di trattamento (dialisi extracorporea e peritoneale)	non realizzabile	cause di morte non registrate
Andamento negli anni della mortalità dei pazienti in dialisi	7	

**Analytic theme: analisi dei pazienti incidenti**

<b>Requested analyses</b>	<b>n° report</b>	<b>Comments</b>
numero dei pazienti incidenti nel 2013	8	tra gli incidenti ci sono i pazienti in predialisi: nessuno schema associato
età media dei pazienti incidenti in dialisi cronica per modalità di trattamento	8	
% pazienti incidenti per fasce di età (< 15, 16-25, 26-45, 46-65, 66-75, >75) e per modalità di trattamento	8	
n° di pazienti incidenti per sesso e fasce di età	8	
% utilizzo della dialisi peritoneale nei pazienti incidenti	8	
% di utilizzo della dialisi peritoneale nei pazienti incidenti per età	8	
% di tipologie di accessi vascolari in dialisi extracorporea (fistola, cateteri estemporanei, cateteri a permanenza, protesi) nei pazienti incidenti e confronto con i pazienti prevalenti	9	per il 70% la tipologia di accesso non è dichiarata (nessuno schema associato)
n° di pazienti incidenti per professione	non realizzabile	oltre il 50% dei pazienti incidenti non ha professione specificata
andamento negli ultimi 10 anni della numerosità di pazienti incidenti	10	

variazione dell'età nei pazienti incidenti negli ultimi 10 anni	10	
% utilizzo della dialisi peritoneale nei pazienti incidenti: andamento negli ultimi 10 anni	10	
altri ingressi in dialisi nell'anno che si aggiungono ai veri incidenti (rientro da trapianto, trasferimento da altri centri)	non realizzabile	dati non disponibili
% dei rientri da trapianto (trapiantati che devono sottoporsi alla dialisi a causa della perdita di funzione del trapianto) sulla popolazione incidente ed intervallo di tempo tra il trapianto e il rientro in terapia sostitutiva	non realizzabile	dati non disponibili

**Analytic theme: analisi di efficacia terapeutica**

<b>Requested analyses</b>	<b>n° report</b>	<b>Comments</b>
Kt/V	12, 13	

**Analytic theme: analisi dei costi**

<b>Requested analyses</b>	<b>n° report</b>	<b>Comments</b>
spesa annua di farmaci	14, 15 ,16	

## PROPOSTE DI SVILUPPO

Il lavoro svolto in questo anno di attività ha permesso di costruire un datawarehouse dei dati clinici relativi alla dialisi e l'elaborazione della reportistica di base, utile all'aggregazione dei dati registrati ed alla loro analisi.

Il lavoro svolto ha inoltre permesso di rilevare alcune criticità legate a:

- qualità, completezza e coerenza dei dati registrati
- modalità di registrazione dei pazienti in predialisi
- mancanza della possibilità di identificare in modo immediato e certo i pazienti trapiantati.

Per risolvere le criticità legate alla qualità del dato, si propone quanto segue:

- ✓ controllo puntuale dei dati anomali e non coerenti e loro correzione. Ad esempio 10 righe con data chiusura e motivo chiusura compilati hanno stato 'A' e risultano quindi aperti;
- ✓ quando possibile, compilazione dei campi vuoti per i quali la compilazione risulta invece necessaria. Ad esempio 15 righe hanno stato 'C' ma hanno motivo chiusura NULL; 3 di queste righe hanno inoltre la data di chiusura vuota;
- ✓ definizione di regole da diffondere al personale sanitario utilizzatore della cartella clinica per il corretto inserimento dei dati: identificazione dei campi obbligatori, condivisione del formato da utilizzare per i campi a testo libero che contengono dati oggetto di analisi.  
Ad esempio attualmente sono registrati oltre 140 valori di peso del paziente anomali. Il peso del paziente prima e dopo la seduta è un dato utilizzato per il calcolo di Kt/V, indice di adeguatezza dialitica. Il campo, essendo un varchar, è un testo libero: in alcuni casi sono presenti caratteri speciali non convertibili in automatico (\*, \, .., +) oppure non è indicato il peso effettivo del paziente, ma la sua variazione rispetto alla misurazione precedente (es. -1kg);
- ✓ monitoraggio periodico dello stato dei dati (coerenza e completezza).

Per risolvere le altre criticità, sarebbe necessaria la modifica, l'integrazione e la messa a punto della cartella clinica attualmente utilizzata per i pazienti di dialisi (soluzione preferibile). Attualmente infatti i pazienti in predialisi vengono registrati come pazienti in dialisi ma ad essi non vengono associati trattamenti dialitici. Occorrerebbe perciò prevedere una sezione della cartella clinica apposita per i pazienti in predialisi.

Per quanto riguarda i pazienti portatori di trapianto sarebbe sufficiente prevedere un flag che consenta di identificare i pazienti trapiantati.

L'attività di progettazione delle messe a punto della cartella clinica di dialisi deve essere effettuata in collaborazione con il personale sanitario, al fine di comprenderne le esigenze e le necessità.

In caso di impossibilità di procedere alla modifica della cartella clinica, si propone una soluzione alternativa, sebbene non ottimale, perché più complessa da gestire e non automatica. Sarebbe possibile infatti aggiungere appositi flag all'interno della tabella TF\_PAZIENTE\_DIALISI che segnalino se il paziente è in predialisi o se è portatore di trapianto. Occorrerebbe perciò una ricognizione su tutti i pazienti ad oggi caricati (attualmente 1138 righe) ed un aggiornamento della tabella. Periodicamente sarebbe poi necessario aggiungere manualmente, agendo direttamente sulla tabella di DW\_CLI, il valore dei flag identificati per i nuovi pazienti, in base alle indicazioni fornite dal personale sanitario.

Si rileva inoltre che quanto ad oggi progettato necessita un'attività di sviluppo, implementazione ed ottimizzazione ulteriore al fine di essere ancora più vicina alle esigenze di analisi dei dati clinici del personale sanitario. L'attività di formazione relativa allo strumento sviluppato ed alla sua successiva modifica e sviluppo deve essere maggiormente capillare.

Si evidenzia infine l'opportunità di diffondere il modello applicato per lo sviluppo del data warehouse di dati clinici per la dialisi ad altre strutture sanitarie, che necessitano di strumenti di analisi delle attività, delle performance e dei costi.

Francesco M. Manno  
