

ANALISI DEI DATI STORICI

File **DTB_Mice.xls**: un solo foglio, contenente le seguenti colonne.

Mouse ID: identificativo del topo nel sistema

ChipBarcode: barcode del microchip inserito nel topo (quelli del file sembrano i barcode fittizi per lo storico)

Strain: razza del topo (**perche' nel file sono tutti della stessa razza?**)

Sex: sesso del topo (**perche' nel file sono tutti maschi?**)

Producer: fornitore del topo? (nel file hanno tutti il valore 'ALLEV')

Date_Available: data dalla quale e' disponibile il topo (*nel db ogni topo ha il valore '/'*)

Age_Available: eta' dalla quale il topo e' disponibile? (*nel db ogni topo ha il valore '/'*)

Treatment: trattamento associato al topo (*nel db ogni topo ha il valore '/'*)

ImplantDetails ID: riferimento al file DTB_Implants.xls, contenente le informazioni degli impianti

ExplantDetails ID: riferimento al file DTB_Explants.xls, contenente le informazioni degli espianti

UniqueGenealogyID (from implant): genID associato al topo dopo l'impianto (**da normalizzare aggiungendo lo '0', come nell'apposita colonna nel file DTB_Implants.xls**)

Dead: indica se il topo e' morto (**nel file ogni topo ha come valore 'no'**)

Corrispondenza tra i vari campi del file e della relativa tabella del database:

Tabella Mice di XMS	DTB_Mice.xls	Commento
id	Mouse ID	Assegnato automaticamente dal sistema, sara' diverso da quello contenuto nel file
id_mouse_strain	Strain	
id_status	'dead'	L'informazione contenuta in Dead viene inglobata nel concetto di status
id_cancer_research_group	/	Nel file non e' specificato a quale gruppo di ricerca sia affidato ogni singolo topo (vedi nota al fondo del documento)
id_source	Producer	
gender	Sex	
barcode	ChipBarcode	
birth_date	/	Proposta: inseriamo una data di default?
available_date	Date_Available	Proposta: inseriamo una data di default? (nel file DTB_Mice.xls il campo e' sempre vuoto)
death_date	/	Proposta: inseriamo una data di default?
id_genealogy	UniqueGenealogyID	Questo genID non e' uniformato con il nuovo formato (aggiunta di uno '0')
notes	/	Non esistono note e/o commenti nel file
-	ImplantsDetailsID	Riferimento all'impianto subito dal topo. Da utilizzare con il file DTB_Implants.xls e gestire con le tabelle Implant_details e Series.
-	ExplantsDetailsID	Riferimento all'espianto subito dal topo. Da utilizzare con il file DTB_Explants.xls e gestire con le tabelle Explants_details e Series.
-	Treatment	<i>Riferimento al trattamento associato al topo. Ogni valore di questa riga ha il valore '/'. Nel caso in cui fosse un valore significativo, si dovrebbe agire sulle tabelle Arms e Protocols di XMS.</i>

Commenti generali:

Come tutti i topi espiantati hanno come voce dead il valore 'no'?

Ipotesi: poniamo come status 'dead' tutti i topi che hanno subito un espianto.

File **DTB_Implants.xls**: due fogli, contenenti le seguenti colonne.

Foglio Implants

ImplantsSeriesID: identificativo della serie di impianto

Operator: indica l'operatore che ha effettuato la serie di impianti

Date: data della serie

Disease, Case, Site: indicano le parti iniziali dei genID utilizzati negli impianti della serie a cui fanno riferimento

Number Of Mice: numero di topi impiantati nella serie

Fresh/Frozen: assume solo valore fresh o frozen (**indicazione sulla piastra utilizzata per la serie?**)

Comments: commento riferito alla serie di impianti

Scope: scopo relativa alla intera serie di impianti (**i valori di questa colonna non corrispondono ai vari scopi attualmente presenti nel database: il valore 'treatment' nel file corrisponde a 'drug testing'? E 'expansion' corrisponde a 'cohort expansion'?**)

Foglio DetailsImplant

ImplantDetailsID: identificativo del relativo record di dati

ImplantSeriesID: riferimento alla serie

UniqueGenealogyID: genID creato con questo singolo impianto (non uniformato con la nuova lunghezza del genID)

ExplantDetailsID: riferimento all'espianto eventualmente subito dal topo impiantato con questo impianto

Aliquot/MouseOfOrigin: indica l'aliquota di origine impiantata nel topo (quindi, indica anche il topo di origine)

Comments: commento relativo al singolo impianto

SiteOfImplant: parte del corpo del topo sul quale si impianta l'aliquota

Inserimento 0 in GenealogyID: contiene gli stessi dati di UniqueGenealogyID, uniformati con l'aggiunta di uno '0' nel contatore (**manca la parte finale del genID, ovvero la parte relativa alla conservazione dell'aliquota (VT, RL,...)**)

Trova trattamento in Explants: solo alcuni record hanno un valore in questo campo (**a cosa serve?**)

Corrispondenza tra i vari campi del file e delle relative tabelle del database:

Tabella Series	DTB_Implants.xls	Commenti
id	ImplantsSeriesID	Assegnato automaticamente dal sistema, sarà diverso da quello contenuto nel file
id_operator	Operator	
id_type	/	Per gli impianti, assume sempre il valore 'implant'
id_scope	Scope	
date	Date	
notes	Comments	
-	Number Of Mice	Dato ricavabile con una query, contando i topi impiantati in questa serie nella tabella Implants_details. Dato da usare per effettuare dei controlli sui dati in fase di inserimento.
-	Disease, Case, Site	Queste voci del file possono essere utilizzate per effettuare dei controlli incrociati sui genID utilizzati in questi impianti
-	Fresh/Frozen	Non esiste corrispondenza di questo dato nel database. Forse può essere utilizzato per ottenere la parte mancante del genID? (ad esempio, 'fresh' = VT)

Tabella Implants_details	DTB_Implants.xls	Commenti
id	ImplantsDetailsID	Assegnato automaticamente dal sistema, sarà diverso da quello contenuto nel file
id_mouse	/	Identificativo del topo impiantato. Non è presente in questo file, ma nel file DTB_Mice.xls viene indicato l'impianto subito da ogni singolo topo.
id_series	ImplantsSeriesID	
aliquots_id	Aliquots/MouseOfOrigin	Riferimento all'aliquota utilizzata per l'impianto (tabella Aliquots)
bad_quality_flag	/	Nel file non si fa riferimento a questo campo. Ipotesi: supponiamo che tutti abbiano avuto esito positivo.
site	SiteOfImplant	
-	UniqueGenealogyID	Da usare per un controllo incrociato con i dati presenti nel foglio Implants e con la colonna "inserimento 0 in GenealogyID"
-	ExplantDetailsID	Dato già presente nel file DTB_Mice.xls, permette di collegare i topi agli espianti. Dato da usare come check dei dati.
-	Inserimento 0 in GenealogyID	Corrisponde a UniqueGenealogyID uniformato al nuovo contatore. Da inserire questo nel db.
-	Trova trattamento in Explants	Serve? Se sì, a cosa lo devo associare nel database? (braccio e/o protocollo?)
-	Comments	Nel database non era stata prevista un campo per i commenti di ogni singolo espianto. E' da inserire?

Tabella Aliquots	DTB_Implants.xls	Commenti
id	/	Identificativo assegnato automaticamente dal sistema
id_explant	/	Utilizzato solo quando si registra un espianto
id_genealogy	Aliquots/MouseOfOrigin	Per registrare le aliquote utilizzate, si inserisce nel database l'aliquota di origine impiantata
idType	/	Utilizzato solo quando si registra un espianto

File **DTB_Explants.xls**: due fogli, contenenti le seguenti colonne.

Foglio Explants:

ExplantID: identificativo della serie di espianto

Operator: operatore che ha effettuato la serie

Date: data della serie (**il formato delle date e' da uniformare; alcune hanno l'anno su 2 cifre, altre su 4**)

NumberOfMice: numero di topi espiantati nella serie di espianti

Scope: scopo per il quale si effettua la serie di espianti

Foglio ExplantDetails:

ExplantDetailsID: identificativo del singolo espianto

ExplantID: identificativo della serie di espianti

UniqueGenealogyID: parte iniziale del genID ottenuto (non trovo corrispondenza con i dati degli impianti (ad es. espianto 136) ---> come mai il genID espiantato corrisponde a quello utilizzato come origine per l'impianto e non al genID aggiornato con l'algoritmo apposito?

MouseID: identificativo del topo espiantato. **Come mai in molti casi e' uguale a UniqueGenealogyID? (---> quindi, anche qui si pone la domanda fatta per il campo precedente) E come mai, quando non e' uguale, questo identificativo non esiste nel file Mice.xls?**

ID gabbie: la maggioranza delle voci di questa colonna e' vuota. Indica la gabbia del topo.

Comments-treatments: annotazione sul trattamento subito da quello xenopaziente. **Corrisponde a un braccio o ad un protocollo?**

FF, VT, RL, SF, OF: indicano il numero delle aliquote collezionate nelle varie tipologie di piastre. Con questo si puo' ricostruire la parte finale del genID partendo da UniqueGenealogyID. In questo modo pero' manca ancora una parte del genID, ovvero il "TissueType" collezionato.

Corrispondenza tra i vari campi del file e delle relative tabelle del database:

Tabella Series	DTB_Explants.xls	Commenti
id	ExplantID	Assegnato automaticamente dal sistema, sara' diverso da quello contenuto nel file
id_operator	Operator	
id_type	/	Per gli espianti, assume sempre il valore 'explant'
id_scope	Scope	Scopo della serie di espianti. <i>Durante l'uso dell'applicazione questo non viene salvato nel db in quanto nella stessa serie possono apparire singoli espianti con scopi programmati tra loro diversi.</i>
date	Date	
notes	/	Nel file DTB_Explants.xls non esistono commenti sulla serie di espianti.
-	NumberOfMice	Dato ricavabile con una query, contando i topi impiantati in questa serie nella tabella Explant_details. Dato da usare per effettuare dei controlli sui dati in fase di inserimento.

Tabella	Explant_details	DTB_Explant.xls	Commenti
id		ExplantDetailsID	Assegnato automaticamente dal sistema, sarà diverso da quello contenuto nel file
id_series		ExplantID	
id_mouse		MouseID	Come mai il mouseID corrisponde a al genID utilizzato come origine per l'impianto e non al genID aggiornato con l'algoritmo apposito?
-		Id gabbie	Ipotesi: non dobbiamo considerare questo dato.
-		Comments_treatments	Questo dato può essere utilizzato per registrare un'associazione tra un topo ed un braccio/protocollo. Il problema è che, rispetto ai campi presenti nell'apposita tabella nel db, mancherebbero vari dati ritenuti attualmente obbligatori (ovvero: id_operator, start_date, expected_end_date)

Tabella	Aliquots	DTB_Implants.xls	Commenti
id		/	Identificativo assegnato automaticamente dal sistema
id_explant		ExplantID	
id_genealogy		UniqueGenealogyID + FF/VT/RL/SF/OF	GenID da calcolare partendo da UniqueGenealogyID e dai vari campi FF/VT/RL/SF/OF. Manca una parte: TissueType. Ipotesi: porre tutti a TUM o a '000'.
idType		/	Questo dato manca nel file Explant_details.xls

Commenti generali:

Come mai diversi topi hanno lo stesso identificativo per i dati ExplantDetails?

ALTRI FILE

File DTB_DevolvedMice.xls: un foglio singolo.

Questo file suppongo raccolga i dati dei topi trasferiti ad altri gruppi di ricerca. Oltre alle intestazioni delle colonne, non contiene informazioni.

File DTB_ExpiredMice.xls: un foglio singolo.

Questo file raccoglie i dati dei topi deceduti. Oltre alle intestazioni delle colonne, non contiene informazioni.

File DTB_Measurements.xls: due fogli.

Questo file raccoglie i dati sulle misurazioni sui vari xenopazienti. Il primo file raccoglie i dati sulla serie di misure, mentre il secondo raccoglie i dati specifici. Oltre alle intestazioni delle colonne, non contiene informazioni.

File DTB_Treatments.xls: un foglio singolo.

Analizzando il contenuto di questo file, suppongo che contenga gli esiti dei vari trattamenti. Quindi, questi dati non trovano corrispondenza nel database di XMS e verranno gestiti dal modulo di integrazione dati.

Nota:

Per rendere XMS operativo, oltre ai dati storici, si devono inserire anche i dati relativi agli utenti ed ai **gruppi di ricerca**, oltre ai dati propedeutici al corretto funzionamento dell'applicazione (ad esempio, via_mode, qualitative_values, drugs, stautus, ...)