



Università degli Studi di Udine

Dipartimento di Studi Umanistici

*Dottorato in Scienze bibliografiche, archivistiche e documentarie
e per la conservazione e restauro dei beni librari ed archivistici*

XXV Ciclo

Strutture tassonomiche e linguaggi specialistici in ambito Biomedico

Candidata

Erika Pasceri

Tutor

Prof. Roberto Guarasci

Co-tutor

Prof. Mauro Guerrini

Coordinatore

Prof.ssa Angela Nuovo

Anno Accademico

2012/2013

Indice

Introduzione	3
Capitolo I	7
1.1 Le malattie rare: Il contesto internazionale	7
1.1.2 Le associazioni sulle malattie rare a livello internazionale	11
1.1.3 I principali studi a livello internazionale sulle malattie rare.....	20
1.2 Analisi del contesto nazionale	22
Capitolo 2	24
2.1 L'importanza degli standard nell'informatica medica	24
2.2 L'Unified Medical Language System	25
2.2.1 IL Metathesaurus	26
2.2.2 Semantic Network	27
2.2.3 Specialist Lexicon.....	28
2.3 Analisi terminologica	29
2.3.1 Acquisizione e analisi dei dati provenienti da Orphanet, dall'ORDR e dalla NORD	30
2.3.2 Le terminologie target: OMIM, SNOMED CT, MeSH, ICD10..	32
2.3.3 La metodologia di lavoro.....	46
2.3.4 Risultati della mappatura verso l'UMLS.....	52
Capitolo 3	58
3.1 Avvio della costruzione del thesaurus: il Registro Nazionale delle Malattie Rare	58
3.2 Dal Registro al thesaurus.....	63
3.2.1 Il thesaurus e le funzioni d'uso	66

3.2.2 Lo strumento utilizzato: il software Multites Pro 2005	71
Capitolo 4	73
4.1 Conclusioni.....	73
4.2 Limiti del lavoro	74
4.3 Prosecuzione del lavoro e prospettive future.....	78
Indice delle Figure	80
Bibliografia	81
Sitografia.....	83
Appendice	84

Introduzione

Obiettivo del lavoro è quello di costruire un modello di classificazione per le malattie rare che possa essere di supporto non solo per i professionisti del settore, ma anche e soprattutto per tutti coloro i quali sono affetti da queste patologie affinché possano utilizzarlo come strumento necessario per la ricerca di informazioni cliniche utili ai percorsi di diagnosi e cura.

Le persone affette da malattie rare si trovano troppo spesso ad affrontare problemi come l'impossibilità di ottenere in tempi rapidi una corretta diagnosi, la mancanza di informazioni esaustive sulla patologia e, spesso, la mancanza di casistica specifica. Tutto ciò non sempre e non solo per oggettiva mancanza di informazioni ma, spesso, per difetto o inesistenza di strumenti di accesso all'informazione stessa. Prendendo come esempio, la sindrome WAGR¹, disturbo patologico estremamente raro legato alla delezione del cromosoma 11 (11q) e caratterizzata dal tumore di Wilms, da anomalie genito-urinarie, e da ritardo nello sviluppo, si scopre che diversi sono stati i termini conosciuti per identificarla, tra cui, appunto, l'acronimo WAGR; termini che si riferiscono alla sua origine genetica (sindrome da delezione 11q, Monosomia 11p13) o alla combinazione di eventi che ne conseguono (tumore di Wilms, anomalie genitourinarie, sindrome da ritardo mentale) senza che siano sempre esplicitate le relazioni di sinonimia esistenti tra i termini.

D'altro canto il problema della non univocità semantica è una questione di non poca rilevanza nel momento in cui si deve compilare, ad esempio, una scheda di dimissione ospedaliera² e ci si trova davanti il

¹ < <http://www.wagr.org/> >

² D'ora in poi SDO.

problema della non esistenza di un termine non ambiguo per identificare una determinata patologia e, e ci si vede, quindi, costretti ad usare termini generici, sintomatologie o perifrasi.

I dati raccolti attraverso le SDO costituiscono un prezioso strumento di conoscenza, di valutazione e di programmazione delle attività di prevenzione, diagnosi e cura ma rischiano di essere fortemente limitati se non spesso falsati dalla poca accuratezza descrittiva derivante da un linguaggio non standardizzato.

La necessità di gestire l'informazione e di creare sistemi di accesso e di organizzazione della conoscenza nello settore specifico trova la sua ragion d'essere nella particolare situazione in cui versa il nostro Paese, caratterizzata dall'assenza o dalla poca accuratezza di schemi di classificazione ai quali sia possibile fare riferimento per la compilazione delle SDO, per lo scambio non ambiguo di informazione, per l'interscambio dei dati in ambiente digitale, ecc.

L'obiettivo principale che ci si propone di raggiungere a seguito del lavoro più propriamente terminologico è, quindi, quello di indirizzare gli specialistici del dominio verso un utilizzo condiviso della terminologia specialistica, che possa garantire l'individuazione chiara ed univoca dei concetti corrispondenti, senza rischi di ambiguità. "Prima di iniziare la costruzione di un thesaurus, è necessario stabilire se il thesaurus in questione è realmente necessario; una volta stabilito ciò, è necessario capire che forma il thesaurus debba avere per venire incontro il più possibile agli utilizzatori finali³."

Partendo dal presupposto che esistono numerose classificazioni dell'ambito biomedico, ed in particolare in quello delle malattie rare, il presente lavoro ha come obiettivo quello di costruire un modello che possa essere in un certo qual modo "unificatore" delle precedenti basi costruite.

³ JWAN AITCHISON, ALAN GILCHRIST, DAVID BAWDEN, *Thesaurus construction and use: a practical manual*, Aslib, 2000, 4th ed., p.5.

La base terminologica di partenza è stato il RNMR essendo il riferimento utilizzato, ad oggi, nel contesto nazionale italiano, soprattutto dai professionisti del settore per la refertazione delle patologie rare. Il RNMR durante la fase di ricognizione e di analisi dello stato dell'arte è risultato essere in alcuni casi poco preciso. È stato effettuato, quindi, il controllo terminologico e semantico avvalendosi di modelli di riferimento quali le risorse terminologiche prese in analisi largamente nel secondo capitolo descrittivo del presente lavoro.

Per meglio comprendere le ragioni di questa scelta metodologica, si riporta qui un esempio estratto dal RNMR ed in particolare dalla categoria *“Malattie delle ghiandole endocrine, della nutrizione, del metabolismo e disturbi immunitari”*, in particolare la *Sindrome di George* che è stata inclusa tra le *“Immunodeficienze primarie”*. In realtà la *Sindrome di George* appartiene alla categoria delle malattie cromosomiche, come si evince dalla sua definizione:

“La diagnosi di sindrome di DiGeorge è clinica. La sindrome di DiGeorge è caratterizzata dall'insieme di diverse malformazioni: ipoplasia del timo e delle ghiandole paratiroidi, cardiopatia congenita di tipo troncoconale e dimorfismi del viso minori ma caratteristici. La sindrome velo-cardio-facciale si distingue per l'associazione di alterazioni cardiache di tipo troncoconale, palatoschisi o insufficienza del velo palatino, dimorfismi del viso e disturbi dell'apprendimento. È noto che queste due sindromi rappresentano espressioni cliniche della stessa condizione, in diverse età. La loro definizione si sovrappone a quella di microdelezione 22q11, che è presente in oltre il 95% dei pazienti. L'acronimo CATCH 22 è stato proposto per descrivere le caratteristiche cliniche della microdelezione 22q11 (Cardiac-Abnormal face-Thymus-Cleft palate-Hypocalcemia). L'evoluzione della sindrome è largamente condizionata dalla natura delle malformazioni congenite. L'ipocalcemia, spesso osservata nel periodo prenatale, di solito scompare, ma in alcuni bambini può persistere un

ipoparatiroidismo che deve essere trattato. L'insufficienza velofaringea spesso esita in voce nasale, anche senza palatoschisi, e può essere accompagnata a difficoltà nel linguaggio. Nella popolazione generale l'incidenza stimata della microdelezione 22q11 è circa 1/5000 nati. La microdelezione 22q11 si trasmette con modalità autosomica dominante e nel 10-20% dei casi uno dei due genitori è portatore della microdelezione. Tuttavia, nella maggior parte dei casi, la mutazione cromosomica è de novo⁴.

Sempre all'interno del RNMR nessun sinonimo è stato associato a tale patologia, mentre ne esistono diversi:

PT *Sindrome di Di George*

UF *CATCH 22*

Microdelezione 22q11

Sequenza di DiGeorge

Sindrome da anomalie facciali e troncoconali

Sindrome di Sedlackova

Sindrome di Shprintzen

Sindrome velo-cardio-facciale

Il thesaurus, organizzato principalmente in base all'ordine alfabetico, necessita di una struttura classificatoria, per raggruppare i termini in base al loro significato, mediante le relazioni semantiche, in modo da facilitare la ricerca e il recupero dell'informazione e, inoltre, per permettere la costruzione, il mantenimento e l'aggiornamento della terminologia.

⁴<[http://www.orphane.net/consor/cgi-bin/Disease_Search.php?lng=IT&data_id=126&Disease_Disease_Search_diseaseGroup=sindrome-di-George&Disease_Disease_Search_diseaseType=Pat&Malattia\(e\)/%20gruppo%20di%20malattie=Monosomia-22q11--Sindrome-di-DiGeorge-&title=Monosomia-22q11--Sindrome-di-DiGeorge-&search=Disease_Search_Simple](http://www.orphane.net/consor/cgi-bin/Disease_Search.php?lng=IT&data_id=126&Disease_Disease_Search_diseaseGroup=sindrome-di-George&Disease_Disease_Search_diseaseType=Pat&Malattia(e)/%20gruppo%20di%20malattie=Monosomia-22q11--Sindrome-di-DiGeorge-&title=Monosomia-22q11--Sindrome-di-DiGeorge-&search=Disease_Search_Simple)>

Capitolo I

1. 1 Le malattie rare: Il contesto internazionale

Una malattia rara, nota anche come “malattia orfana”, è una patologia che colpisce ogni anno una piccola percentuale della popolazione mondiale. La maggior parte delle patologie rare sono di origine genetica e, molto spesso, non curabili. In molti casi tali patologie, si manifestano nei primissimi anni di vita e circa il trenta per cento dei bambini affetti da simili patologie, non arriva a compiere il quinto anno di età. La definizione di malattia rara non è univoca, per cui una malattia può essere considerata rara in alcune aree geografiche e “frequente” (o comune) in altre.

Il riferimento all’intera popolazione di un continente, come avviene in Europa, consente il verificarsi di situazioni particolari date dal fatto che una determinata patologia può essere diffusa localmente per motivi genetici, ambientali, ecc. In questo caso la malattia, pur essendo globalmente rara e quindi definita tale, non lo sarà in particolari zone geografiche. Per molti anni, l’AIDS è stata una malattia estremamente rara, solo successivamente divenuta rara ed oggi il numero degli affetti di questa patologia cresce sempre di più, specialmente in alcune popolazioni. Un altro esempio è quello della lebbra, una malattia rara in Italia, ma frequente in Africa centrale. La talassemia, un’anemia di origine genetica, è rara nel Nord Europa, ma è frequente nelle regioni del Mediterraneo; la “malattia periodica⁵” è rara in

⁵ La malattia periodica o meglio conosciuta come febbre Mediterranea familiare (FMF) si manifesta con brevi episodi febbrili che insorgono a intervalli variabili con dolori intermittenti di tipo addominale, toracico, articolare e/o cutaneo, accompagnati dal possibile sviluppo di amiloidosi renale. Questa malattia colpisce soprattutto i soggetti originari dell’area del Mediterraneo.

Italia, ma è molto diffusa in Armenia. Allo stesso modo, esistono molte malattie normalmente diffuse che possiedono varianti rare.

Una patologia è considerata rara quando ha una prevalenza nella popolazione generale inferiore ad una data soglia, codificata dalla legislazione di ogni singolo paese: l'Unione europea definisce tale soglia allo 0,05% della popolazione, ossia 1 caso su 2000 abitanti, l'Italia si attiene a tale definizione. Molte patologie sono però molto più rare, arrivando alla frequenza di un caso su 100.000 persone (0,001%) o più.

Nel gennaio 2004, la Direzione Generale per la salute pubblica della *Commissione Europea* ha costituito la *Task Force delle Malattie Rare (RDTF)*, con l'obiettivo di consigliare ed assistere la Direzione Generale per la salute pubblica della *Commissione Europea* nel promuovere una ottimale azione di prevenzione nonché le attività di diagnosi e cura. Tra i principali obiettivi della RDTF vi è anche l'impegno per costituire e gestire un forum di discussione e favorire uno scambio di opinioni ed esperienze nella comunità degli esperti. La *Task Force* è oggi guidata da Ségolène Aymé, genetista e direttore di *Orphanet*, portale delle malattie rare e dei farmaci orfani, gestito da un consorzio europeo di cui fanno parte una quarantina di paesi e il cui coordinamento ha sede in Francia. Ciascun team facente parte del consorzio si impegna, a livello nazionale, a raccogliere informazioni sulle consulenze specialistiche, sui laboratori di diagnosi, sulle attività di ricerca in corso e sulle associazioni di pazienti nei rispettivi paesi. Tutti i team del consorzio si impegnano, altresì, a rispettare le specifiche di qualità di Orphanet.

La gestione del database, del sito web e il coordinamento di Orphanet sono cofinanziati dall'*Institut National de la Santé et de la Recherche Médicale*, dalla direzione Generale della Sanità francese e dalla Commissione Europea. Le attività di Orphanet a livello nazionale sono allo stesso modo finanziate dalle istituzioni dei rispettivi paesi e/o da fondi specifici.

Orphanet offre una serie di servizi liberamente accessibili che comprendono: un classificazione delle malattie rare in diverse lingue

revisionata periodicamente da un comitato di esperti, un elenco dei farmaci orfani con tutte le loro fasi di sviluppo (dall’Agenzia Europea per il Farmaco all’autorizzazione europea all’immissione in commercio), un elenco dei servizi specialistici presenti nei paesi che ne fanno parte (con informazioni sulle consulenze, sui laboratori di diagnosi, sui progetti di ricerca in corso, sulle sperimentazioni cliniche, sui registri, sui network, sulle piattaforme tecnologiche e sulle associazioni di pazienti), un servizio di assistenza alla diagnosi che permette la ricerca per segni e sintomi, un’Enciclopedia professionale comprendente testi specialistici e un’Enciclopedia dedicata ai pazienti che contiene informazioni sulle singole patologie redatte con un linguaggio accessibile agli utenti generici ed alle relative associazioni, una raccolta di studi e articoli tematici, i Quaderni di Orphanet⁶ che trattano argomenti trasversali, direttamente scaricabili dal sito.

Il database è posto sotto la responsabilità di un comitato scientifico e di un comitato editoriale, i cui membri sono nominati sulla base della loro esperienza e competenza nel campo delle malattie rare, su proposta delle società scientifiche, delle autorità sanitarie pubbliche dei paesi impegnati nel progetto, o di qualsiasi altra associazione del settore. Tutte le informazioni disponibili al pubblico sono validate almeno da un componente del Comitato Scientifico prima della pubblicazione on-line.

Negli Stati Uniti, la legge sulle Malattie Rare⁷ del 2002 ha stabilito che:

“(1) Rare diseases and disorders are those which affect small patient populations, typically populations smaller than 200,000 individuals in the United States. Such diseases and conditions include Huntington’s disease, amyotrophic lateral sclerosis (Lou

⁶ <http://www.orpha.net/consor/cgi-bin/Education_Home.php?lng=IT>

⁷ *Rare Diseases Act of 2002*, (Public Law 107-280 107th Congress).

Gehrig's disease), Tourette syndrome, Crohn's disease, cystic fibrosis, cystinosis, and Duchenne muscular dystrophy⁸.

definendo, quindi, rara "una malattia o condizione che colpisca meno di 200.000 persone negli Stati Uniti", o circa 1 persona su 1.500. La definizione deriva dall'Orphan Drug Act del 1983, legge federale emanata per incoraggiare la ricerca sulle malattie rare e le possibili cure. Dieci anni dopo, nel 1993, è stato istituito anche l'NIH Office of Rare Diseases Research (ORDR), all'interno del National Institute of Health (NIH), l'agenzia federale responsabile della ricerca biomedica negli Stati Uniti e non solo.

Lo scopo dell'ORDR è quello di coordinare e sostenere la ricerca sulle malattie rare, cercando di fornire quante più possibili informazioni, ma soprattutto di promuovere la collaborazione e l'interoperabilità a livello internazionale.

Esistono diverse associazioni anche di stampo totalmente volontario che nascono dalla necessità di voler fare qualcosa di realmente concreto per chi è affetto da tali patologie, come ad esempio la National Organization for Rare Disorders⁹. I leader delle organizzazioni di pazienti affetti da malattie rare avevano iniziato a rendersi conto che vi sono problemi comuni a tutte le persone affette da malattie rare e hanno, quindi, unito le proprie forze per chiedere l'approvazione di una specifica normativa nazionale che si è concretizzata nell'Orphan Drug Act.

Oggi, la NORD fornisce informazioni, sensibilizzazione, ricerca e servizi ai pazienti per aiutare tutti i pazienti e le famiglie colpite da malattie rare.

⁸ Public Law 107-280, Nov. 6, 2002, Sec. 2. *Findings and Purposes*.

⁹ D'ora in poi NORD.

1.1.2 Le associazioni sulle malattie rare a livello internazionale

Sono numerose le associazioni che si occupano di malattie rare, non solo da un punto di vista medico, ma anche e soprattutto come punto di riferimento per chi volesse ottenere informazioni, mettersi in contatto con altre associazioni, o con specialisti del settore; tali associazioni, infatti, hanno sviluppato dei sistemi classificatori, più o meno dettagliati, o, in alcuni casi, semplici liste terminologiche con annesse descrizioni sulle malattie e relativi riferimenti verso centri specializzati di cure.

Tali associazioni sono sia governative, ma anche private o semplicemente gestite da comunità spontanee formatesi dalle famiglie dei pazienti affetti da tali patologie.

1.1.2.1 European Union Committee of Experts on Rare Diseases (EUCERD)

Il comitato di esperti dell'Unione europea sulle malattie rare è stato istituito formalmente con la decisione della Commissione europea del 30 novembre 2009 (2009/872/EC)¹⁰. La missione principale del comitato è quella di aiutare la Commissione Europea per la preparazione e l'attuazione delle azioni comunitarie nel settore delle malattie rare, in collaborazione e con delle consultazioni con gli organismi specializzati negli Stati membri, le autorità europee competenti in materia di ricerca e di azione per la salute pubblica e altre parti interessate che operano nel settore¹¹.

L'EUCERD promuove l'interoperabilità per favorire continui scambi di esperienze, politiche e pratiche tra le parti che collaborano al progetto.

In particolare, è impegnato nelle seguenti azioni:

¹⁰ Commission Decision, 30th November 2009, establishing a European Union Committee of Experts on Rare Diseases, (2009/872/EC), Art. 1.

¹¹ *Ivi*, art.2.

- Assistere la Commissione nel monitoraggio, valutazione e diffusione dei risultati delle misure adottate a livello comunitario e nazionale nel campo delle malattie rare;
- Contribuire nella realizzazione di azioni comunitarie nel settore, in particolare per l'analisi dei risultati, proponendo miglioramenti delle misure adottate;
- Fornire pareri, raccomandazioni o report alla Commissione, su richiesta di quest'ultima o di propria iniziativa;
- Assistere la Commissione nella cooperazione internazionale in materia di malattie rare;
- Assistere la Commissione nell'elaborazione di linee guida, suggerimenti e qualsiasi altra azione definita nella comunicazione della Commissione e nella raccomandazione del Consiglio;
- Fornire una relazione annuale sulle proprie attività alla Commissione;
- Il comitato ha il potere di istituire gruppi di lavoro temporanei composti da esperti esterni per missioni specifiche;

1.1.2.2 European Organization for Rare Diseases (EURORDIS)

L'Organizzazione Europea per le Malattie Rare¹² è una alleanza di organizzazioni di pazienti e individui attivi nel settore delle malattie rare, guidata dai pazienti.

La missione di Eurordis è quella di costruire una comunità pan-Europea di organizzazioni di pazienti e persone che vivono con le malattie rare, per essere la loro voce a livello Europeo, e direttamente o indirettamente, lottare

¹² D'ora in poi EURORDIS.

contro l'impatto delle malattie rare sulle loro vite. A questo fine, Eurordis ha intrapreso diverse attività in nome dei suoi membri, in favore di:

- Rafforzare dei gruppi di pazienti di malattie rare;
- Promuovere le malattie rare come argomento di sanità pubblica;
- Accrescere la sensibilizzazione pubblica sulle malattie rare, e anche quella delle Istituzioni Nazionali e Internazionali;
- Migliorare l'accesso all'informazione, trattamento, assistenza, e sostegno per le persone che vivono con le malattie rare;
- Incoraggiare le buone pratiche sulle malattie rare;
- Promuovere la ricerca scientifica e clinica sulle malattie rare;
- Sviluppare trattamenti e farmaci orfani;
- Migliorare la qualità della vita attraverso il sostegno ai pazienti nel sociale, welfare (benessere) e servizi educativi¹³.

EURORDIS promuove, inoltre, lo sviluppo di servizi adatti alla situazione e alle esigenze specifiche delle singole persone affette da malattie rare. È inoltre, al centro del processo legislativo. Il suo contributo è stato di fondamentale importanza per l'adozione di un'importante legislazione sulle malattie rare e i farmaci orfani a livello europeo, come il Regolamento Europeo sui Farmaci Orfani, i Farmaci Pediatrici e le Terapie Avanzate. Ha numerose collaborazioni con le federazioni nazionali per le malattie rare in diversi paesi, aumentando così la propria influenza nei processi nazionali e per l'adozione e sviluppo di piani o strategie nazionali per le malattie rare in ogni Paese europeo.

L'attività di EURORDIS si svolge anche attività di formazione per rafforzare la capacità dei rappresentanti dei malati di partecipare alle sperimentazioni cliniche e di essere coinvolti nello sviluppo dei farmaci e nel

¹³ Cfr. <<http://www.eurordis.org/it/content/la-nostra-missione>>

processo legislativo europeo. la formazione rafforza la capacità dei rappresentanti dei malati di promuovere con efficacia la causa delle malattie rare¹⁴.

1.1.2.3 Orphanet

Orphanet, è gestito da un consorzio europeo di cui fanno parte una quarantina di paesi, il cui coordinamento ha sede in Francia. I team nazionali hanno il compito di raccogliere informazioni sulle consulenze specialistiche, sui laboratori di diagnosi, sulle attività di ricerca in corso e sulle associazioni di pazienti nei rispettivi paesi. Tutti i team aderiscono alle norme di qualità di Orphanet. Il team coordinatore francese è responsabile della gestione del database e del sito web, del controllo di qualità, dell'elenco delle malattie rare, delle classificazioni e dell'edizione dell'enciclopedia di Orphanet.

Orphanet è gestito, inoltre, da diversi comitati che si occupano in modo indipendente della supervisione del progetto, per assicurarne la coerenza, lo sviluppo tecnologico e la continuità. è impegnato sia a livello europeo che internazionale in diverse attività. In particolare, a livello europeo esiste un direttivo costituito dai rappresentanti delle agenzie che finanziano la gestione del database e del sito web e il coordinamento europeo di Orphanet; un comitato di coordinamento costituito dai coordinatori nazionali di Orphanet, con a capo il direttore dell'unità Inserm-Orphanet; un comitato editoriale composto da oltre 100 esperti internazionali. A livello nazionale, esiste un comitato direttivo e/o scientifico a seconda del paese (uno per ognuno dei 38 paesi partecipanti), costituito da esperti nazionali con competenze che coprono tutte le aree mediche.

La gestione del database e del sito web e il coordinamento sono cofinanziati dall'Inserm (*Institut National de la Santé et de la Recherche Médicale*), dal Dipartimento Generale della Sanità francese e dalla Commissione

¹⁴ Cfr. <<http://www.eurordis.org/it/training-resources>>

Europea, mentre alcuni servizi specifici sono finanziati da altri partner.

Le attività di Orphanet a livello nazionale sono allo stesso modo finanziate dalle istituzioni dei rispettivi paesi e/o da fondi specifici.

Orphanet si impegna a sostenere, aggiornare e sviluppare un database on-line, totalmente libero e gratuito, dedicato alle malattie rare e ai farmaci orfani. La raccolta dei dati e la diffusione delle informazioni si attengono alle disposizioni legali in vigore nei vari Paesi impegnati nel progetto: codice etico professionale, legge sull'elaborazione dati e libertà, sui diritti di proprietà intellettuale e qualsiasi altra legge o regolamento applicabile. Il comitato francese di protezione dei dati personali (*Commission Nationale de l'Informatique et des Libertés; CNIL*) ha dato parere favorevole alla creazione di *Orphanet* il 5 Maggio 1997.

Tutte le informazioni e i servizi presenti in Orphanet sono conformi ai codici e alle indicazioni emanati da comitati etici ad hoc, riconosciuti a livello nazionale e internazionale, riguardanti specialmente il rispetto dei diritti dei pazienti, il rispetto della confidenzialità delle informazioni, la pratica della medicina on-line e la sicurezza dei network.

Ad oggi, Orphanet ha aderito ai seguenti codici e norme:

- l'*HONcode 'HONConduct385483'*¹⁵, l'*eHealth Code of Ethics*¹⁶;
- le '*Guidelines for Medical and Health Information Sites on the Internet*' dell'*Associazione Americana dei Medici*¹⁷;

¹⁵ L'Health On the Net Foundation (HON) è stata fondata con lo scopo di promuovere la diffusione di informazioni sanitarie di qualità per i pazienti, i professionisti e il grande pubblico, facilitando l'accesso ai dati più recenti e rilevanti attraverso Internet.

La certificazione HONcode è uno standard etico finalizzato a fornire informazioni sanitarie di qualità. Esso evidenzia l'intento di un sito web di voler pubblicare informazioni trasparenti. La trasparenza del sito migliora l'utilità dello stesso e l'obiettività delle informazioni, garantisce, inoltre la correttezza dei dati pubblicati. <www.hon.ch/HONcode>

¹⁶ <ihealthcoalition.org/ethics/ehcode.html>

- le indicazioni del *Consiglio Nazionale Francese dell'Ordine dei Medici (Conseil National de l'Ordre des Médecins)* ¹⁸.

Il database è posto sotto la responsabilità di un comitato scientifico e di un comitato editoriale, i cui membri sono nominati sulla base della loro esperienza e competenza nel campo delle malattie rare, su proposta delle società scientifiche, delle autorità della salute dei Paesi impegnati nel progetto, o di qualsiasi altra associazione del settore. Ogni qualvolta vengono inserite nuove informazioni, un membro del Comitato Scientifico ne valida la pertinenza e la correttezza. Di seguito in figura 1, la schermata di rappresentazione di un termine all'interno del database Orphanet.

:: Aniridia

Numero Orpha	: ORPHA77	ICD-10	: Q13.1
Sinonimo/i	: Delezione 11p	OMIM	: 106210.1 ¹⁷
Prevalenza	: 1-9 / 100 000	UMLS	: C0003076 C2931795
Trasmissione	: Autosomica dominante Sporadica	MeSH	: C538293 D015783
Età di esordio	: Infanzia	MedDRA	: 10002532
		SNOMED CT	: 69278003

RIASSUNTO

L'aniridia è un difetto congenito dell'occhio, caratterizzato dall'assenza parziale o totale dell'iride. Può essere isolata o fare parte di una sindrome (aniridia isolata e sindromica; si vedano questi termini).

L'incidenza annuale dell'aniridia isolata è stimata in 1/64.000-1/96.000.

L'aniridia isolata può associarsi ad altre anomalie oculari, come la cataratta, il glaucoma (di solito durante l'adolescenza), il panno corneale, l'ipoplasia del nervo ottico, l'assenza del riflesso maculare, l'ectopia del cristallino, il nistagmo e la fotofobia, che in genere esitano di per sé in una diminuzione della vista. L'aniridia sindromica si associa ad altre anomalie non oculari come il nefroblastoma (tumore di Wilms) o altre anomalie urogenitali, il deficit cognitivo nel caso della sindrome WAGR (si veda questo termine) o altre anomalie dello sviluppo nelle sindromi più rare.

Revisore(i) esperto(i)
Pr Michael PETERSEN
Ultimo aggiornamento: Marzo 2012

Informazioni supplementari
Ulteriori informazioni su questa malattia
> Classificazione(i) (2)
> Gene(i) (3)
> Pubblicazioni in PubMed [¹⁷]
> Altri siti (7)
Risorse mediche per questa malattia
> Centri specializzati (76)
> Test diagnostici (103)
> Associazioni dei pazienti (20)
> Farmaco(i) orfano(i) (0)
Attività di ricerca su questa malattia
> Progetti di ricerca (11)
> Ricerche cliniche (1)
> Registri & biobanche (10)
> Network (2)
Quaderni di Orphanet
> Prevalenza

Figura 1 Schermata di rappresentazione di un termine all'interno del database Orphanet.

¹⁷ <www.ama-assn.org/ama/pub/category/1905.html>

¹⁸ <www.conseil-national.medecin.fr>

Come evidenziato nella figura 1, ogni termine inserito all'interno del database di Orphanet è corredato da:

- Numero ORPHA: identificativo all'interno del database;
- Sinonimo/i: tutti i sinonimi del termine in oggetto;
- Prevalenza: stima della prevalenza della singola patologia;
- Trasmissione: come la malattia si trasmette e la frequenza con la quale si trasmette;
- Età di esordio: quando la malattia si manifesta
- I codici corrispondenti del termine all'interno delle principali classificazioni internazionali (nell'esempio: ICD10, OMIM, UMLS, MeSH, MeDRA, SNOMED CT);
- Riassunto: descrizione dettagliata della patologia in oggetto;
- Revisore/i esperto/i: identificativo dell'esperto che ha revisionato la malattia;
- Ultimo aggiornamento: data ultima di revisione della malattia: le informazioni vengono aggiornate ogni qual volta i progressi scientifici lo richiedano, o almeno una volta l'anno per tutti i dati presenti nel database.

1.1.2.4 Office of Rare Diseases Research (ORDR)

L'Office of Rare Diseases Research¹⁹ presso il National Center for Advancing Translational Sciences (NCATS) sostiene e coordina la ricerca sulle malattie rare. L'ORDR, inoltre, fornisce il proprio supporto a pazienti che sono affetti da una delle migliaia tra le malattie rare conosciute oggi. Per poter svolgere al meglio le sue attività, ORDR ha attivato una rete di

¹⁹ D'ora in poi ORDR.

collaborazione con vari istituti con il compito di coordinare e promuovere i rapporti tra i vari soggetti così come altri Istituti e Centri di NIH.

The screenshot shows the ORDR (Office of Rare Diseases Research) website. At the top, it features the U.S. Department of Health & Human Services logo and the NIH motto "Turning Discovery Into Health". The main header includes the ORDR logo and a search bar. Below the header is a navigation menu with categories like "Rare Diseases Information", "Patient Advocacy Groups", "Research & Clinical Trials", "Genetic & Rare Diseases Information Center", and "Scientific Conferences". The main content area is titled "Genetic and Rare Diseases Information Center (GARD)" and displays a page for "Aniridia". The page includes a breadcrumb trail, a "printer friendly version" link, and a detailed description of the condition. The description states that Aniridia is an eye disorder characterized by a complete or partial absence of the colored part of the eye (the iris). It also mentions associated symptoms like increased sensitivity to light (photophobia) and reduced visual acuity. The page concludes with a "References" section listing two sources: a Genetics Home Reference entry from June 2009 and a GeneReviews entry from August 2008.

Figura 2 Schermata di visualizzazione della rappresentazione di un termine all'interno dell'ORDR

All'interno del sito dell'ORDR è possibile visualizzare, in ordine alfabetico, le malattie presenti, come si evince dalla figura, il termine ha una descrizione dettagliata e i riferimenti bibliografici per ulteriori ricerche attraverso la rete della National Library of Medicine.

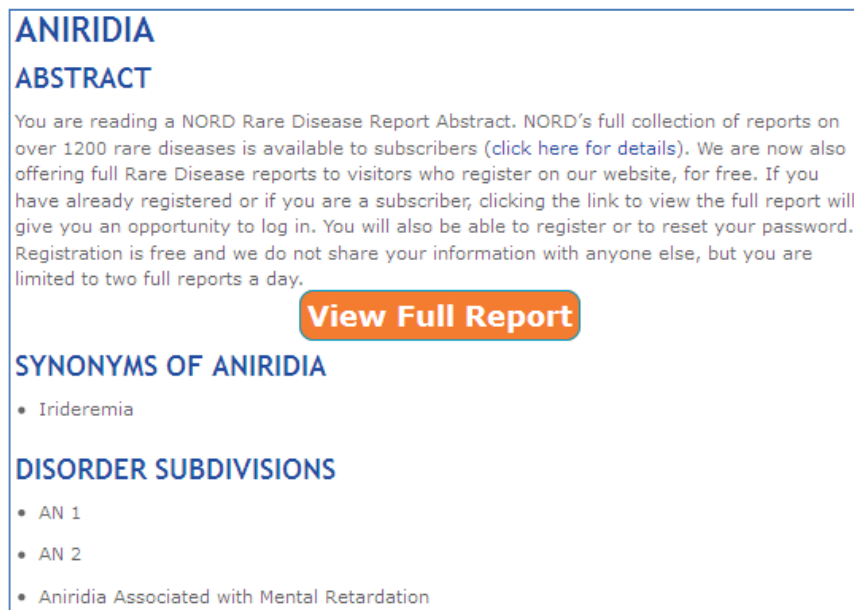
1.1.2.5 La National Organization for Rare Disorders (NORD)

L'Organizzazione Nazionale per le Malattie Rare²⁰, è una federazione di organizzazioni sanitarie di volontariato dedicate ad aiutare le persone con rare malattie "orfane" e di assistenza alle organizzazioni che li servono. NORD è impegnata nella identificazione, il trattamento e la cura delle malattie rare attraverso programmi di educazione, sensibilizzazione, ricerca e servizio.

I punti di forza della NORD sono:

- La coscienza nazionale e il riconoscimento del peso delle sfide affrontate dalle persone affette da malattie rare ed i relativi costi per la società;
- La convinzione che una nazione in cui le persone affette da malattie rare può e deve garantire l'accesso alla diagnostica e alle terapie che estendono e migliorare la loro vita;
- Una cultura sociale, politica e finanziaria di innovazione che supporta la ricerca di base necessaria per creare dei test diagnostici e terapie per tutte le malattie rare;
- Un contesto normativo che favorisca lo sviluppo e l'approvazione tempestiva di strumenti diagnostici sicuri ed efficaci e trattamenti per i pazienti affetti da malattie rare.

²⁰ D'ora in poi NORD.



ANIRIDIA
ABSTRACT

You are reading a NORD Rare Disease Report Abstract. NORD's full collection of reports on over 1200 rare diseases is available to subscribers (click [here](#) for details). We are now also offering full Rare Disease reports to visitors who register on our website, for free. If you have already registered or if you are a subscriber, clicking the link to view the full report will give you an opportunity to log in. You will also be able to register or to reset your password. Registration is free and we do not share your information with anyone else, but you are limited to two full reports a day.

[View Full Report](#)

SYNONYMS OF ANIRIDIA

- Irideremia

DISORDER SUBDIVISIONS

- AN 1
- AN 2
- Aniridia Associated with Mental Retardation

Figura 3 Visualizzazione della rappresentazione di un termine all'interno del database della NORD.

Come è possibile vedere in figura 3, anche qui nel database della NORD è possibile visualizzare per ogni termine diversi dettagli:

- Una lista di sinonimi, laddove presenti, del termine in oggetto;
- Un *abstract* con tutti i dettagli sulla patologia per gli utenti registrati, che possono farlo attraverso il sito senza dover pagare nulla. L'unico svantaggio è che la NORD limita gli accessi di visualizzazione dei report a due al giorno.

1.1.3 I principali studi a livello internazionale sulle malattie rare

Gli studi effettuati finora sulla rappresentazione delle malattie rare nelle terminologie biomediche sono stati relativamente esigui, molti dei quali motivati dal fatto che i creatori di risorse, come Orphanet e ORDR,

hanno iniziato ad inserire dei riferimenti per rimandare alle risorse terminologiche presenti nell'UMLS.

Lo studio condotto da Milicic et al.²¹ ha portato allo sviluppo di una metodologia per l'identificazione automatica di corrispondenze da Orphanet verso UMLS sulla base di tecniche sofisticate per il confronto dei termini che possa essere il più possibile automatico, riducendo così al minimo il confronto manuale. Questo tipo approccio è prevalentemente dedicato alle risorse presenti nel database di Orphanet ottimizzando in modo specifico il processo di allineamento per questa risorsa poiché Orphanet fornisce, per la maggior parte dei termini in esso contenuti, il riferimento verso i codici di OMIM e quelli della ICD, individuati manualmente dagli esperti del dominio.

L'analisi condotta da Merabti et al.,²² inoltre, ha portato all'implementazione di una metodologia basata sull'elaborazione del linguaggio naturale per la mappatura della terminologia di Orphanet verso il MeSH²³. I risultati ottenuti con questa metodologia sono stati infine confrontati con quelli ottenuti con la corrispondenza esatta basata su stringhe verso UMLS e ICD utilizzando i riferimenti preesistenti verso gli identificativi dei singoli termini, evidenziando come, con la nuova metodologia implementata, possano essere individuati dei nuovi termini che possano essere aggiunti al database.

Al contrario delle metodologie fin qui descritte, studiate focalizzando l'attenzione su una singola risorsa, l'approccio utilizzato nello studio che ha preceduto la costruzione del thesaurus, per l'analisi della rappresentazione delle terminologie rare nei principali sistemi di classificazione esistenti, è stato pensato per poter essere applicato su molteplici risorse di ambito

²¹ M. MILICIC-BRANDT, A. RATH, A. DEVEREAU, S. AYME, *Mapping Orphanet Terminology to UMLS*, in AIME 2011 Proceedings of the 13th conference on Artificial intelligence in medicine, pp. 194-203, Springer-Verlag:Berlin, Heidelberg, 2011.

²² T. MERABTI, M. JOUBERT, T. LECROQ, A. RATH, S.J. DARMONI, *Mapping biomedical terminologies using natural language processing tools and UMLS: mapping the Orphanet thesaurus to the MeSH*, IRBM, Vol. 31, Issue 4, September 2010, pp. 221-225.

²³ La ricerca condotta da Merabti et al. è stato svolto unicamente per la lingua francese.

specifico (in questo caso ORDR, Orphanet e NORD) individuando per le stesse, mappature verso una serie di altre terminologie che ricoprono uno spettro percentualmente più ampio della terminologia biomedica (SNOMED, ICD, OMIM, MeSH). Come nei precedenti studi, anche in questo è stato utilizzato l'UMLS come matrice per effettuare i confronti, con il vantaggio di poter mantenere, nei risultati ottenuti, le relazioni tra i termini.

1.2 Analisi del contesto nazionale

L'analisi della situazione esistente nel panorama italiano si è rivelata una fase essenziale, sia per la corretta definizione dell'approccio metodologico da adottare, privilegiando la costruzione di un thesaurus che vuole essere un completamento delle attuali classificazioni esistenti in lingua italiana, partito dalla necessità di intervenire per un'unificazione a livello terminologico, sia per un'esigenza degli utenti finali della risorsa, i pazienti affetti da patologie rare, ma anche da come esigenza emersa tra gli esperti di dominio coinvolti nelle diverse fasi dell'attività di ricerca.

Nel già citato sistema di Orphanet esiste la versione italiana, anche se non per la totalità dei termini.

Il coordinamento centrale in Italia avviene grazie al Centro Nazionale delle Malattie rare²⁴ dell'Istituto Superiore di Sanità. Il centro opera a livello nazionale svolgendo attività di censimento, aggiornamento e ricerca nel campo delle malattie rare.

Tra i vari progetti coordinati dal CNMR, citiamo EPIRARE²⁵ (European Platform for Rare Registries), piattaforma europea per registri e database finalizzata alla costruzione di un *set di dati* che –condivisibile tra i diversi Paesi.

²⁴ D'ora in poi CNMR.

²⁵ <<http://www.epirare.eu/>>

Il progetto EPIRARE è co-finanziato dall'Unione Europea e coordinato dal Centro Nazionale Malattie Rare dell'ISS. Oltre che con il Ministero della Salute e l'Istituto di Fisiologia Clinica del Consiglio Nazionale delle Ricerche, le attività sono svolte in associazione con l'Istituto di Sanità Carlos III (ES), la London School of Economics and Political Sciences (UK), l'Organizzazione Europea per le Malattie Rare (FR), l'Istituto Scientifico di Salute Pubblica (BE), il Centro Ellenico per il controllo e la prevenzione delle malattie (GR), l'Associazione bulgara per la promozione dell'Educazione e della Scienza (BG), l'Università di Maastricht (NL), il Ministero della Salute e delle Politiche Sociali (ES). Attualmente, al progetto hanno aderito numerosi altri partners, inclusi gli USA. Il Registro Nazionale delle Malattie Rare²⁶ è stato istituito presso l'Istituto Superiore di Sanità²⁷ con il Decreto Ministeriale 279/2001 (Art. 3) e ha avuto successive implementazioni mediante gli Accordi Stato-Regioni del 2002 e 2007. Il RNMR costituisce lo strumento principale di sorveglianza delle Malattie Rare su scala nazionale; l'obiettivo prioritario è la produzione di evidenze epidemiologiche a supporto sia della definizione e dell'aggiornamento dei livelli essenziali di assistenza (LEA), sia delle politiche e della programmazione nazionale. Inoltre il RNMR ha i seguenti ulteriori obiettivi:

- stimare la prevalenza e l'incidenza delle diverse malattie rare;
- caratterizzarne la distribuzione sul territorio nazionale;
- stimare la migrazione sanitaria dei pazienti;
- stimare il ritardo diagnostico²⁸.

²⁶ D'ora in poi RNMR.

²⁷ D'ora in poi ISS.

²⁸ <http://www.salute.gov.it/dettaglio/dettaglioNews.jsp?id=2233&tipo=new>. Del RNMR si dà ampio dettaglio nel Capitolo 3.

Capitolo 2

2.1 L'importanza degli standard nell'informatica medica

Il settore sanitario è uno forse dei più complessi dal punto di vista terminologico e contenutistico. Le informazioni in esso contenute necessitano di essere raccolte, interpretate, usate per la risoluzione di numerosi problemi. L'ausilio delle recenti innovazioni tecnologiche possono certamente velocizzare e snellire queste operazioni, ma ovvia e necessaria condizione è che le informazioni devono essere rappresentate in una forma che possa essere interoperabile, quindi condivisa da un certo numero di utenti. Nella maggior parte dei casi ciò significa codificare le informazioni utilizzando un linguaggio definito "controllato".

L'informatica medica lavora da decenni in questo senso, ma solo in tempi recenti tutta la comunità medica ha finalmente riconosciuto l'importanza delle terminologie e la necessità di avere degli standard che rendano possibile il "dialogo" tra diversi operatori e strutture. Spesso risulta assai più conveniente crearsi le proprie terminologie piuttosto che cercare di utilizzare quelle standardizzate, che nella maggior parte dei casi hanno anche dei costi di acquisizione e gestione notevoli.

Un rilevante tentativo in questo settore è quello della National Library of Medicine che con l'Unified Medical Language System raccoglie oltre 100 terminologie e le rende gratuitamente disponibili. Funzioni statistiche ed amministrative sono sempre state importanti motivazioni per la codifica dei dati sanitari, sia per remunerazioni sia per refertazione obbligatoria. Questa è stata sicuramente una delle principali forze di spinta verso la digitalizzazione delle informazioni sanitarie e biomediche.

2.2 L'Unified Medical Language System

L'Unified Medical Language System²⁹ è un sistema compositivo che prevede una rete semantica che categorizza un insieme di descrittori provenienti dall'analisi di diversi sistemi di prima generazione (cross-thesaurus).

Nel settore biomedico, come accennato, esistono diversi tipi di basi di conoscenza create e curate da diversi soggetti ed in diversi contesti geografici. La diversità di linguaggi e di codifiche rende naturalmente disomogenea l'informazione strutturata. L'UMLS è stato creato proprio per ovviare a questo problema, poiché lavora in modo da gestire le sinonimie e le quasi sinonimie multilingua in modo che, ad esempio, il termine *schistorrhachis* e il suo sinonimo italiano *spina bifida* siano riferibili alla stessa unità di significato.

La costruzione delle basi terminologiche di dominio è stata avviata nel 1986 dalla National Library of Medicine e conta, ad oggi, più di centoventi lingue che vengono però sempre gestite partendo da un master terminologico inglese nativo³⁰.

La NLM è stata designata come ente di coordinamento e/o sviluppo e diffusione di mappature fra terminologie cliniche standardizzate e altre importanti terminologie all'interno del Metathesaurus UMLS.

L'UMLS è diretto da un team multidisciplinare della NLM e riceve il sostegno per l'attività di ricerca grazie a dei finanziamenti pubblici.

Non esistono, ad oggi, limitazioni riguardo a chi possa suggerire di includere un vocabolario nell'UMLS, anche se la NLM potrebbe porre dei vincoli di inclusione rispetto alla tempistica richiesta per le fasi di analisi, processing o editing. Gli aggiornamenti sono distribuiti due volte l'anno, aprile e novembre, e sono rilasciati gratuitamente.

²⁹ D'ora in poi UMLS.

³⁰ <<http://www.nlm.nih.gov/research/umls/>>

Il punto di forza dell'UMLS risiede, comunque, nella struttura che risulta essere così composta:

- Metathesaurus
- Semantic Network
- Specialist Lexicon& Tools

2.2.1 IL Metathesaurus

Il Metathesaurus è organizzato per concetti: ne riflette e conserva il significato, la denominazione e le sue relazioni dal vocabolario da cui lo trae. Ogni concetto o significato ha nel Metathesaurus un Concept Unique Identifier³¹ permanente, ovvero un identificativo assegnato. Il CUI non ha significato intrinseco e non cambia mai, nonostante i cambiamenti nel tempo, nel Metathesaurus o nei vocabolari da cui provengono i termini. Se un CUI viene rimosso dal o duplicato, se ne tiene uno solo dei due, al quale si collegano tutte le relazioni, e l'altro viene eliminato.

Il Metathesaurus è un vocabolario multi-lingue e rappresenta la base di dati contenente le informazioni sulla molteplicità dei concetti rappresentanti l'ambito biomedico, le differenze terminologiche e le relazioni che intercorrono tra essi. Lo scopo per cui è stato progettato è quello di uniformare tutte le versioni elettroniche dei sistemi di classificazione, thesauri, set di codici e liste controllate di termini appartenenti al settore biomedico (analisi statistiche, codici per la refertazione clinica ecc.), inseriti al suo interno.

Esso è, ovviamente, legato ad altre basi di conoscenza presenti all'interno dell'UMLS: la *Semantic Network* e lo *Specialist Lexicon* che consentono il legame tra i diversi concetti, poiché ogni termine inserito

³¹ D'ora in poi CUI.

all'interno dell'UMLS viene assegnato ad una precisa categoria semantica, pur mantenendo i significati di ogni singolo termine inserito nel database, così come per relazioni con altri termini e l'appartenenza ai singoli vocabolari da cui ogni termine proviene. Quando due vocabolari utilizzano termini differenti per descrivere lo stesso concetto, il Metathesaurus li include entrambi, associandoli come sinonimi e differenziandoli a seconda del vocabolario di appartenenza, includendo, ovviamente, laddove presente, la gerarchia all'interno del quale il termine è inserito.

2.2.2 Semantic Network

Lo scopo della *Semantic Network* è di fornire una categorizzazione coerente di tutta la molteplicità dei concetti inseriti all'interno del Metathesaurus, fornendo, inoltre, tutte le relazioni che intercorrono tra i concetti, siano esse di sinonimia o di gerarchia, diventando quindi una sorta di *authority control*. La definizione di queste relazioni è stabilita grazie a delle descrizioni testuali e mediante le informazioni inerenti le gerarchie ereditate dai sistemi classificatori da cui i termini provengono.

Le tipologie semantiche rappresentano i nodi della *Semantic Network* mentre le relazioni tra i termini rappresentano i collegamenti tra gli stessi. Le categorie semantiche principali rappresentano organismi, strutture anatomiche, funzione biologica, prodotti chimici, oggetti fisici e concetti o idee. Il campo di applicazione attuale delle diverse categorie semantiche è molto ampia, consentendo la categorizzazione semantica di una vasta gamma di terminologia in domini diversi³².

Il significato di ogni termine inserito all'interno del Metathesaurus è contestualizzato in relazione al vocabolario di appartenenza, attraverso delle

³² ALEXA T. MCCRAY, ANITA BURGUN, OLIVIER BODENREIDER, *Aggregating UMLS Semantic Types for Reducing Conceptual Complexity*, IOS Press, Volume 84/2001, MedINFO, Amsterdam: IOS Press, pp. 216-220.

definizione (ove presenti) o individuando il suo posto in un eventuale gerarchia³³, sinonimi e altre relazioni tra i termini indicati.

Ad ogni termine viene assegnato una tipologia semantica quanto più specifica possibile, altrimenti viene associato ad una più generica. Ad esempio, il termine "Macaco"³⁴ viene assegnato alla categoria "mammifero" perché non è presente nella *Semantic Network* una tipologia più specifica, che dovrebbe essere, in questo caso "primati". Il livello di granularità varia a seconda della rete, ma questo aspetto ha delle implicazioni notevoli per la corretta interpretazione del significato che è stato assegnato ai termini all'interno del Metathesaurus.

2.2.3 Specialist Lexicon

Lo *Specialist Lexicon* è composto da lessico di inglese generale che include al suo interno termini di ambito biomedico. Questo strumento lessicale è stato progettato per affrontare l'alto grado di variabilità dei termini presenti nel linguaggio naturale che, molto spesso, hanno diverse forme flesse e che, a tale proposito, possono anche essere considerate istanze della stessa parola.

L'unità lessicale di base è una struttura che consiste di spazi (*slots*) e da riempitivi (*fillers*). *"Each lexical record has a base= slot whose filler indicates the base form, and optionally a set of spelling_variants= slots to indicate spelling variants. An "entry=" slot records the unique identifier (EUI) of the record. EUI numbers are seven digit numbers preceded by an "E". Each*

³³ Non tutti i termini inseriti all'interno del Metathesaurus provengono da sistemi classificatori in cui si presente una gerarchia, alcuni come già accennato in precedenza, sono semplici liste di termini controllati.

³⁴ Il Macaco (*Macaca*) è un genere di primati della famiglia dei Cercopitecidi. I macachi sono il genere di primati a più ampia diffusione dopo l'uomo. Sono diffusi in larga parte dell'Asia: dall'Afghanistan alla Cina, dall'India al sudest asiatico al Giappone. Inoltre una specie (*Macaca sylvanus*, detta comunemente bertuccia) vive nell'Africa nord-occidentale. In Europa è presente una sola specie, *Macaca sylvanus*, diffusa a Gibilterra.

*record has a cat= slot indicating part of speech. The lexical record is delimited by braces {...}*³⁵.”

L'esempio sotto riportato mostra un esempio delle caratteristiche possedute dalla singola unità lessicale dello Specialist Lexicon:

```
{base=anesthetic spelling_variant=anaesthetic entry=E0354094 cat=noun  
variants=regvariants=uncount}{base=anestheticspelling_variant=anaestheticentry  
=E0330019 cat=adj variants=inv position=attribposition=pred stative}36
```

La forma di base "anesthetic" e la sua variante "anaesthetic" appaiono in due record lessicali distinti, una come aggettivo, l'altra come sostantivo. I termini sono quindi divisi dal punto di vista morfologico, ma non sul piano semantico.

2.3 Analisi terminologica

L'UMLS è stato lo strumento che ha reso possibile l'avvio reale del presente progetto di ricerca. La base di partenza è stata, infatti, l'analisi puntuale della rappresentazione delle malattie rare, dal punto di vista terminologico, all'interno dei principali sistemi di classificazione esistenti a livello internazionale; il lavoro di analisi è stato condotto presso la National Library of Medicine.

Considerata la vastità dell'argomento in relazione ai limiti temporali imposti, il lavoro è stato focalizzato solo su alcune terminologie specifiche dell'ambito biomedico³⁷ quali: il Medical Subject Headings³⁸, la terminologia di riferimento per l'indicizzazione e la ricerca nella letteratura biomedica;

³⁵ < <http://www.ncbi.nlm.nih.gov/books/NBK9680/>>

³⁶ *Ibidem*.

³⁷ Le terminologie selezionate sono state anche quelle che hanno riportato una maggiore percentuale di rappresentazione dei termini delle malattie rare.

³⁸ D'ora in poi MeSH.

l'Online Mendelian Inheritance in Man,³⁹ base di conoscenza di riferimento sulle malattie genetiche; la Systematized Nomenclature of Medicine-Clinical Terms⁴⁰, la più vasta terminologia clinica del mondo, e la International Classification of Diseases 10th revision⁴¹, la classificazione delle malattie, dei traumatismi, degli interventi chirurgici e delle procedure diagnostiche e terapeutiche.

L'analisi terminologica si è svolta a partire da un corpus terminologico costruito dalle basi di dati provenienti dal database di Orphanet, dell'ORDR e del NORD. I corpus sono stati mappati attraverso l'UMLS, applicando appositi filtri, per effettuare una prima scrematura in modo tale da comprendere quali terminologie fossero maggiormente rappresentative del dominio oggetto di studio.

2.3.1 Acquisizione e analisi dei dati provenienti da Orphanet, dall'ORDR e dalla NORD

Per le fonti terminologiche del dominio specifico delle malattie rare sia per quanto riguarda i dati provenienti dal database di Orphanet, dall'ORDR e dalla NORD, è stato necessario, ovviamente, acquisire le prescritte autorizzazioni.

I dati provenienti da Orphanet sono stati acquisiti in formato tabellare, come mostrato nella figura sottostante:

ID Orphanet	Tipologia termine	Descrizione termine
ORPHA000881	PT	Turner syndrome

³⁹ D'ora in poi OMIM.

⁴⁰ D'ora in poi SNOMED CT.

⁴¹ D'ora in poi ICD.

ORPHA000095	PT	Friedreich ataxia
ORPHA000586	PT	Cystic fibrosis
ORPHA000846	PT	Alpha-thalassemia
ORPHA000848	SYN	Cooley anemia
ORPHA000848	PT	Beta-thalassemia
ORPHA000262	SYN	Duchenne and Becker dystrophinopathy
ORPHA000262	PT	Duchenne and Becker muscular dystrophy
ORPHA000261	SYN	EDMD
ORPHA000261	SYN	Emerinopathy
ORPHA000261	PT	Emery-Dreifuss muscular dystrophy
ORPHA000480	PT	Kearns-Sayre syndrome
ORPHA000269	SYN	Landouzy-Dejerine myopathy
ORPHA000269	PT	Facioscapulohumeral muscular myopathy
ORPHA000269	SYN	Facioscapulohumeral muscular dystrophy

Figura 4 Campione dei dati provenienti dal database di Orphanet

Come si evince dalla tabella, nella prima colonna sono presenti gli ID dei concetti⁴² all'interno del database, nella seconda colonna la tipologia del termine: PT sta ad indicare il termine principale (7,715 in totale) e SYN il relativo sinonimo (5,224 in totale), i termini che si riferiscono allo stesso concetto, ovvero alla medesima patologia, sono associati allo stesso identificativo.

Un'altra caratteristica molto importante dei termini di Orphanet è che la maggior parte dei termini hanno un collegamento con il corrispondente termine, nell'ICD-10 e OMIM, ciò è stato di fondamentale importanza in questo lavoro di ricerca per un maggiore controllo in termini di *quality assurance* in fase di valutazione dei risultati ottenuti.

I termini dell'ORDR, invece, sono stati acquisiti direttamente attraverso il sito web⁴³. I termini nella loro composizione originaria non sono legati da alcun tipo di relazione, fatta eccezione per quella di sinonimia.

⁴² Per concetto in questa sede si intende identificare la patologia all'interno del database, ogni concetto può essere rappresentato da uno o più termini, ovvero da un termine principale e da uno o più sinonimi.

⁴³ <<http://rarediseases.info.nih.gov/>>

La lista di termini, aggiornata periodicamente dall'ORDR, sul piano semantico è rappresentata da un totale di 6,857 concetti principali e 11,803 sinonimi. Per ognuno di essi, tramite il sito web, vengono fornite informazioni dettagliate sulle risorse informative liberamente accessibili al pubblico che possano aiutare quanto più possibile l'utente ad ottenere quanti più dettagli su ogni singola patologia. I termini inclusi nella lista rappresentano le malattie rare per cui è stata avanzata specifica richiesta da parte di gruppi di pazienti per ottenere maggiori informazioni presso l'ORDR ed, in particolare, presso il Genetic and Rare Diseases Information Center (GARD), centro finanziato dal ORDR, e presso il National Human Genome Research Institute (NHGRI), oppure sono termini, di diversa provenienza, che rappresentano patologie che nel corso degli ultimi 10 anni sono state definite rare dalla comunità scientifica.

La stessa metodologia è stata applicata per i termini provenienti dalla NORD, essi, seppur in numero esiguo rispetto alle prime due risorse, sono stati comunque analizzati poiché il campione dei termini rappresenta la terminologia "più vicina" a quella dell'utente finale, il paziente o familiare del paziente affetto da patologia rara. La copertura semantica in questo caso ha registrato un totale di 1,236 concetti principale con 4,562 sinonimi.

2.3.2 Le terminologie target: OMIM, SNOMED CT, MeSH, ICD10

In questo paragrafo si riporta descrizione dettagliata delle risorse utilizzate per l'analisi della mappatura della terminologia specifica del dominio delle patologie rare. Le terminologie su cui è stata focalizzata l'attenzione sono state Online Mendelian Inheritance in Man (OMIM), SNOMED-Clinical Terms (SNOMED CT), Medical Subject Headings (MeSH), e l'International Classification of Diseases 10th revision (ICD10),

2.3.2.1 Online Mendelian Inheritance in Man (OMIM)

La Online Mendelian Inheritance in Man è la terminologia sui geni e fenotipi umani più completa e aggiornata con descrizioni dettagliate sulle patologie genetiche mendeliane e su oltre 12,000 geni. Uno dei punti di forza di OMIM, in particolare, è che evidenzia le relazioni tra fenotipi e genotipi. Il vantaggio di questa terminologia è che viene quotidianamente aggiornata e ciascun record contiene dei riferimenti verso altre risorse genetiche.

Il lavoro di costruzione del database iniziò agli inizi del 1960 grazie al dott. Victor A. McKusick, che cominciò a costruire il catalogo delle malattie mendeliane e lo intitolò *Mendelian Inheritance in Man* (MIM), che vide nel tempo dodici riedizioni tra il 1966 e il 1998.

La versione che diede vita a quella attualmente in uso, fu creata nel 1985 grazie ad una collaborazione, tuttora esistente, con la National Library of Medicine e William H. Welch della *Welch Medical Library* presso la *John Hopkins University School of Medicine*, sotto la direzione della dott.ssa Ada Hamosh. La nuova ottica di lavoro fu orientata alla realizzazione di una versione che potesse essere pubblicata sulle prime reti internet, cosa che venne attuata poi nel 1987.

OMIM, all'atto della creazione, fu pensato in particolare per l'utilizzo da parte dei medici di base, ma anche per favorire la ricerca genetica, quindi per specialisti del settore, ricercatori, e giovani studenti. Il database oggi è liberamente consultabile da parte del pubblico, senza restrizioni, tuttavia non sempre le informazioni pubblicate sono di immediata comprensione se non da parte del personale estremamente qualificato; non è adatto quindi a fornire risposte immediate a cittadini/pazienti che necessitano informazioni su particolare condizioni medico-genetiche.

2.3.2.2 Systematized Nomenclature of Medicine (SNOMED-CT)

La *Systematized Nomenclature of Medicine* (SNOMED-CT) è la più vasta terminologia al mondo in ambito clinico creata dalla *International Health Terminology Standard Development Organization*⁴⁴ (IHTSDO) per l'utilizzo all'interno delle cartelle cliniche elettroniche. La terminologia di SNOMED-CT copre la maggior parte delle aree di informazione clinica come: patologie, sintomi, procedure diagnostiche, ecc.

SNOMED-CT è uno strumento utile per l'indicizzazione e recupero di informazioni, ma anche per l'aggregazione di dati clinici, per ridurre il rischio di errore nei casi di differenti sistemi di codifica utilizzati nel *clinical care* nella ricerca e per la salute dei pazienti. SNOMED-CT è fondamentale nell'interoperabilità e nello scambio dei dati clinici, l'elevato livello di dettaglio al suo interno contribuisce al supporto decisionale, percorsi di cura, rilevazioni statistiche ecc.⁴⁵.

Esistono, inoltre, numerose traduzioni che contribuiscono ad aumentare l'internazionalità della classificazione. Le traduzioni in diverse lingue contribuiscono altresì ad aumentare i termini che vengono utilizzati a livello locale nelle singole nazioni.

Da un punto di vista quantitativo SNOMED CT contiene più di 311,000 concetti, associati ad un identificativo univoco (ad esempio il concetto 22298006 si riferisce solo ed esclusivamente a *Myocardial infarction*). Tutti i concetti all'interno della classificazione sono organizzati grazie ad una classificazione gerarchica con una relazione di tipo *IS-A*, ad esempio:

Viral pneumonia IS-A Infectious pneumonia

Infectious pneumonia IS-A Pneumonia

Pneumonia IS-A Lung disease.

⁴⁴ D'ora in poi IHTSDO.

⁴⁵ RUCH PATRICK, GOBEILL JULIEN, LOVIS CHRISTIAN, GEISSBÜHLER ANTOINE, (2008) *Automatic medical encoding with SNOMED categories*, BMC Medical Informatics and Decision Making, 2008.

I concetti possono avere più legami gerarchici all'interno della classificazione, ad esempio, *Infectious pneumonia* è anche un termine più specifico di *Infectious disease*.

I concetti all'interno di SNOMED CT sono collegati tra di loro con più di 1,360,000 relazioni, mentre ogni singolo concetto contiene dettagli grazie a diversi termini o descrizioni ad esso collegato che si dividono in *Fully Specified Names* (FSNs), *Preferred Terms* (PTs), and *Synonyms*. Ciascun concetto ha uno ed uno solo FSN, che è unico all'interno di SNOMED CT, come anche per il PT, stabilito da un comitato di esperti come il termine più comunemente utilizzato dalla comunità scientifica che allo stesso tempo è anche il più rappresentativo.

È possibile, per ciascun termine, avere da zero a più sinonimi, questi ultimi non devono per forza avere la prerogativa di essere unici, poiché sono dei rimandi che riportano al concetto principale.

SNOMED-CT è, inoltre, un thesaurus multilingue con alla base un'ontologia; grazie a questo sistema è possibile gestire la sinonimia di termini provenienti anche da diverse lingue, ad esempio:

*"Acute coryza", "Acute nasal catarrh",
"Acute rhinitis", "Common cold"*

e le relative versioni in spagnolo:

"resfrío común", "rinitis infecciosa"

hanno il medesimo ID 82272006.

Le primissime versioni di SNOMED avevano una struttura a faccette ordinata sulla base di assi semantici che richiedevano un'organizzazione dei concetti molto più complessa e un differente sistema di codifica. Questo

sistema, oltre ad essere più costoso in termini di gestione globale dell'intero sistema, non era abbastanza *user-friendly* così fu deciso di optare per il corrente sistema classificatorio per non venire meno agli obiettivi base per cui SNOMED-CT è stato creato. SNOMED CT, ad oggi, è gestito e distribuito per il suo utilizzo dall'*International non-profit Standards Development Organization*, con sede a Copenhagen, Danimarca. L'utilizzo di tale sistema è vincolato dall'acquisto di una licenza. I LDCs⁴⁶ possono utilizzare SNOMED CT senza alcuna licenza a scopo di ricerca scientifica nell'informatica medica, per studi su nuovi possibili utilizzi della classificazione.

La versione completa di SNOMED CT è disponibile all'interno dell'Unified Medical Language System per gli utenti che hanno sottoscritto l'accordo⁴⁷.

2.3.2.3 International Classification of Diseases 10th revision ICD-10

L'ICD-10 è la decima revisione della ICD, ossia la classificazione internazionale delle malattie e dei problemi correlati, proposta dall'Organizzazione Mondiale della Sanità in cui sono classificate ben oltre duemila malattie. La traduzione ufficiale in lingua italiana, è stata effettuata a cura dell'ISTAT e dell'Ufficio di Statistica del Ministero della Salute, ed è stata pubblicata dall'Organizzazione mondiale della sanità a Ginevra nel 2000 e a Roma nel 2001⁴⁸.

Le differenze rispetto alla revisione precedente, denominata ICD-9, sono l'aumento dei settori portati da 17 a 19, e l'utilizzo di un codice alfanumerico in sostituzione del precedente codice numerico. Non vi è una

⁴⁶ *Least Developed Countries*, i paesi meno sviluppati.

⁴⁷ Proprio per i motivi sopra citati, la versione utilizzata in questo studio risale al 31 luglio 2010, poiché era la versione disponibile al momento dello stage effettuato presso la National Library of Medicine, dove è stata svolta ampia parte dell'analisi terminologica utile per il presente lavoro di ricerca.

⁴⁸ MINISTERO DELLA SANITÀ, *ICD-10 : classificazione statistica internazionale delle malattie e dei problemi sanitari correlati:10^a revisione*, 3 volumi, Roma: Istituto poligrafico e Zecca dello Stato, 2001.

corrispondenza tra settori e lettere dell'alfabeto, poiché alcuni capitoli ne utilizzano più di una.

Di seguito, nella tabella 1, l'elenco dei capitoli⁴⁹:

Capitoli	Sezioni	Titolo
I	A00-B99	Malattie infettive e parassitarie
II	C00-D48	Neoplasie
III	D50-D89	Malattie del sangue e del sistema immunitario
IV	E00-E90	Malattie endocrine, nutrizionali e metaboliche
V	F00-F99	Patologie mentali e del comportamento
VI	G00-G99	Neuropatie
VII	H00-H59	Oftalmopatie
VIII	H60-H95	Patologie dell'orecchio
IX	I00-I99	Disturbi del sistema circolatorio
X	J00-J99	Disturbi del sistema respiratorio
XI	K00-K93	Malattie gastroenterologiche

⁴⁹ <http://it.wikipedia.org/wiki/ICD-10>.

XII	L00-L99	Dermatopatie
XIII	M00-M99	Patologie muscoloscheletriche e connettivali
XIV	N00-N99	Disturbi genitourinari
XV	O00-O99	Gravidanza, parto e puerperio
XVI	P00-P96	Disturbi congeniti
XVII	Q00-Q99	Malformazioni e anomalie cromosomiche
XVIII	R00-R99	Segni, sintomi e dati di laboratorio patologici non altrimenti classificati
XIX	S00-T98	Traumatologi, tossicologia e altre cause esterne di malattia
XX	V01-Y98	Altre cause esterne di morbilità e mortalità
XXI	Z00-Z99	Fattori che possono influenzare lo stato di benessere
XXII	U00-U99	Codici speciali

Tabella 1 Elenco dei capitoli presenti in ICD-10.

Di seguito, si dà una breve descrizione per ogni capitolo sulle principali malattie incluse:

Capitolo I: malattie infettive e parassitarie. Del primo gruppo fanno parte le malattie infettive intestinali (A00-A09), la tubercolosi (A15-A19), le

malattie batteriche trasmesse da animali (A20-A28), infezioni a trasmissione sessuale (A50-A64), malattie da spirochete (A65-A69) e da Chlamydia (A70-A74), rickettsiosi (A75-A79), infezioni al sistema nervoso centrale (A80-A89) e febbri virali da artropodi ed emorragiche (A90-A99). Nel secondo gruppo figurano infezioni che interessano la cute (B00-B09), l'epatite virale (B15-B19), l'infezione da HIV (B20-B24), malattie virali come la mononucleosi infettiva (B25-B34), micosi (B35-B49), malattie protozoarie (B50-B64), elmintiasi (B65-B83), infestazioni come la scabbia (B85-B89), batteri e virus (B95-B97).

Capitolo III: tumori divisi in maligni (C00-C97), benigni (D10-D36) ed in situ (D00-D09). I tumori maligni sono ulteriormente suddivisi per zona d'interesse. La categoria include il mesotelioma (C45), il sarcoma di Kaposi (C46) ed il morbo di Hodgkin (C81).

Capitolo III: malattie del sangue ed alcuni disturbi legati al sistema immunitario. Le anemie sono distinte in nutrizionali (D50-D53) ed emolitiche (D55-D59). La categoria include l'agranulocitosi (D70), la sarcoidosi (D86) ed altre immunodeficienze.

Capitolo IV: disturbi della tiroide (E00-E07), il diabete mellito (E10-E14) e disturbi di varie ghiandole endocrine. Nella categoria figurano anche malnutrizione (E40-E46), obesità (E65-E68), disturbi metabolici (E70-E90) e deficit nutrizionali (E50-E64).

Capitolo V: disturbi psichici e comportamentali di natura organica (F00-F09), dovuti all'uso di sostanze psicoattive (F10-F19), affettivi (F30-F39), nevrotici (F40-F48), legati a disfunzioni fisiologiche (F50-F59), disturbi della personalità (F60-F69), dello sviluppo psicologico (F80-F89) e comportamentali (F90-F98). Sono elencate tra le malattie la schizofrenia (F20) e vari ritardi mentali (F70-F79).

Capitolo VI: malattie del sistema nervoso e comprende malattie infiammatorie del sistema nervoso centrale (G00-G09), disturbi episodici e parossistici (G40-G47), disturbi dei nervi, delle radici e dei plessi nervosi

(G50-G59), polineuropatie (G60-G64) e paralisi cerebrale (G80-G83). Nella categoria figurano la malattia di Huntington (G10), il morbo di Parkinson (G20), il morbo di Alzheimer (G30), la sclerosi multipla (G35), l'epilessia (G40), l'emicrania (G43), l'emiplegia (G81) e l'idrocefalo (G91).

Capitolo VII: malattie dell'occhio. Comprende disturbi delle palpebre, dell'apparato lacrimale e dell'orbita oculare (H00-H06), della congiuntiva (H10-H13), della sclera, della cornea, dell'iride e del corpo ciliare (H15-H22), del cristallino (H25-H28), della coroide e della retina (H30-H36), del corpo vitreo (H43-H45) e del nervo ottico (H46-H48).

Capitolo VIII: malattie dell'orecchio e del processo mastoideo. I disturbi elencati coinvolgono l'orecchio esterno (H60-H62), l'orecchio medio (H65-H75) e l'orecchio interno (H80-H83).

Capitolo IX: malattie dell'apparato circolatorio. Tra i vari tipi di cardiopatie figurano le cardiopatie reumatiche (I05-I09) e le cardiopatie ischemiche (I20-I25). Include anche malattie che coinvolgono arterie, arteriole, capillari, vene e vasi linfatici.

Capitolo X: malattie dell'apparato respiratorio. Comprende infezioni e disturbi delle vie respiratorie, influenza, polmonite oltre che malattie della pleura.

Capitolo XI: malattie del sistema digerente. I disturbi elencati coinvolgono la bocca, le ghiandole salivari, l'esofago, lo stomaco, il fegato, il pancreas, l'intestino, il duodeno, l'appendice ed il peritoneo. **Capitolo XII:** disturbi che coinvolgono la pelle. Comprende infezioni della cute (L00-L08), dermatite ed eczema (L20-L30), orticaria ed eritema (L50-L54).

Capitolo XIII: malattie delle ossa, dei muscoli e dei tessuti connettivi che sono suddivise in artropatie (M00-M25), dorsopatie (M40-M54), osteopatie e condropatie (M80-M94), disturbi del tessuto connettivo (M30- M36) e dei tessuti molli (M60-M79).

Capitolo XIV: malattie dell'apparato uro-genitale. Comprende calcoli urinari (N20-N23), malattie dell'apparato genitale maschile (N40-

N51), delle mammelle (N60-N64) e disturbi che colpiscono i reni e gli ureteri (N25-N29).

Capitolo XV: gravidanza, parto e puerperio. Vi sono elencati vari tipi di aborto (O00-O08), disturbi materni (O20-O29), complicanze del travaglio e del parto (O60-O75) e vari tipi di parto (O80-O84). Comprende l'eclampsia (O15), il polidramnios (O40) e la placenta previa (O44).

Capitolo XVI: malattie del periodo perinatale. Comprende disturbi legati a complicanze della gravidanza, del travaglio e del parto (P00-P04), traumi da parto (P10-P15), infezioni (P35-P39), disturbi respiratori e cardiovascolari (P20-P29), emorragici ed ematologici (P50-P61), endocrini e metabolici (P70-P74).

Capitolo XVII: malformazioni e deformazioni congenite ed anomalie cromosomiche. Sono divise per area d'interesse: sistema nervoso (Q00-Q07), occhio, orecchio, viso e collo (Q10-Q18), apparato circolatorio (Q20-Q28), apparato respiratorio (Q30-Q34), sistema digerente (Q38-Q45), apparato genitale (Q50-Q56) ed apparato urinario (Q60-Q64).

Capitolo XVIII: i sintomi non elencati nelle altre categorie, divisi per area d'interesse: apparato circolatorio e respiratorio (R00-R09), sistema digerente (R10-R19), pelle ed ipoderma (R20-R23), sistema nervoso R25-R29) ed apparato urinario (R30-R39). Sono inoltre inclusi i risultati anormali di alcuni esami, quali gli esami del sangue (R70-R79) e delle urine (R80-R82).

Capitolo XIX: traumatismi ed avvelenamenti. I traumatismi sono divisi per zona d'interesse: della testa (S00-S09), del collo (S10-S19), del torace (S20-S29), della spalla e del braccio (S40-S49), del gomito e dell'avambraccio (S50-S59), del polso e della mano (S60-S69), dell'anca e della coscia (S70-S79), del ginocchio e della gamba (S80-S89), della caviglia e del piede (S90-S99).

Capitolo XX: cause esterne di morbosità e mortalità. Include incidenti da trasporto (V01-V99), cadute (W00-W19), annegamento (W65- W74),

contatto con animali e piante velenose (X20-X29), autolesione intenzionale (X60-X84) ed aggressione (X85-Y09).

Capitolo XXI: fattori influenzanti lo stato di salute ed il ricorso a servizi medici. Sono elencate visite mediche, necessità di vaccinazione, misure contraccettive, donazione di organi, convalescenza e anamnesi.

Capitolo XXII: malattie con codici speciali, spesso assegnati temporaneamente. Un esempio è la SARS (U04).

In Italia la Direzione Centrale Salute Integrazione Sociosanitaria e Politiche Sociali della Regione Friuli Venezia Giulia è riconosciuta – dal gennaio 2010 - come Centro Collaboratore italiano dell'Organizzazione Mondiale della Sanità per la Famiglia delle Classificazioni Internazionali .

Il Centro, riconosciuto anche dal Ministero della Salute, si occupa dello sviluppo, implementazione, utilizzo e diffusione delle classificazioni internazionali, in primis ICD e ICF⁵⁰.

2.3.2.4 Medical Subject Headings (MeSH)

Il MeSH è un thesaurus costituito da oltre 24.000 descrittori, prodotto e gestito dalla National Library of Medicine (NLM), utilizzato nell' indicizzazione degli articoli delle riviste biomediche di PubMed/Medline ed altri archivi della NLM.

Il MeSH ha una struttura ad albero e i termini accolti al suo interno sono collegati tra loro attraverso un sistema di relazioni semantiche:

- ***Relazione di equivalenza:*** questa tipologia di relazione rimanda da una voce non accolta ad una che è considerata, all'interno del dominio specifico di appartenenza, quella più diffusa o maggiormente utilizzata, ad esempio:

⁵⁰ Classificazione Internazionale del Funzionamento, della Disabilità e della Salute.

Clinical Markers **see** *Biological Markers*;

➤ **Relazioni Gerarchiche:** questa tipologia di relazione rinvia da un termine gerarchicamente superiore ad uno inferiore e viceversa, ad esempio,

○ *Apparato digerente*

▪ *Vie biliari*

▪ *Fegato*

➤ **Relazioni Associative:** questa tipologia di relazione rimanda da un termine ad un altro correlato, ad esempio:

Veins **see related** *Phlebography*;

Neoplasms **see related** *Antineoplastic agents*;

➤ **Altri tipi di relazione:** la voce *consider also terms* collega termini linguisticamente correlati, aventi radici latine o greche o anglosassoni, questa tipologia di relazione è usata principalmente per i termini anatomici, ad esempio:

Liver **consider also terms** *at hepat-*

L'aggiornamento dei descrittori all'interno del MeSH viene effettuato con cadenza semestrale, sempre dalla NLM, sulla base dell'evoluzione delle scienze biomediche: gli esperti e specialisti di ogni area della biomedicina effettuano un monitoraggio costante della letteratura scientifica e tengono

sotto osservazione le aree emergenti della ricerca. La National Library of Medicine si avvale della collaborazione di un comitato di esperti (*Literature Selection Technical Review Committee*), che si riunisce periodicamente per esaminare le nuove pubblicazioni in ambito scientifico, valutando eventuali nuovi inserimenti all'interno di Medline.

Il thesaurus MeSH è composto da:

- Oltre venticinquemila descrittori principali (*main headings*);
- Ottantatre sottodescrittori (*subheading* o *qualifiers*);
- Oltre centomila voci supplementari (*Supplementary Concept Records*), che comprendono nomi di sostanze chimiche, numeri di registro CAS etc.

Dal 1999 la NLM ha dato la possibilità di consultare liberamente il Thesaurus MeSH nella sua versione originale senza alcun costo di licenza attraverso le seguenti modalità di ricerca: *MeSH Browser*⁵¹ che consente la ricerca all'interno del vocabolario attraverso vari punti d'accesso; *MeSH Database* per la selezione dei descrittori da utilizzare per la ricerca in MEDLINE/PubMED; il *Metathesaurus UMLS*, in cui convergono, oltre che il MeSH, diversi vocabolari del dominio biomedico.

⁵¹ <<http://www.nlm.nih.gov/mesh/MBrowser.html>>

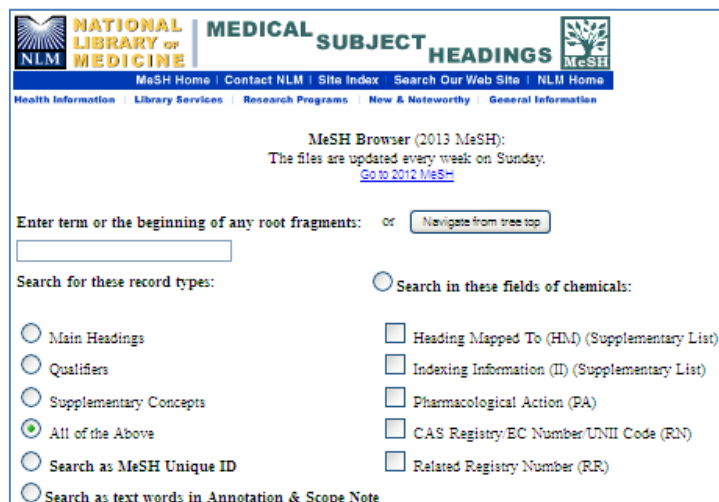


Figura 5 Schermata di visualizzazione del MeSH browser.

All'interno del MeSH Browser (Fig.5) è possibile effettuare la ricerca all'interno del thesaurus attraverso diverse tipologie:

- *MeSH Heading* (Main headings)
- *Supplementary Concept* (sostanze chimiche)
- *Qualifiers* (Subheadings)
- La ricerca può essere effettuata, inoltre, avviando la navigazione dall'inizio della struttura ad albero o inserendo un termine nella sua interezza o semplicemente parte di esso, ad esempio:

- Diseases
 - Bacterial Infections and Mycoses
 - Virus Diseases
 - Parasitic Diseases
 - ...

La traduzione del MeSH nelle diverse lingue nasce nel quadro del progetto UMLS della NLM, avviato nel 1986 per la creazione di un linguaggio

medico accessibile universalmente attraverso i sistemi informativi computerizzati. Il lavoro di traduzione iniziò nel 1998 partendo dalle categorie relative all'anatomia e alle malattie e fu completato tra il 1999 e il 2001, con il conseguente invio della traduzione alla National Library of Medicine per l'approvazione definitiva⁵².

La versione italiana del MeSH è a cura del settore di Documentazione dell'Istituto superiore di Sanità⁵³, sotto la direzione della dott.ssa Maurella Della Seta, che, in collaborazione con la NLM da ormai un trentennio, revisionano ed aggiornano la traduzione con cadenza annuale.

Dal sito dell'ISS è possibile accedere alla pagina di consultazione del MeSH nella sua versione italiana⁵⁴. Dalla stessa pagina è, inoltre, possibile visualizzare ed accedere direttamente al termine nella sua versione originale con un link che rimanda alla pagina ufficiale del MeSH⁵⁵.

2.3.3 La metodologia di lavoro

Nonostante diversi vocabolari, terminologie e sistemi di classificazione siano chiaramente necessari, presentano diverse difficoltà soprattutto nella loro implementazione. È possibile, tuttavia, ovviare a questi problemi attraverso il *mapping*, che significa collegare i contenuti da una terminologia o schema di classificazione ad un altro. Operazione fondamentale del *mapping*, durante la fase di analisi dei dati, è quella di decidere le regole da applicare, ovvero i parametri da utilizzarsi, che influiranno fortemente sui risultati ottenuti.

Il *mapping* considera scopi differenti, livelli di dettagli e linee guida della codifica source (di partenza) e della codifica target (di arrivo). Il

⁵² <http://www.iss.it/binary/sdoc/cont/mesh_Gruppo.pdf>

⁵³ D'ora in poi ISS.

⁵⁴ <<http://www.iss.it/site/mesh/Index.aspx>>

⁵⁵ <<http://www.nlm.nih.gov/mesh/MBrowser.html>>

processo utilizza un metodo standard in cui il concetto della terminologia o la descrizione della classificazione sono interpretate fra sistemi.

L'analisi dei dati ha previsto due macrofasi: una prima fase è stata dedicata a capire quanto le risorse individuate si sovrapponevano e se fossero più o meno omogenee dal punto di vista della selezione del termine principale.

C'è da sottolineare, tuttavia, che nonostante lo scopo finale sia il medesimo per tutte le associazioni da cui sono state acquisite le basi terminologiche, ovvero fornire informazioni dettagliate agli utenti coinvolti nel mondo delle patologie rare, siano essi professionisti del settore o pazienti, la formazione dei rispettivi database e la conseguente pubblicazione sono avvenute in modalità differenti. È per questo motivo, principalmente, che sussistono delle disomogeneità soprattutto a livello quantitativo.

Tuttavia, anche se la base terminologica della NORD è esigua in termini quantitativi si è rivelata importante nei casi in cui non si avevano informazioni sulle relazioni di sinonimia per termini presenti sia in ORDR che in Orphanet, ma in NORD erano presenti, come mostrato in Fig.6:

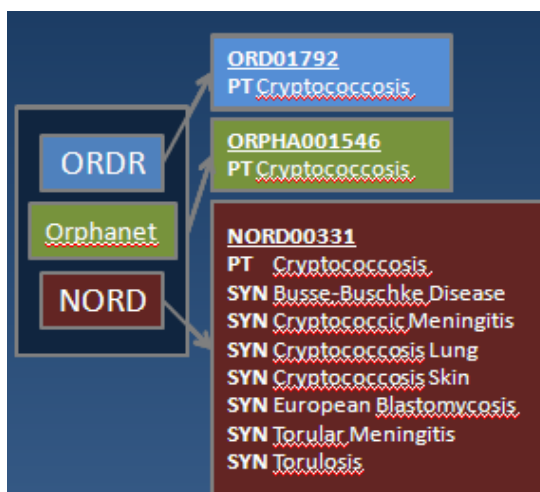
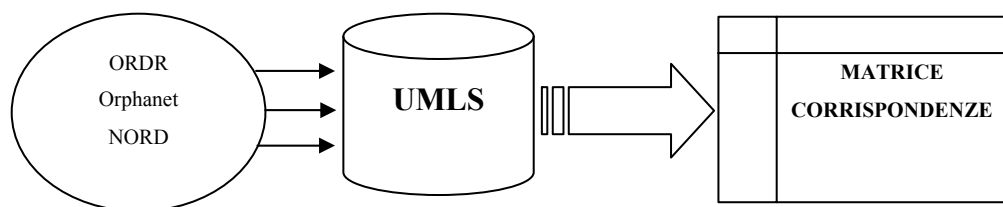


Figura 6 Rappresentazione delle patologie rare nelle risorse di partenza

Una volta creata la base terminologica di tutti i termini provenienti dalle tre risorse fin qui esposte, si è proceduto ad effettuare una mappatura verso i termini presenti all'interno dell'UMLS.

I termini facenti parte del corpus sono stati utilizzati come discriminante all'interno di tutti i vocabolari e classificazioni esistenti all'interno del database dell'UMLS, una volta effettuata la mappatura, è stato possibile ottenere una matrice di corrispondenza che ha evidenziato quali e quanti termini siano presenti in ogni singola risorsa terminologica.



Prima di effettuare la mappatura, tuttavia, è stato necessario applicare degli appositi filtri per evitare che ci fosse troppo rumore informativo. All'interno dell'UMLS vi sono numerosi vocabolari altamente specifici, che non potevano, per la loro intrinseca natura, essere sicuramente utili per lo scopo del presente lavoro.

Al momento della fase di analisi dell'ipotesi di lavoro, sono stati fissati i seguenti obiettivi:

- Sovrapposizione delle terminologie di partenza (per comprendere, in termini quantitativi e qualitativi, quanto le tre risorse di partenza avessero in comune;
- Presenza/non presenza dei termini all'interno dell'UMLS dei termini provenienti dalle risorse di partenza;

- Una volta ritrovato un termine all'interno delle risorse terminologiche, acquisire eventuali termini aggiuntivi che possano alimentare il thesaurus, obiettivo finale del lavoro di ricerca.

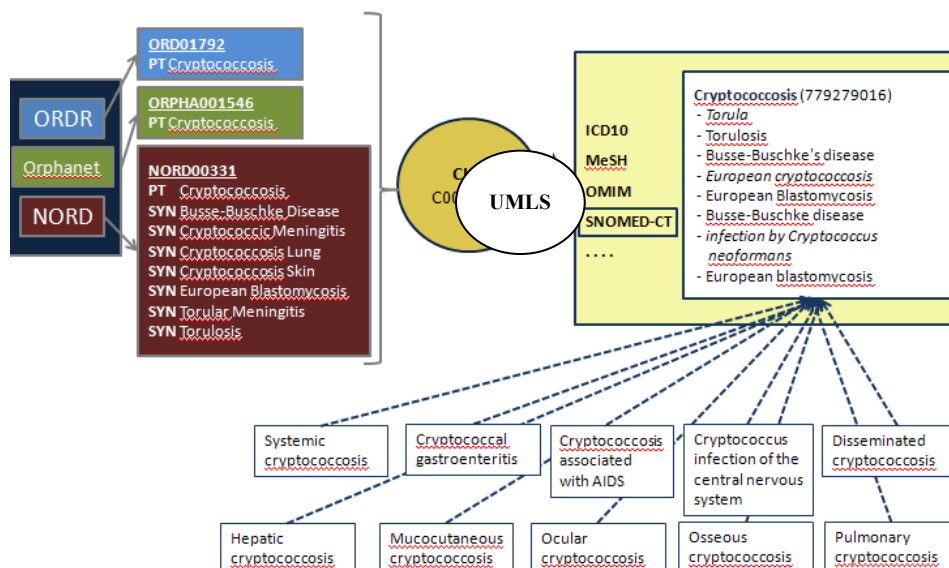


Figura 7 Mappatura delle risorse di partenza attraverso l'UMLS.

La Fig. 7 mostra l'ipotesi di lavoro appena enunciata. Ad esempio, partendo dal termine *Cryptococcosis*⁵⁶, si evince che il termine è presente in tutte e tre le risorse di partenza, ma in varia misura: in Orphanet e in ORDR è presente il termine singolo, mentre in NORD, sono presenti diversi sinonimi.

Una volta individuato il termine è stata effettuata la mappatura verso l'UMLS, rintracciando, quindi, il termine all'interno di diverse risorse terminologiche. Il risultato ottenuto è che il termine *Cryptococcosis* è presente in diverse terminologie, ma è importante notare come SNOMED CT fornisca maggiori informazioni rispetto a quelle di partenza.

⁵⁶ La criptococcosi è un'infezione micotica rara cosmopolita dovuta al *Cryptococcus neoformans*. L'incidenza annuale è stata stimata tra 1/5.000 e 1/50.000, a seconda della regione geografica. Le principali vittime dell'infezione sono i pazienti immunodeficienti.

Questo esempio rispetto al singolo termine sta a sottolineare l'importanza del lavoro svolto, ovvero, come, si possano ottenere delle risorse terminologiche arricchite che portino, quindi, svolgere funzioni ad alto livello di dettaglio.

Poiché i termini di ORDR, Orphanet e NORD si riferiscono a disturbi o patologie (rari), è stato ristretto il campo di ricerca all'interno dell'UMLS (questo perché l'UMLS comprende nomi di farmaci, analisi di laboratorio ecc.) per effettuare la mappatura solo verso uno specifico gruppo di termini. Il filtro semantico applicato fornisce un discreto livello di disambiguazione tra i termini: quando un termine appartenente alle risorse originarie (ORDR e Orphanet), come NF2, viene mappato verso l'UMLS, può trovare corrispondenza in una patologia (neurofibromatosi di tipo 2) oppure in un gene (NF2, cromosoma 22, di cui le mutazioni causano la neurofibromatosi di tipo 2). Applicando il filtro semantico Semantic Disorders (Fig.8), si otterrà come risultato solo il primo, corrispondente alla patologia.

Andersen disease	C0017923	Disorders
Andersen disease	C1415001	Genes & Molecular Sequences

Figura 8 Applicazione del filtro semantico.

Come si evince dalla figura, partendo dallo stesso termine, identifichiamo due diversi CUI, ma solo il primo è stato preso per questo studio, poiché il secondo appartiene al gruppo semantico *Genes & Molecular Sequences*.

Tuttavia, pur applicando tale filtro potrebbe, comunque, persistere un certo grado di ambiguità, quando ad esempio un concetto delle risorse originarie (ORDR, Orphanet e NORD), viene mappato verso più concetti delle risorse presenti nell'UMLS. Il secondo filtro applicato è quello dei termini aggregati a livello concettuale. Il vantaggio, in questo caso, è stato quello di

avere i termini raggruppati per concetti, oltre che nell'UMLS, anche nelle risorse originarie. Le operazioni di mappatura possono definirsi corrette nel caso in cui termini appartenenti allo stesso concetto di malattia rara delle risorse di partenza sono stati mappati verso lo stesso concetto nell'UMLS, ma, al contrario, quando i termini appartenenti allo stesso concetto delle risorse di partenza risultano avere più corrispondenze nei concetti presenti nell'UMLS, ci si trova nel caso in cui la mappatura è ambigua.

Un'attenzione particolare è stata rivolta anche agli acronimi presenti in un numero non trascurabile all'interno delle classificazioni. Per ovviare a problemi di ambiguità semantica, sono state applicate delle regole di disambiguazione assegnando un "peso" ai termini in forma estesa ed un altro, di misura inferiore, agli acronimi.

L'assegnazione dei valori è stata effettuata con delle regole grammaticali grazie alle quali è stato assegnato valore "0,33" alle parole che avessero almeno una o più lettere tutte maiuscole, considerate, quindi, acronimi; mentre alle parole più lunghe, con carattere minuscolo, è stato assegnato valore pieno, ovvero "1".

Id Orphanet	Tipologia termine	CUI UMLS	Peso
ORPH001	T1	CUI1	1
ORPH001	SYN1	CUI2	1
ORPH001	SYN2 [ACRO]	CUI2	0,33
ORPH001	SYN3	CUI1	1

Il codice ORPH001 è l'identificativo assegnato al termine proveniente da Orphanet, nella seconda colonna è rappresentata la tipologia del termine analizzato:

- T1 : termine principale
- SYN 1 ,3 = sinonimi del termine principale

- SYN 2 [ACRO] = sinonimo del termine principale sotto forma di acronimo

Nell'ultima colonna, è stato inserito il peso associato ad ogni termine. Dalla tabella si può notare come, a partire dallo stesso termine, si ha come risultato l'associazione a due diversi CUI all'interno dell'UMLS, e ciò sembra essere abbastanza improbabile.

La formula che segue è stata applicata come regola di disambiguazione per i risultati ottenuti, la prima volta è stata applicata al CUI1 e la seconda volta al CUI2:

$$W_{\text{CUI1}} = \frac{\sum_{w\text{CUI1}}}{\sum_{w\text{CUI1}} + \sum_{w\text{CUI2}}} = \frac{2}{2 + 1,33} = \frac{2}{3,33} = \mathbf{0,60}$$

La formula chiarisce come è stato calcolato il peso (W = Weight) di ogni CUI individuato. In questo caso il risultato è di 0,60, quindi, poiché il valore soglia è stato fissato a 0.6, il peso ottenuto stabilisce che in questo caso il CUI1 identificato non genera ambiguità nei risultati ottenuti.

$$W_{\text{CUI2}} = \frac{\sum_{w\text{CUI2}}}{\sum_{w\text{CUI2}} + \sum_{w\text{CUI1}}} = \frac{1,33}{1,33 + 2} = \frac{1,33}{3,33} = \mathbf{0,39}$$

Nel secondo caso il valore ottenuto è di 0.39, quindi è considerato ambiguo ed è stato analizzato attraverso revisione manuale.

2.3.4 Risultati della mappatura verso l'UMLS

Effettuata la mappatura si sono ottenuti risultati di tre tipologie:

1. Nessun CUI associato: nessun termine è stato individuato all'interno dell'UMLS;
2. Un solo CUI: la corrispondenza è stata rintracciata in modo esatto.
3. Più di un CUI: è stata individuata più di una corrispondenza a partire da un solo termine, questi casi sono stati analizzati manualmente per individuare eventuali problemi nella query di partenza o semplicemente per studiarne la casistica.

La matrice in figura 9 evidenzia, se pur in piccola parte, i risultati della mappatura delle risorse terminologiche di partenza verso l'UMLS: il numero uno sta ad indicare per ogni singolo termine la presenza all'interno della risorsa terminologica, al contrario, la cella vuota ne indica l'assenza.

In verde i termini provenienti da Orphanet, ORDR e NORD, in rosso i termini all'interno delle differenti risorse terminologiche. I termini mappati verso l'UMLS hanno trovato corrispondenza per un totale di 8,435 CUI.

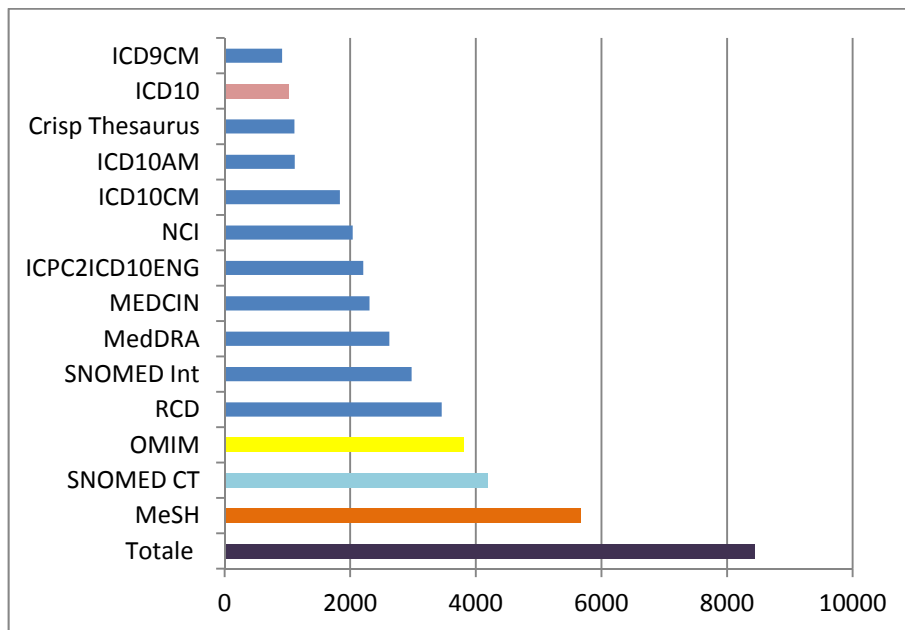


Figura 10 Percentuali di rappresentazione delle risorse terminologiche all'interno dell'UMLS

La risorsa terminologica all'interno dell'UMLS che meglio rappresenta l'ambito delle malattie rare è il MeSH, nella versione del 2011, con il 67%, seguito da SNOMED CT con il 50% e OMIM con il 45%. L'ICD ne contiene solo il 12%.

2.2.3.1 Problemi di ambiguità durante le operazioni di mappatura

Durante le operazioni di mappatura, come accennato, sono stati riscontrati dei problemi di ambiguità nei risultati ottenuti. Possiamo riassumere i risultati in tre categorie distinte:

- **Concetti non ambigui (Fig.11):** La maggior parte dei termini appartenenti ad uno stesso concetto mappano verso lo stesso CUI.

Code Source	Source Preferred Name	CUI	Source Preferred Name
ORU0011/	Acrodysostosis	C0220659	Acrodysostosis
ORPHA001248	Maxillo-rasal dysplasia	C0220692	MAXILLONASAL DYSPLASIA, BINDER TYPE
NORD00312	Corn Syndrome	C1384514	Corn Syndrome

Figura 11 Esempio di mappatura non ambigua.

- Mappatura ambigua (Fig.12): i termini di un dato concetto mappano verso più di un CUI all'interno dell'UMLS.

	CUI1 C0268128	CUI2 C0220987	CUI3 C0268131
Orotic aciduria	Orotic aciduria		
Orotic aciduria hereditary		Hereditary orotic aciduria	
Orotidylic decarboxylase deficiency			Hereditary orotic aciduria, type 2
Uridine monophosphate synthetase deficiency	---	---	---

Figura 12 Esempio di mappatura ambigua

In questo caso la mappatura risulta ambigua a causa del diverso scopo dei due sistemi di classificazione. Orphanet raggruppa dei termini che considera appartenenti alla stessa categoria semantica, mentre SNOMED CT che è una classificazione di ambito clinico altamente specializzata con un livello di dettaglio molto elevato, mette gli stessi termini in gerarchia per evidenziare che uno è subordinato all'altro.

- Concetti non mappati (Fig.13): Nessuna corrispondenza individuata all'interno dell'UMLS.

Source Code	Full Name	Possible Explanation
ORPHA002096 ORDR05685	Richieri Costa-Guion Almeida-Cohen syndrome	Extremely rare disease
ORPHA002369	Lateral body wall complex	Extremely rare disease
ORD03770 NORD00694	Levy-Yeboa Syndrome	Recently recognized disorder

Figura 13 Esempio di termini non mappati.

Nella figura sopra sono riportati degli esempi di casi in cui non è stata individuata nessuna corrispondenza all'interno dell'UMLS.

Per quanto riguarda i concetti all'interno della NORD solo il 2% non hanno trovato alcuna corrispondenza, il 22% per ciò che concerne l'ORDR e il 47% per Orphanet.

I casi appena descritti sono stati revisionati manualmente per individuare le cause della mancata mappatura, ed è stato riscontrato, infatti, che le patologie che non hanno trovato corrispondenza all'interno dell'UMLS sono state identificate come eccessivamente rare e, quindi, non inserite in nessuna classificazione di tipo specialistico, come nel caso della *Richieri Costa-Guion Almeida-Cohem syndromee* la *Lateral body wall complex*.

Nel caso della *Levy-Yeboa Syndrome* non è stata rintracciata la corrispondenza all'interno dell'UMLS poiché è una sindrome di recente scoperta e quindi non presente all'interno del database dell'UMLS in base agli ultimi aggiornamenti.

Capitolo 3

3.1 Avvio della costruzione del thesaurus: il Registro Nazionale delle Malattie Rare

La desiderata interoperabilità ed i relativi risultati (che molto spesso si traducono, o si dovrebbero tradurre, in riduzione dei costi, aumento dell'efficienza, migliore qualità nello scambio dei dati) non si possono raggiungere senza usare vocabolari specialistici, terminologie e sistemi di classificazione, che lavorino insieme in ambiente digitale.

L'analisi dello stato dell'arte svolta in questi anni ha portato dei significativi risultati nell'ambito delle malattie rare, specialmente nel contesto italiano.

Rispetto alla peculiarità dei problemi connessi a tali patologie si prevede un'azione di coordinamento delle diverse iniziative in atto promuovendo quelle volte a garantire: diagnosi appropriata e tempestiva; facilitazioni dell'accesso a centri specialistici per programmi terapeutici; attività di prevenzione; sostegno alla ricerca scientifica, con particolare riferimento allo sviluppo di nuove terapie. Attualmente nel nostro paese non c'è una legislazione organica per le malattie rare e soprattutto una specifica per i farmaci orfani.

L'industria produttrice di farmaci in Italia dedica in media 8-12% del fatturato alle spese di ricerca complessiva. Considerando che le spese di ricerca e sviluppo per le malattie rare ammontano a circa cinquanta milioni di dollari spesi in dieci anni, settecento milioni di dollari investiti complessivamente dalle ditte farmaceutiche italiane per la ricerca nel 1987 sarebbero stati teoricamente sufficienti ad assicurare lo sviluppo di 14 farmaci con caratteristiche fortemente innovative nello stesso anno. Nel complesso delle aziende farmaceutiche italiane esiste un notevole patrimonio culturale e tecnologico, eppure nessuno potrebbe sostenere che negli ultimi

dieci anni siano stati messi a punto in Italia 140 farmaci con caratteristiche fortemente innovative, ma neanche centoundici o quarantasette.

In verità non è più pensabile che si possa risolvere il problema a livello nazionale, ma occorre pensare ad una legge unica per la C.E. Nel 1993 a Londra fu fondato EMEA (agenzia europea per la valutazione dei prodotti medici). Nel 1995 entrò in vigore un nuovo sistema europeo per l'autorizzazione della registrazione dei medicinali. La necessità di una legge europea riguarda non solo il campo dello sviluppo e della ricerca, ma anche quello dell'informazione e della distribuzione dei costi. Un farmaco costa oggi 250-500 mila dollari e in media dalla scoperta di un principio attivo alla sua commercializzazione possono passare da 10 a 14 anni.

Quando viene messo in commercio, l'obiettivo è quello di recuperare le spese e quindi rendere molto e per molto tempo, ma per definizione gli acquirenti saranno pochi (farmaci orfani per patologie orfane o rare e/o per popolazioni con limitate risorse economiche). Molti ricercatori si occupano poco volentieri di malattie rare perché ci sono scarse probabilità di finanziamento e le cure rimangono allo stato embrionario. Nonostante siano rare, sono almeno 5.000 e costituiscono quindi un grosso peso per la Comunità Europea e un grosso gruppo di consumatori delle risorse della salute; ecco perché il loro trattamento può essere un buon investimento per la collettività.

Nell'ambito delle attività previste dalla "Ridefinizione del sistema di partecipazione al costo delle prestazioni sanitarie e del regime delle esenzioni" operata con il Decreto legislativo n.124 del 29 aprile 1998, il Consiglio Superiore di Sanità ha approvato uno schema di regolamento relativo alla "Istituzione della rete nazionale delle malattie rare e di esenzione dalla partecipazione al costo delle relative prestazioni sanitarie".

Con il Decreto Legislativo del 18 maggio 2001, n. 279, il Ministero della Sanità adotta il "Regolamento di istituzione della rete nazionale delle malattie rare e di esenzione alla partecipazione al costo delle relative

prestazioni sanitarie, ai sensi dell'articolo 5, comma 1, lettera b) del decreto legislativo 29 aprile 1998, n. 124":

Art. 5. (esenzione della partecipazione in relazione a particolari condizioni di malattia). – 1. Con distinti Regolamenti del Ministero della sanità da emanarsi ai sensi dell'art. 17, comma 3, della legge 23 agosto 1988, n.400, sono individuate rispettivamente: a) le condizioni di malattie croniche o invalidanti; b) le malattie rare. le condizioni e malattie di cui alle lettere a) e b) danno diritto all'esenzione dalla partecipazione per le prestazioni di assistenza sanitaria indicate da medesimi regolamenti. Nell'individuare le condizioni di malattia, il ministro della sanità tiene conto della gravità clinica, del grado di invalidità, nonché della onerosità della quota di partecipazione derivante dal costo del relativo trattamento.

Il suddetto Regolamento è teso a:

- disciplinare "le modalità di esenzione dalla partecipazione al costo delle malattie rare per le correlate prestazioni di assistenza sanitaria incluse nei livelli essenziali di assistenza", considerando ciascuna condizione morbosa e relative complicanze, le prestazioni di elevata frequenza, ed escludendo dall'esenzione le prestazioni caratterizzanti il percorso diagnostico di patologie esenti ad eccezione delle malattie rare (*art. 1*);
- individuare "specifiche forme di tutela per i soggetti affetti", sottolineando la necessità di garantire la disponibilità dei farmaci orfani e l'erogazione delle prestazioni assistenziali (*art. 1*).

Per raggiungere gli scopi sopraindicati il Regolamento prevede l'istituzione di una Rete Nazionale per la prevenzione, la sorveglianza, la diagnosi e la terapia delle malattie rare (*art. 2*): tale Rete sarà costituita da presidi accreditati (preferibilmente ospedalieri) segnalati dalle regioni, tra

cui il Ministro della sanità individuerà i Centri Interregionali di riferimento per le malattie rare. Sorveglianza e monitoraggio delle malattie rare sono compiti dei Centri interregionali di riferimento (*art. 2*), le informazioni e la documentazione da essi raccolta verranno centralizzate nel Registro Nazionale delle Malattie Rare istituito presso l'Istituto Superiore di Sanità in attuazione dell'articolo 3 del D.M 279/2001, al fine di consentire una programmazione nazionale e regionale degli interventi, e regolamentando lo scambio ed il trasferimento di dati sensibili nel pieno rispetto delle leggi sulla privacy (*art. 5*).

Ai fini del riconoscimento del diritto all'esenzione il regolamento individua 284 codici di esenzione riferibili a malattie e/o gruppi di malattie, afferenti a 47 categorie di malattie (*art. 4*, allegato 1): per i soggetti riconosciuti affetti si prevede l'esenzione, presso i presidi della rete, per prestazioni, prescrivibili secondo criteri di efficacia ed appropriatezza (*art. 7*), riguardanti accertamento, trattamento ivi compresi i farmaci specifici, monitoraggio e prevenzione delle complicanze, per le malattie rare individuate dal regolamento stesso (*art. 6*).

Il Registro ha come obiettivi generali di effettuare la sorveglianza delle malattie rare e di supportare la programmazione nazionale e regionale degli interventi per i soggetti affetti da malattie rare. Il registro mira, infatti, ad ottenere informazioni epidemiologiche (in primo luogo il numero dei casi di una determinata malattia rara e relativa distribuzione sul territorio nazionale) utili a definire le dimensioni del problema; si tratta, inoltre, di uno strumento utile per stimare il ritardo diagnostico e la migrazione sanitaria dei pazienti, supportare la ricerca clinica e promuovere il confronto tra operatori sanitari per la definizione di criteri diagnostici.

L'attività del RNMR è iniziata nel 2001 e, per aumentare la copertura e l'efficienza della raccolta dei dati epidemiologici, il CNMR, a partire dall'inizio del 2007, ha messo in atto una nuova modalità di raccolta dati che include un nuovo software: uno strumento che può essere utilizzato sia dai singoli

presidi/centri abilitati alla diagnosi e al trattamento dei pazienti affetti da malattie rare, sia sui Responsabili dei Centri di Coordinamento Regionale che coordinano le attività e fanno da tramite tra il CNMR e i singoli presidi/centri.

Il software è sviluppato su piattaforma web, di semplice utilizzo, realizzato rispettando gli standard di sicurezza e di riservatezza per il trattamento dei dati sensibili. Il CNMR mette, inoltre, a disposizione il software gratuitamente sia alle Regioni che non hanno ancora attivato un proprio Registro Nazionale, sia quelle che ne sono già in possesso.

Con tutte le Regioni è stato condiviso e concordato all'interno dell'Accordo Stato-Regioni del 10 maggio 2007 un elenco di variabili obbligatorie (un data set minimo) da inviare al Registro Nazionale Malattie Rare.

Il data set minimo prevede campi obbligatori sia per la parte anagrafica di arruolamento del paziente, sia per la parte relativa alla patologia; nella sezione facoltativa del Registro Nazionale Malattie Rare è possibile specificare i criteri diagnostici, gli esami clinici di laboratorio e strumentali alla diagnosi e al follow-up ed è possibile segnalare il decesso del paziente⁵⁷.

Rispetto alla peculiarità dei problemi connessi a tali patologie si prevede un'azione di coordinamento delle diverse iniziative in atto promuovendo quelle volte a garantire: diagnosi appropriata e tempestiva; facilitazioni dell'accesso a centri specialistici per programmi terapeutici; attività di prevenzione e sostegno alla ricerca scientifica con particolare riferimento allo sviluppo di nuove terapie.

Un farmaco, ad oggi, costa 250-500 mila dollari e in media dalla scoperta di un principio attivo alla sua commercializzazione possono passare da dieci a quattordici anni. Quando viene messo in commercio, l'obiettivo è quello recuperare le spese e quindi rendere molto e per molto tempo, ma per

⁵⁷ Cfr. <www.iss.it/cnmr/regi/cont.php?id=860&lang1&tipo=14>

definizione gli acquirenti saranno pochi (farmaci orfani per patologie rare/ o per popolazioni con limitate risorse economiche). Molti ricercatori si occupano poco volentieri di malattie rare perché ci sono scarse probabilità di finanziamento e le cure rimangono allo stato embrionario. Nonostante siano rare, sono almeno cinquemila e costituiscono quindi un grosso peso per la Comunità Europea e un grosso gruppo di consumatori delle risorse della salute; ecco perché il loro trattamento può essere un buon investimento per la collettività.

3.2 Dal Registro al thesaurus

La scelta di riorganizzare la conoscenza in questo ambito specifico è ricaduta sul thesaurus in ragione dei risultati che lo stesso permette di ottenere in termini di controllo terminologico, recupero e strutturazione dell'informazione sia per un utilizzo da parte degli esperti di dominio, che da parte del cittadino/paziente. Per tali motivi, esso risponde agli obiettivi propri dell'attività in oggetto, che prevedono la sua per un corretto recupero di informazione che sia univoca e non ambigua.

Il concetto di thesaurus e le operazioni necessarie alla sua realizzazione sono oggetto di una normativa tecnica aggiornata molto recentemente: la ISO 25964-1:2011, *Information and documentation – Thesauri and interoperability with other vocabularies, Part 1: Thesauri for information retrieval*, sostituisce infatti le precedenti norme in materia, ormai datate, ovvero la ISO 2788:1986, *Documentation – Guidelines for the establishment and development of monolingual thesauri*, la ISO 5964:1985, *Documentation – Guidelines for the establishment and development of multilingual thesauri*.

Il panorama normativo è caratterizzato anche dalla presenza degli standard ANSI/NISO Z39-19 2005, *Guidelines for the construction, format, and management of monolingual controlled vocabularies*, che si interessa anche ad

altri tipi di vocabolari controllati e si apre a problematiche di interoperabilità tra i vocabolari stessi attraverso la definizione di un modello basato su XML (eXtensible Markup Language)⁵⁸ e detto Zthes⁵⁹, e dalla BS 8723, *Structured vocabularies for information retrieval – Guide*.

Nella definizione del concetto di thesaurus è possibile comprendere come lo strumento si sia evoluto nel corso del tempo per rispondere alle esigenze di recupero e gestione dell'informazione che evolvono in fretta nel tempo, specialmente nell'ultimo decennio, sicuramente a seguito dell'avvento del web e dell'aumento della quantità di informazione disponibile in ambienti prevalentemente digitali, che hanno indubbiamente contribuito alla valorizzazione delle potenzialità del thesaurus stesso.

Secondo la ISO 2788:1985:

il thesaurus è un vocabolario di un linguaggio di indicizzazione controllato, organizzato formalmente in maniera da rendere esplicite le relazioni “a priori”⁶⁰ fra i concetti

mentre la definizione fornita dalla ISO 25964:2011 è la seguente:

controlled and structured vocabulary in which concepts are represented by terms, organized so that relationships between concepts are made explicit, and preferred terms are accompanied by lead-in entries for synonyms or quasi-synonyms

⁵⁸ <<http://www.w3.org/XML/>>

⁵⁹ <<http://zthes.z3950.org/>>

⁶⁰ Le relazioni paradigmatiche o “a priori” tra i concetti sono così definite dalla ISO 25964:2011: “*Relationship between concepts which is inherent in the concepts themselves*”. Si tratta cioè di relazioni che, contrariamente a quelle sintagmatiche, sono sempre valide indipendentemente dai contesti specifici di indicizzazione o di definizione del thesaurus. Le relazioni sintagmatiche sono perciò sconsigliate all'interno dei thesauri.

È evidente come gli sviluppatori di thesauri si siano dovuti, necessariamente adeguare ai nuovi “ambienti” su cui costruire le basi terminologiche, con l’evolversi del tempo. Pratiche quotidiane che imponevano l’utilizzo del cartaceo sono ormai largamente superate con l’avvento del digitale che ha permesso di migliorare notevolmente le ricerche in termini di efficacia e tempo impiegato, laddove per avere accesso ad un termine basta solo digitare la sequenze dei caratteri effettuando la ricerca per parola chiave o semplicemente scorrendo l’albero navigando liberamente in tutte le direzioni, gerarchiche e non, semplicemente con un tocco di mouse.

Per tali ragioni è stato necessario dover adeguare le norme vigenti al nuovo ambiente digitale creatosi: la definizione fornita dalla ISO 25964:2011 introduce la differenza tra termini preferiti e non preferiti e il fatto che i sinonimi e i quasi sinonimi siano legati tra di loro per mezzo di rinvii, non riferendosi esclusivamente alla pratica dell’indicizzazione, ma anche al recupero dell’informazione, che può avvenire a partire da tutti i termini appartenenti al vocabolario d’accesso⁶¹. In questo senso, obiettivo del thesaurus è quello di far sì che sia il professionista dell’informazione, che l’utente utilizzino lo stesso termine preferito per individuare un dato concetto o che l’utente arrivi al recupero dell’informazione pur utilizzando una chiave di ricerca diversa da quella preferita.

La ISO 25964:2011 introduce delle modifiche e delle novità legate al nuovo ruolo che si riconosce al thesaurus in ambiente digitale. Rispetto, ad esempio, alla norma ISO 2788:1896, focalizzata principalmente sul controllo terminologico, e quindi sulla scelta della forma dei termini da inserire nel thesaurus, sulle tipologie di relazioni che possono essere stabilite tra un termine e l’altro e sulla presentazione delle due modalità di visualizzazione del thesaurus stesso, alfabetica e sistematica, la nuova norma prevede: regole

⁶¹ “Costituito sia dai termini preferiti che dai termini non preferiti, cioè dai termini che non possono essere utilizzati per l’indicizzazione e che rimandano a termini preferiti”
SERAFINA SPINELLI, *Introduzione all’indicizzazione*, 2006.
< <http://biocfarm.unibo.it/~spinelli/indicizzazione/>>.

e raccomandazioni riguardo ai software per la costruzione e la gestione di un thesaurus; modelli di dati per il thesaurus monolingue e per quello multilingue; attenzione rivolta verso l'interoperabilità tra sistemi basati sull'uso di thesauri, sistemi di indicizzazione e di IR, con le conseguenti raccomandazioni relative ai formati e ai protocolli di scambio delle informazioni e ai modelli per le operazioni di importazione ed esportazione dei dati; maggiore rilevanza attribuita alla ricerca di informazione con scelte più accurate anche nell'inserimento dei termini composti e nella scelta dei quasi sinonimi; introduzione della differenza tra concetto e termine, quest'ultimo inteso come forma linguistica per rappresentare il primo; possibilità di specificare la generica relazione RT precisando se si tratta di causa/effetto, processi/prodotti, ecc.

Nella norma stessa, infatti, si precisa che, oltre ad essere strumenti elettronici, realizzati con l'ausilio di software dedicati, i thesauri vengono sempre più spesso integrati con altre applicazioni web, quali motori di ricerca, centri di documentazione e sistemi di *content management* in generale.

3.2.1 Il thesaurus e le funzioni d'uso

Il thesaurus, fin dalla sua creazione, ha da sempre assolto delle funzionalità di fondamentale importanza che grazie alle nuove tecnologie sono state aumentate di valore e ha consentito, di conseguenza, di poter allargare la sua intrinseca funzione ai diversi contesti d'uso nella gestione dell'informazione in ambiente digitale. Una risorsa che raccolga e strutturi i concetti rappresentativi di un dominio, più o meno specialistico, attraverso relazioni di equivalenza, gerarchiche e associative definite sulle proprietà intrinseche dei concetti stessi, comporta un valore aggiunto per qualsiasi contesto che richieda descrizione, rappresentazione e recupero di informazione e documenti.

Nello specifico, le principali funzionalità, strettamente interrelate, del thesaurus possono essere così sintetizzate: controllo terminologico, indicizzazione, supporto nella definizione dei metadati e classificazione sono attività che competono al professionista dell'informazione, mentre navigazione, ricerca ed espansione dei risultati delle ricerche coinvolgono direttamente l'utente come utilizzatore di tale strumento.

3.1.1.1 Controllo terminologico

La funzione di controllo viene esercitata prevalentemente per i fenomeni di sinonimia e di polisemia. All'interno di un thesaurus il significato di un termine deve essere assolutamente non ambiguo e la definizione della rete semantica tra i concetti permette di specificare quanto più possibile il significato che si vuole privilegiare in uno specifico contesto.

In tal senso è importante sottolineare anche l'utilizzo di note d'ambito, che consentono di inserire delle definizioni o delle specificazioni ulteriori circa un dato termine, e di qualificatori, che permettono di specificare l'ambito o il sottoambito al quale un termine appartiene, contribuisce alla determinazione univoca di un concetto.

Allo stesso modo, le relazioni di equivalenza gestiscono la presenza di sinonimi, varianti lessicali, quasi - sinonimi, attraverso l'elezione, sulla base di diversi criteri, di un termine preferito al quale gli altri restano legati in quanto punti di accesso all'informazione.

Tale legame permette di essere indirizzati verso il termine preferito e di arrivare comunque all'informazione ricercata, anche se indicizzata tramite i descrittori.

Il concetto di controllo si traduce soprattutto in un requisito fondamentale per garantire l'incontro fra lessico dell'indicizzatore e lessico del ricercatore, e cioè la relazione biunivoca fra termine e concetto, fra significante e significato: ciò significa che in un thesaurus un termine esprime

sempre uno ed un solo concetto, e che un concetto è sempre espresso da uno ed un solo termine.

Poiché questa condizione è tutt'altro che vera nel linguaggio naturale, affetto, come ovvio, da ridondanze, ambiguità, polisemie, omonimie, omografie ed altre disfunzioni che ne garantiscono ricchezza ed espressività, il raggiungimento del controllo, cioè dell'univocità semantica, nella costruzione e manutenzione del thesaurus comporta due tipi di accorgimenti:

1. La raccolta e la collazione dei sinonimi, dei quasi sinonimi (cioè termini non sinonimi in senso proprio, ma considerabili tali ai fini della rappresentazione dei concetti dell'ambito coperto dal thesaurus) e degli antonimi (cioè degli opposti, o meglio dei termini collocati in diversi punti dello stesso continuum semantico: non solo *calore* e *freddo*, ma anche *epore*, *fresco* e così via) atti a descrivere il medesimo concetto e la scelta di uno solo di questi termini: il termine prescelto diventa nel vocabolario termine preferito, TP (o PT, preferred term), descrittore, cioè termine abilitato a descrivere quel determinato concetto; tutti gli altri termini (termini non preferiti, NPT, non descrittori), che non possono essere assegnati ai documenti per esprimerne il contenuto concettuale, possono però vantaggiosamente entrare nel vocabolario controllato come punti di accesso che rinviano al termine preferito;
2. La riduzione del contenuto semantico del termine preferito ad un solo significato, di solito il più tipico nell'ambito disciplinare del thesaurus, per cui ad esempio in un thesaurus di ornitologia il termine *gru* esprimerà il concetto corrispondente all'"uccello dei Ralliformi" e non alla "macchina dal braccio girevole per sollevare pesi", né al "carrello mobile per riprese cinematografiche".

3.1.1.2 Indicizzazione

L'indicizzazione è la funzionalità principale riconosciuta ad un thesaurus. Tale strumento, infatti, struttura un set di termini che possono essere selezionati ed attribuiti ad un documento in quanto rappresentativi del suo contenuto concettuale.

Le relazioni tra i concetti consentono al professionista dell'informazione, in un contesto di indicizzazione manuale, di scegliere i termini in base al livello di specificità che più si adatta alle caratteristiche del corpus documentale, sia questo in forma cartacea o digitale. I descrittori associati ai documenti ne permettono poi il recupero a partire da differenti modalità di accesso all'informazione.

3.1.1.3 Navigazione e ricerca

Nei sistemi di *content management* che prevedono l'integrazione di un thesaurus, la presentazione sistematica che mette in evidenza la struttura classificatoria definita, può rappresentare il punto di partenza per operazioni di *browsing*.

Anche la presentazione alfabetica può tuttavia essere navigata per orientare la ricerca di informazione da parte dell'utente, soprattutto in contesti in cui il suo bisogno informativo non è perfettamente definito. La visualizzazione delle relazioni tra i termini fornisce un indubbio supporto anche per la migliore conoscenza del dominio, soprattutto per utenti con poche competenze specialistiche. La presenza di relazioni di equivalenza permette principalmente di far fronte alla disomogeneità di competenze e di conoscenze da parte degli utenti, ma anche alla varietà della lingua, che comporta, anche in contesti specialistici, la presenza massiccia di sinonimi e varianti, al fine di recuperare l'informazione indipendentemente dal termine utilizzato per la ricerca o di guidare l'utente verso l'utilizzo del termine scelto come preferito.

3.1.2.1 Le relazioni del thesaurus

La metodologia usuale per navigare il formato alfabetico del thesaurus è attraverso l'impiego di un insieme standard di relazioni o tipi di rimandi.

Le relazioni utilizzate sono⁶²:

USE	<i>Use</i>	Usa	Rinvio da un termine non preferito ad uno preferito
UF	<i>Used for</i>	Usato per	Richiamo da un termine preferito ad uno non preferito
BT	<i>Broader Term</i>	Termine più generale	Rimando a termini dal significato più generale
NT	<i>Narrower Term</i>	Termine più specifico	Rimando a termini dal significato più specifico
RT	<i>Related Term</i>	Termine correlato	Rimando ad un termine che è collegato al primo in modo diverso da BT e NT

L'ordine delle relazioni thesaurali è sempre il medesimo, il primo gruppo di relazioni (USE e UF) è legato al controllo del vocabolario, mentre le

⁶² VANDA BROUGHTON, *Costruire Thesauri: strumenti per indicizzazione e metadati semantici*, (a cura di) P. Cavaleri, (traduzione di) L. Ballestra e L. Venuti, Milano, Editrice Bibliografica, 2008, p.88.

altre (BT,NT, e qualche volta RT) si riferiscono alle relazioni gerarchiche nel thesaurus⁶³.

Le note d'uso (o note d'ambito) sono utilizzate solo nel caso in cui il termine descritto possa avere più significati risultando, di fatto, ambiguo. Molto spesso si tratta di una breve definizione del termine, ma vi sono casi in cui un esempio è necessario per chiarire al meglio il significato e l'inquadramento generale di un termine.

3.2.2 Lo strumento utilizzato: il software Multites Pro 2005

Per quanto riguarda il software di supporto alla costruzione del thesaurus, la scelta è ricaduta su Multites Pro2005, che risponde ad un insieme di requisiti fondamentali per garantire una corretta gestione di tale risorsa.

Tra le caratteristiche più rilevanti del software si possono annoverare:

- Disponibilità delle relazioni standard previste dalla normativa in materia di vocabolari controllati (BT, NT, RT, ...) e possibilità di aggiungerne delle altre, definendone la natura: nel caso specifico, ad esempio, la relazione gerarchica è stata ulteriormente specializzata attraverso l'aggiunta di BTG, BTP e delle relazioni inverse;
- Possibilità di definire delle categorie all'interno delle quali collocare i termini inseriti nel thesaurus: le faccette (al livello più generico) corrispondono alle categorie, per ciascuna delle quali è possibile visualizzare i termini che le appartengono;
- Possibilità di ricorrere alla poligerarchia, e quindi di attribuire ad uno stesso termine più di un concetto sovraordinato;

⁶³ *Ivi*, p.89.

- Inserimento automatico delle relazioni inverse e controllo della coerenza delle relazioni inserite (es. non è possibile associare delle relazioni gerarchiche o associative ai termini non preferiti, ad esclusione della relazione *Subject Category* – SC che indica la faccetta di appartenenza);
- Nessun limite sulla quantità di termini, etichette di snodo, livelli gerarchici e relazioni di vario tipo che possono essere inserite;
- Possibilità di definire e di visualizzare alcune caratteristiche per ciascun termine inserito, tra cui: lo status, per cui un termine può essere *accepted, candidate, provisional* o *not valid*; il tipo, che permette di stabilire se un termine è un descrittore o un'etichetta di snodo; in quante e che tipo di relazioni è coinvolto il termine; la presenza di eventuali note d'ambito e infine la categoria alla quale il termine appartiene. Per ciascuno di essi, poi, è possibile visualizzare schematicamente le relazioni che lo interessano e navigare a partire dai termini che fanno parte della rete semantica definita attorno a ciascun concetto.
- Possibilità di visualizzare il thesaurus tramite diverse modalità di presentazione:
 - Alfabetica;
 - Gerarchica;
 - Sistemica.

Capitolo 4

4.1 Conclusioni

L'ambito delle malattie rare è un ramo che, per definizione, risulta essere di "nicchia" rispetto alla totalità delle scienze mediche. Il lavoro sin qui svolto vuole essere un punto di partenza per continuare ad arricchire lo strumento di accesso all'informazione creato, partendo dalla concezione che seppur "raro" questo dominio non va per nulla tralasciato o traslato in secondo piano.

Una buona rappresentazione è stata riscontrata invece all'interno delle terminologie analizzate e inserite all'interno dell'UMLS, in particolar modo in quelle su cui sono stati fatti ulteriori approfondimenti (SNOMED-CT, MeSH, OMIM), il MeSH più degli altri con il 67% dei termini individuati.

L'obiettivo iniziale che ci si era posti è stato parzialmente raggiunto perché molte sono state le informazioni aggiuntive ottenute attraverso le operazioni di *mapping* effettuate, a partire dal corpus creato estraendo i termini dal database di Orphanet, ORDR e NORD e creando così la base di partenza per la costruzione del thesaurus.

La mappatura ha evidenziato come oltre ad ottenere le informazioni desiderate, il *mapping* è stato anche di fondamentale importanza per l'invio di *feedback* ai gestori dell'UMLS: alcune relazioni, infatti, grazie al supporto delle terminologie specifiche di partenza, sono state modificate in base alle informazioni di granularità fornite. Inoltre, il lavoro di analisi svolto grazie alla preziosa collaborazione della NLM, di Orphanet e dell'ORDR, ha contribuito a sancire un gruppo di lavoro fondamentale per l'inizio e per il prosieguo del lavoro di interoperabilità tra diverse Istituzioni di diversi Paesi, ma che in comune hanno lo stesso obiettivo e, nel caso specifico, quale quello delle malattie rare, di raggiungere un alto livello di comunicazione

verso i professionisti del settore e tanto più verso gli utilizzatori finali, ovvero i pazienti/cittadini.

Il lavoro svolto ha permesso di riconfermare, comunque, le potenzialità di uno strumento di organizzazione della conoscenza quale il thesaurus il cui scopo e i cui principi sono stati ampiamente rivalutati con l'avvento del web e l'esigenza di creare sistemi che strutturino e organizzano le informazioni altrimenti difficilmente recuperabili.

È doveroso sottolineare, tuttavia, come la costruzione di un sistema di organizzazione della conoscenza, nel caso specifico di un thesaurus, richieda, un margine di soggettività che si manifesta nelle scelte compiute per la selezione dei termini, per la loro disposizione al suo interno e per le modalità di applicazione dei principi propri alla metodologia e via dicendo.

Il modello creato, sulla base della ridefinizione del Registro Nazionale delle Malattie Rare, vuole essere un primo passo verso quello che potrebbe diventare uno strumento di pieno utilizzo a livello nazionale.

4.2 Limiti del lavoro

I problemi riscontrati durante l'analisi terminologica sono stati, in particolare, durante la prima fase, sui termini per i quali, durante la mappatura, non è stata rintracciata alcuna corrispondenza.

L'ambiguità evidenziata in alcuni risultati può anche trovare giustificazione nei seguenti casi: la patologia presa in questione potrebbe essere estremamente rara e, quindi, non essere stata introdotta nelle risorse terminologiche analizzate; le patologie recentemente scoperte sono state introdotte nelle liste terminologiche specialistiche dell'ambito specifico delle malattie rare, poiché aggiornate di continuo, ma non nelle risorse terminologiche analizzate che comprendono in generale tutto il dominio biomedico (ICD, SNOMED, OMIM, MeSH), che hanno, nella maggior parte dei casi aggiornamenti semestrali o annuali. Un'altra ragione può dipendere

anche dalla definizione della malattia rara in sé. Come è stato già detto, la definizione di “rara” può dipendere da fattori geografici, alcune patologie possono essere considerate rare in alcune aree e frequenti in altre.

Le risorse utilizzate (Orphanet, ORDR e NORD) non forniscono, nei loro database terminologici, un allineamento verso tutti i concetti presenti nell’UMLS. Valutando la sovrapposizione tra le fonti terminologiche utilizzando l’UMLS come matrice di sovrapposizione, tra i 6841 concetti provenienti dall’ORDR, solo il 39% è stato mappato verso quelli di Orphanet. La percentuale dei concetti provenienti dal database di Orphanet verso quelli dell’ORDR è del 31%.

I risultati ambigui ottenuti possono essere parzialmente spiegati se si pensa al differente orientamento che i termini possono avere all’interno di un sistema di classificazione. Ad esempio, il concetto di Orphanet con identificativo 30 "*Oroticaciduria*" viene mappato verso tre concetti nell’UMLS: C0268128, C0220987 e C0268131. Nel database di Orphanet il concetto *Oroticaciduria* risulta avere quattro relazioni di sinonimia con altri termini, mentre nell’UMLS, i diversi termini che fanno parte sempre dello stesso concetto, sono isolati e, quindi, non aggregati, come mostrato in figura 14:

Orphanet 30	CUI 1 C0268128	CUI 2 C0220987	CUI 3 C0268131
Oroticaciduria	Orotic aciduria	/	/
Orotic aciduria hereditary	/	Hereditary orotic aciduria	/
Orotidylic decarboxylase deficiency	/	/	Hereditary orotic aciduria, type 2
Uridine monophosphate synthetase deficiency	/	/	/

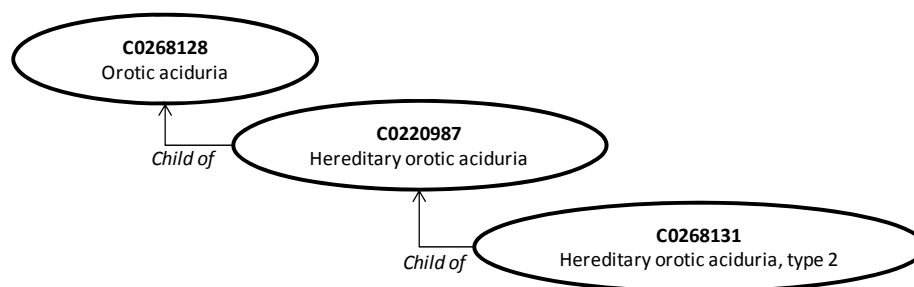


Figura 14 Rappresentazione dei termini non mappati.

L'ambiguità è stata riscontrata poiché i tre termini per cui viene rintracciata la corrispondenza nell'UMLS provengono da SNOMED CT, che essendo la classificazione più dettagliata esistente al mondo in ambito clinico, per il proprio livello di granularità, organizza e classifica, quindi, i tre termini secondo un ordine gerarchico padre/figlio a seconda della tipologia della malattia.

4.3 Un modello funzionale di thesaurus

Il thesaurus creato nel presente lavoro di ricerca vuole essere un modello di come l'attuale RNMR possa essere integrato, anche in virtù delle nuove regolamentazioni in ambito sanitario e sugli aggiornamenti che lo stesso probabilmente subirà nel prossimo biennio.

Per evidenziare l'effettiva efficacia ed efficienza del thesaurus e delle potenzialità che lo stesso offre, si prenda ad esempio il termine "Disautonomia familiare" (Fig.15):

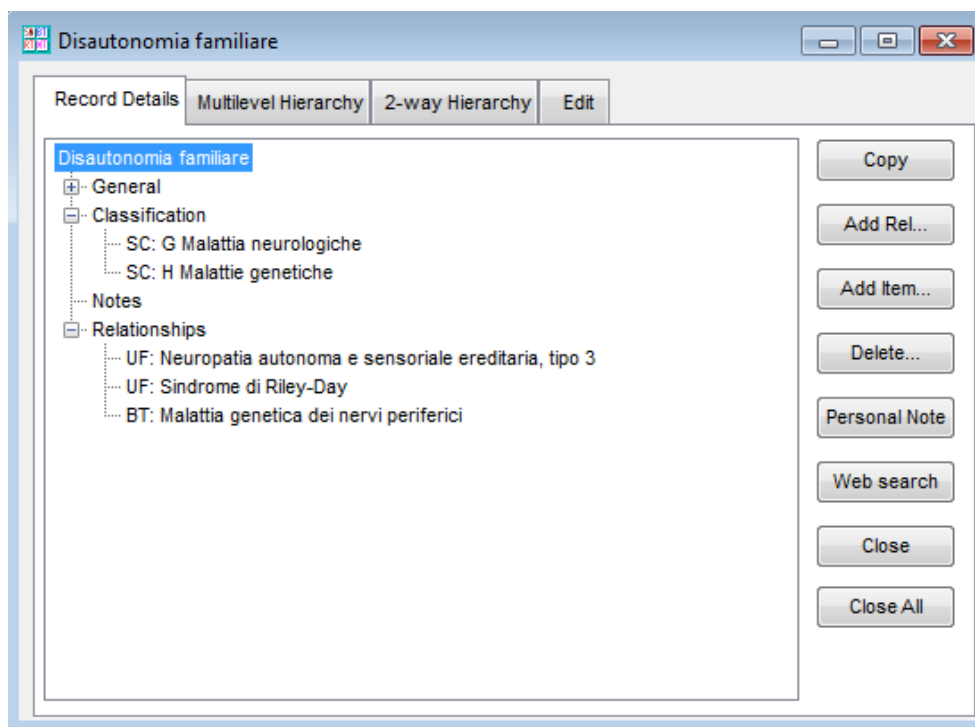


Figura 15 Scheda del termine "*Disautonomia familiare*" strutturato attraverso il software MultiTes.

Il termine qui evidenziato, presente all'interno del RNMR come termine singolo, privo di qualunque tipo di relazione, è stato arricchito di non solo di relazioni gerarchiche, ma anche con l'indicazione di più categorie di appartenenza permettendo così di ottenere più chiavi di accesso per la ricerca ed eventuali approfondimenti. Inoltre le relazioni sinonimiche permettono di evidenziare il termine utilizzato dalla comunità scientifica (identificato dalla sigla USE) e il/i termine/i utilizzati dalla comunità non scientifica (identificato dalla sigla UF), ma che comunque vengono utilizzati da utenti non specialistici per effettuare ricerche e approfondimenti.

Il thesaurus potrà, quindi, essere utilizzato dalla comunità scientifica per l'accesso rapido alle informazioni in rete. Parte dei termini integrati nel thesaurus, infatti, fanno parte dei principali sistemi di classificazione utilizzati dalla comunità scientifica, come ad esempio SNOMED CT, OMIM o MeSH. Quindi ciascun termine sarà collegato al rispettivo codice in SNOMED CT, OMIM e così via. Questo collegamento dà inoltre la possibilità di avere la denominazione

standardizzata per ciascuna patologia, aspetto non da trascurare in fase di trascrizione, ad esempio, della Scheda di Dimissione Ospedaliera. Una trascrizione non standardizzata, come spesso avviene nell'attuale panorama sanitario, può portare sia ad errori di interpretazione di tipo semantico, ma anche ad una non corretta corrispondenza verso i codici dei tariffari regionali/nazionali ai fini di rimborso per le prestazioni cliniche.

4.4 Prosecuzione del lavoro e prospettive future

Il Servizio Sanitario è in una fase di grande rinnovamento e sempre maggiore risulta essere la necessità di una razionalizzazione e maggiore efficienza dei servizi erogati ai cittadini. Il raggiungimento degli obiettivi passa attraverso un miglioramento della comunicazione delle diverse parti che agiscono all'interno del Servizio Sanitario Nazionale. È auspicabile, quindi, in tal senso che la scarsa comunicazione fra medici di famiglia e quelli ospedalieri si possa superare con la sintesi ideale fra i bisogni dei cittadini/pazienti e la tecnologia dei grandi ospedali attraverso, ad esempio, l'applicazione diffusa di sistemi di comunicazione digitale.

L'efficienza del servizio sanitario in ogni momento del percorso sanitario e socio-sanitario si incrementa attraverso azioni quali l'alleggerimento dell'onere documentale, la personalizzazione delle cure, la riduzione dell'errore umano e lo sviluppo di una sanità centrata sul cittadino come sottolineato anche nelle indicazioni dell'OMS dal 2002 (*Towards a Common Language for Functioning, Disability and Health: ICF The International Classification of Functioning, Disability and Health*).

A tal fine, una raccolta corretta ed il più esauriente possibile dei dati clinici di un paziente ha un ruolo importante sia nella pratica medica quotidiana sia nella gestione clinica del malato, nonché nel corretto iter delle prestazioni fornite dal sistema sanitario. Ciò con il fine di migliorare il servizio reso all'utenza evitando il ricorso alle diverse committenze che convergono sullo stesso assistito attraverso una pluralità di attori,

competenze e responsabilità. Le procedure di integrazione consentiranno la continuità assistenziale e terapeutica fra centro clinico di riferimento, Azienda Sanitaria, MMG e livello domiciliare attraverso una rete di condivisione delle informazioni cliniche necessarie all'attuazione del piano assistenziale per definire un Piano di Assistenza Individualizzato.

Un thesaurus appare come lo strumento adeguato a tale scopo in quanto aiuta a gestire la complessità di una terminologia specialistica tramite l'utilizzo di relazioni concettuali, fornendo al contempo una struttura gerarchica per la classificazione delle conoscenze di dominio, una via di accesso efficiente alle informazioni ricercate ed un utile strumento di discernimento e disambiguazione dei significati.

L'idea è quella di proseguire con l'alimentazione dei termini mancanti e di costruire una matrice di corrispondenza con la versione inglese per ciascun termine. Laddove non dovesse esistere la traduzione italiana, ne verrà creata una nuova, previa validazione da parte di un comitato di esperti del settore. Tuttavia, al fine di non realizzare un mero lavoro di traduzione e di adeguare il thesaurus alle necessità del contesto nazionale, si farà ricorso alla letteratura scientifica di settore in lingua italiana per la costruzione di corpora documentali su cui effettuare operazioni di estrazione terminologica al fine di arricchire il lessico già individuato per questo lavoro di ricerca con patologie rare peculiari dell'area mediterranea.

Una volta completo e con periodici aggiornamenti, il thesaurus potrebbe essere applicato sia a livello ospedaliero per un supporto terminologico durante la codifica delle SDO da parte dei professionisti del settore, ma anche da parte degli utenti finali, quali i pazienti, come strumento di ricerca di informazioni. L'utilizzo in questo senso potrebbe portare ad un arricchimento del thesaurus con termini utilizzati più comunemente dai "non addetti ai lavori" quali i cittadini/pazienti, inserendoli come chiavi di ricerca collegato con delle relazioni sinonimiche ai termini preferiti.

Indice delle Figure

Figura 1 Schermata di rappresentazione di un termine all'interno del database Orphanet.....	16
Figura 2 Schermata di visualizzazione della rappresentazione di un termine all'interno dell'ORDR.....	18
Figura 3 Visualizzazione della rappresentazione di un termine all'interno del database della NORD.	20
Figura 4 Campione dei dati provenienti dal database di Orphanet	31
Figura 5 Schermata di visualizzazione del MeSH browser.	45
Figura 6 Rappresentazione delle patologie rare nelle risorse di partenza	47
Figura 7 Mappatura delle risorse di partenza attraverso l'UMLS.	49
Figura 8 Applicazione del filtro semantico.....	50
Figura 9 Matrice della m appatura delle terminologie di partenza verso l' UMLS	54
Figura 10 Percentuali di rappresentazione delle risorse terminologiche all'interno dell'UMLS.....	55
Figura 11 Esempio di mappatura non ambigua.....	56
Figura 12 Esempio di mappatura ambigua	56
Figura 13 Esempio di termini non mappati.....	57
Figura 14 Rappresentazione dei termini non mappati.....	76

Bibliografia

AITCHISON, JEAN, ET AL., *Thesaurus Construction and use: a practical manual*, 4th Edition, London, ASLIB, 2000.

ANSI/NISO Z39.19-2005, *Guidelines for the construction, format, and management of monolingual controlled vocabularies*, Bethesda MD, NISO Press, 2005

BIBLIOTECA NAZIONALE CENTRALE DI FIRENZE, *Nuovo Soggettario*, Milano, Editrice Bibliografica, 2006.

BROUGHTON VANDA, *Costruire Thesauri: strumenti per indicizzazione e metadati semantici*, P. Cavaleri (a cura di), L. Ballestra e L. Venuti (traduzione di), Milano: Editrice Bibliografica, 2008.

ISO 25964-1:2011, *Information and documentation – Thesauri and interoperability with other vocabularies, Part 1: Thesauri for information retrieval*, 2011.

ISO 2788:1986, *Documentation - Guidelines for the establishment and development of monolingual thesauri*, 1986.

LENCI, ALESSANDRO, et alii, *Testo e computer: Elementi di Linguistica Computazionale*, Roma, Carocci Editore, 2005.

MCCRAY ALEXA T., BURGUN ANITA, BODENREIDER OLIVIER, *Aggregating UMLS Semantic Types for Reducing Conceptual Complexity*, IOS Press, Volume 84/2001, MedINFO, Amsterdam: IOS Press, pp. 216-220.

MERABTI TAYEB, JOUBERT MICHEAL, LECROQ THIERRY, RATH ANA, DARMONI STEPHAN, *Mapping biomedical terminologies using natural language processing tools and UMLS: mapping the Orphanet thesaurus to the MeSH*, IRBM, Vol. 31, Issue 4, September 2010.

MILICIC-BRANDT MAYA, RATH ANA, DEVEREAU A., AYME S., *Mapping Orphanet Terminology to UMLS*, in AIME 2011 Proceedings of the 13th conference on Artificial intelligence in medicine, pp. 194-203, Springer-Verlag: Berlin, Heidelberg, 2011.

PUBLIC LAW 107-280, Nov. 6, 2002, Sec. 2. Findings and Purposes.
Commission Decision, 30th November 2009 (2009/872/EC).

RUCH PATRICK, GOBEILL JULIEN, LOVIS CHRISTIAN, GEISSBÜHLER ANTOINE,
(2008) *Automatic medical encoding with SNOMED categories*, BMC Medical
Informatics and Decision Making, 2008.

UNI ISO 5963:1985, Documentazione - *Metodi per l'analisi dei documenti, la
determinazione del loro soggetto e la selezione dei termini di
indicizzazione*, 1985.

MINISTERO DELLA SANITÀ, *ICD-10 : classificazione statistica internazionale delle
malattie e dei problemi sanitari correlati:10ª revisione*, 3 volumi, Roma:
Istituto poligrafico e Zecca dello Stato, 2001.

SERAFINA SPINELLI, *Introduzione all'indicizzazione*, 2006.
<<http://www.biocfarm.unibo.it/~spinelli/indicizzazione/>>

JWAN AITCHISON, ALAN GILCHRIST, DAVID BAWDEN, *Thesaurus construction and use:
a practical manual*, London: Aslib, 2000, 4th ed.

Sitografia

<<http://www.orpha.net>>

<<http://www.eurordis.org>

<<http://www.eurordis.org/it>

<<http://www.hon.ch/HONcode>>

<www.ama-assn.org>

<www.conseil-national.medecin.fr>

<<http://www.epirare.eu>>

<<http://www.nlm.nih.gov>>

< <http://www.ncbi.nlm.nih.gov>>

<<http://rarediseases.info.nih.gov>>

<<http://it.wikipedia.org>>

<<http://www.iss.it> >

<<http://www.w3.org/XML/>>

<<http://zthes.z3950.org/>>

<<http://www.orpha.net> >

Data di ultima consultazione 22/12/2012

Appendice

(Presentazione alfabetica del thesaurus)

Aceruloplasminemia

NT: Aceruloplasminemia congenita

Neurodegenerazione con accumulo cerebrale di ferro

Aceruloplasminemia congenita

BT: Aceruloplasminemia

Acrocefalosindattilia

BT: Craniosinostosi sindromica

NT: Sindrome di Apert

Sindrome di Pfeiffer

Sindrome di Saethre-Chotzen

Acrodisostosi

BT: Disostosi mandibolofacciale

NT: Disostosi acrofacciale postassiale

Disostosi acrofacciale, tipo Nager

Disostosi mandibolofacciale legata all'X

SC: Anomalie dello sviluppo durante l'embriogenesi

Adams-Oliver, sindrome di

BT: Sindromi con difetti di riduzione degli arti

Adiposi dolorosa

UF: Dercum, malattia di

Neurolipomatosi

BT: Malattia del tessuto sottocutaneo

Adrenoleucodistrofia legata all'X

BT: Leucodistrofie

Adrenoleucodistrofia neonatale

BT: Leucodistrofie

Agenesia pancreatica e cerebellare

USE: Diabete mellito neonatale permanente - agenesia pancreatica
e cerebellare

Alopecia

BT: Malattia cutanea genetica rara

NT: Alopecia - deficit immunitario

Alopecia areata totale

Alopecia areata universale

Alopecia focale congenita - megalencefalia

Alopecia - deficit immunitario

BT: alopecia

Alopecia - epilessia - piorrea - ritardo mentale

BT: alopecia

Alopecia areata totale

BT: alopecia

Alopecia areata universale

BT: alopecia

Alopecia focale congenita - megalencefalia

BT: alopecia

Alpers, sindrome di

BT: Malattia mitocondriale associata a neuropatia periferica

Altra sindrome con lissencefalia come segno cardinale

BT: Lissencefalia

Amiloidosi

NT: Amiloidosi primarie e familiari

Amiloidosi da beta2-microglobulina

BT: Amiloidosi primarie e familiari

Amiloidosi da transtiretina

BT: Amiloidosi primarie e familiari

Amiloidosi familiare, tipo Finlandese

BT: Amiloidosi primarie e familiari

Amiloidosi primarie e familiari

BT: Amiloidosi

NT: Amiloidosi da beta2-microglobulina

Amiloidosi da transtiretina

Amiloidosi familiare, tipo Finlandese

Amiloidosi primitiva

Amiloidosi renale familiare

Amiloidosi primitiva

BT: Amiloidosi primarie e familiari

Amiloidosi renale familiare

BT: Amiloidosi primarie e familiari

Aminoaciduria, tipo Hartnup

USE: Hartnup, sindrome di

Anemia deficitaria acquisita rara

BT: Anemie ereditarie

Anemie ereditarie

NT: Anemia deficitaria acquisita rara

Angioedema

NT: angioedema acquisito

Angioedema ereditario

SC: I Malattia reumatologica e sistemica

Angioedema acquisito

BT: angioedema

Angioedema ereditario

BT: angioedema

Angioma o difetto vascolare

NT: Malformazione vascolare

Aniridia

BT: Cataratta sindromica

NT: Aniridia isolata

Aniridia sindromica

Aniridia isolata

BT: Aniridia

Aniridia sindromica

BT: Aniridia

Ano imperforato

BT: Malformazione anorettale

SC: L Anomalie dello sviluppo durante l'embriogenesi

Anomalia della pubertà e/o del ciclo mestruale di origine genetica

NT: Ipogonadismo ipogonadotropo congenito

Anomalia di Axenfeld-Rieger

UF: Anomalia di Rieger

BT: Iridogoniodisgenesia

SC: H Malattie genetiche

Anomalia di migrazione delle cellule della cresta neurale con glaucoma

BT: Glaucoma secondario disgenetico

NT: Iridogoniodisgenesia

Anomalia di Peters

UF: Glaucoma congenito di Peters

BT: Difetto non sindromico dello sviluppo oculare

NT: Anomalia di Peters - cataratta

Anomalia di Von Hippel

SC: Malattie genetiche

Anomalia di Peters - cataratta

BT: Anomalia di Peters

Anomalia di Rieger

USE: Anomalia di Axenfeld-Rieger

Anomalia di Von Hippel

BT: Anomalia di Peters

Anomalie congenite multiple

NT: Filippi, sindrome di

Anomalie congenite multiple senza deficit cognitivo

NT: Microcefalia

Anomalie dei piedi

RT: Jackson-Weiss, sindrome di

Anomalie di chiusura del tubo neurale

NT: Malformazione del canale neuroenterico, del midollo spinale
e della colonna

Anomalie ossee dei cingoli

BT: Malattia genetica rara

NT: Sprengel, anomalia di

Arnold-chiari, sindrome di

UF: Malformazione di Arnold-Chiari, tipo 1

Malformazione di Chiari, tipo 1

Malformazione di Chiari, tipo I

BT: Malformazione del canale neuroenterico, del midollo spinale
e della colonna

SC: Malattie genetiche

Atresia biliare

UF: BA

BT: Malformazioni viscerali non sindromiche

Atresia del digiuno

USE: Atresia del piccolo intestino

Atresia del piccolo intestino

UF: atresia del digiuno

BT: Malformazione intestinale

Malformazione intestinale non sindromica

BA

USE: Atresia biliare

Bean, sindrome di

UF: Sindrome del Nevo blu

BT: Malformazione venosa

SC: L Anomalie dello sviluppo durante l'embriogenesi

Behcet, malattia di

BT: Malattia infiammatoria e autoimmune con epilessia

SC: G Malattia neurologiche

Borreliosi di Lyme

USE: Lyme, malattia di

Camptodattilia - displasia ossea

UF: Camptodattilia di Goodman

BT: Condrodisplasia

SC: L Anomalie dello sviluppo durante l'embriogenesi

Camptodattilia di Goodman

USE: Camptodattilia - displasia ossea

Caroli, malattia di

BT: Malformazioni viscerali non sindromiche

Cataratta rara

BT: Malattia oculare genetica

NT: Cataratta sindromica

SC: H Malattie genetiche

Cataratta sindromica

BT: Cataratta rara

NT: Aniridia

Chavany-Marie, Sindrome di

USE: Chediak-higashi, malattia di

Chediak-higashi, malattia di

UF: Chavany-Marie, Sindrome di

Sindrome del nucleo rosso superiore

BT: Difetti congeniti del metabolismo

SC: H Malattie genetiche

Cheratosi follicolare

USE: Darier, malattia di

Chiray-Foix, sindrome di

SC: G Malattia neurologiche

CLS

USE: Coffin-Lowry, sindrome di

Coffin-Lowry, sindrome di

UF: CLS

BT: Deficit cognitivo sindromico legato all'X

SC: H Malattie genetiche

Coffin-Siris, sindrome di

BT: Malattia neurologica genetica

SC: H Malattie genetiche

Coloboma corio-retinico

BT: Coloboma oculare

Coloboma del cristallino

BT: Coloboma oculare

Coloboma della papilla ottica

BT: Coloboma oculare

Coloboma delle palpebre

BT: Coloboma oculare

Coloboma dell'iride

BT: Coloboma oculare

Coloboma maculare

BT: Coloboma oculare

Coloboma oculare

BT: Difetto non sindromico dello sviluppo oculare

NT: Coloboma corio-retinico

Coloboma del cristallino

Coloboma della papilla ottica

Coloboma delle palpebre

Coloboma dell'iride

Coloboma maculare

SC: Malattie genetiche

Condrodisplasia

NT: Camptodattilia - displasia ossea

Dyggve-Melchior-Clausen, sindrome di

Conkrite-Canada, malattia di

UF: pigmentazione cutanea

Poliposi gastrointestinale

SN: La Sindrome di Cronkhite-Canada è una rara sindrome acquisita, non familiare, che sembra verificarsi dopo importanti stress di tipo mentale, fisico ed emotivo. Tale malattia è caratterizzata dalla presenza di una diffusa poliposi gastrointestinale, alterazioni distrofiche delle unghie, alopecia, ipogeusia, parestesie, xerostomia, atrofia ungueale, iperpigmentazione cutanea, diarrea anoressia, perdita di peso, vomito, dolore addominale ed altre complicazioni gastrointestinali come un'enteropatia con perdita proteica e malnutrizione.

Corea di Huntington

UF: Còrea Maior

Còrea Maior

USE: Corea di Huntington

CPD IV

USE: Joubert, sindrome di

Craniosinostosi

NT: Craniosinostosi sindromica

RT: Jackson-Weiss, sindrome di

SC: E Malformazioni congenite

Craniosinostosi sindromica

BT: Craniosinostosi

NT: Acrocefalosindattilia

Crouzon, malattia di

Crioglobulinemia essenziale

USE: Crioglobulinemia mista

Crioglobulinemia mista

UF: Crioglobulinemia essenziale

Crioglobulinemia mista essenziale

Crioglobulinemia primitiva

Vasculite crioglobulinemica

NT: Crioglobulinemia mista, tipo II

Crioglobulinemia mista, tipo III

Crioglobulinemia mista essenziale

USE: Crioglobulinemia mista

Crioglobulinemia mista, tipo II

BT: Crioglobulinemia mista

Crioglobulinemia mista, tipo III

BT: Crioglobulinemia mista

Crioglobulinemia primitiva

USE: Crioglobulinemia mista

Crouzon, malattia di

UF: Disostosi craniofacciale di Crouzon

BT: Craniosinostosi sindromica

SC: Malattie genetiche

Cutis Laxa

NT: Cutis laxa - anomalie gravi polmonari, gastrointestinali e
delle vie urinarie

Cutis laxa autosomica dominante

Cutis laxa autosomica recessiva, tipo 1

Cutis laxa autosomica recessiva, tipo 2

Gerodermia osteodisplastica

Sindrome di Cantu

Sindrome MACS

Sindrome SCARF

SC: Anomalie dello sviluppo durante l'embriogenesi

Cutis laxa - anomalie gravi polmonari, gastrointestinali e delle vie urinarie

BT: Cutis Laxa

Cutis laxa autosomica dominante

BT: Cutis Laxa

Cutis laxa autosomica recessiva, tipo 1

BT: Cutis Laxa

Cutis laxa autosomica recessiva, tipo 2

BT: Cutis Laxa

Darier, malattia di

UF: Cheratosi follicolare

SC: Malattie genetiche

Deficienza di Acth

BT: Deficit acquisito degli ormoni pituitari, escluso il
craniofaringioma

SC: Malattie endocrine

Deficit acquisito degli ormoni pituitari, escluso il craniofaringioma

BT: Insufficienza ipofisaria

NT: Deficienza di Acth

Deficit cognitivo sindromico legato all'X

NT: Coffin-Lowry, sindrome di

Deficit di alfa-1-antitripsina

BT: Nefropatia secondaria a malattia metabolica da accumulo o
altra

SC: Malattie genetiche

Deficit di citocromo ossidasi di Saguenay-Lac-St. Jean

BT: Leigh, malattia di

Dercum, malattia di

USE: Adiposi dolorosa

Diabete mellito neonatale

BT: Diabete mellito raro

NT: Diabete mellito neonatale permanente - agenesia pancreatica
e cerebellare

Diabete mellito neonatale permanente - agenesia pancreatica e cerebellare

UF: Agenesia pancreatica e cerebellare

BT: Diabete mellito neonatale

SC: Malattie endocrine

Diabete mellito raro

NT: Diabete mellito neonatale

Difetti congeniti del metabolismo

NT: Chediak-higashi, malattia di

Malattia del trasporto degli aminoacidi

Difetto del metabolismo degli aminoacidi o di altri acidi organici

NT: Malattia del trasporto degli aminoacidi

Difetto della fosforilazione ossidativa mitocondriale da anomalie del DNA
mitocondriale

NT: Kearns-Sayre, sindrome di

Difetto della fosforilazione ossidativa mitocondriale da delezione singola a grande scala del DNA mitocondriale

NT: Kearns-Sayre, sindrome di

Difetto non sindromico dello sviluppo oculare

NT: Anomalia di Peters

Coloboma oculare

Disautonomia familiare

UF: Neuropatia autonoma e sensoriale ereditaria, tipo 3

Sindrome di Riley-Day

BT: Malattia genetica dei nervi periferici

SC: Malattia neurologiche

Malattie genetiche

Discheratosi congenita

UF: DKC

Sindrome di Zinsser-Engman-Cole

BT: Sindrome dismorfica con amartosi

SC: Anomalie dello sviluppo durante l'embriogenesi

Disostosi acrofaciale postassiale

BT: Acrodisostosi

Disostosi acrofacciale, tipo Nager

BT: Acrodisostosi

Disostosi con prevalente coinvolgimento costale e vertebrale

NT: Klippel-Feil, sindrome di

Disostosi craniofacciale di Crouzon

USE: Crouzon, malattia di

Disostosi mandibolofacciale

NT: Acrodisostosi

Disostosi mandibolofacciale legata all'X

BT: Acrodisostosi

Displasia caudale

USE: Sequenza da regressione caudale

Displasia craniometafisaria

BT: Sindrome dismorfica che interessa le ossa

Distonia di torsione idiopatica

BT: Malattia distonica

SC: Malattie genetiche

Disturbo cerebelloparenchimale IV

USE: Joubert, sindrome di

Disturbo dello sviluppo sessuale 46,XY

BT: Malattia dello sviluppo sessuale di origine endocrina

NT: Sindrome da insensibilità agli androgeni

DKC

USE: Discheratosi congenita

Dyggve-Melchior-Clausen, sindrome di

BT: Condrodisplasia

SC: Anomalie dello sviluppo durante l'embriogenesi

Ectasia canalicolare precaliceale

USE: Rene con midollare a spugna

Ehlers danlos, sindrome di

NT: Sindrome della cornea fragile

Sindrome di Ehlers-Danlos con eterotopia periventricolare

Sindrome di Ehlers-Danlos, tipo 10

Sindrome di Ehlers-Danlos, tipo 11

Sindrome di Ehlers-Danlos, tipo 5

Sindrome di Ehlers-Danlos/osteogenesi imperfetta

Embriofetopatia infettiva

BT: Malattia teratologica

NT: Embriopatia da citomegalovirus

Embriopatia da parvovirus

Embriopatia da virus della varicella

Infezione congenita da enterovirus

Infezione congenita da virus herpes

Infezione congenita da virus di Epstein-Barr

Sindrome da rosolia congenita

Toxoplasmosi congenita

SC: Malattie infettive

Anomalie dello sviluppo durante l'embriogenesi

Embriopatia da citomegalovirus

BT: Embriofetopatia infettiva

Embriopatia da parvovirus

BT: Embriofetopatia infettiva

Embriopatia da virus della varicella

BT: Embriofetopatia infettiva

Emoglobinuria parossistica notturna

BT: Malattia ematologica genetica

SC: Malattie genetiche

Encefalopatia con corpi inclusi di neuroserpina, forma familiare

BT: Epilessia mioclonica progressiva

Epidermolisi bollosa acquisita

BT: Malattia autoimmune bollosa cutanea

Epilessia mioclonica progressiva

BT: Sindrome epilettica con esordio nell'adolescenza

NT: Encefalopatia con corpi inclusi di neuroserpina, forma
familiare

Epilessia mioclonica progressiva, tipo 3

Epilessia mioclonica progressiva, tipo 6

Epilessia mioclonica progressiva, tipo Unverricht-Lundborg

Malattia di Lafora

SC: Malattie genetiche

Epilessia mioclonica progressiva, tipo 3

BT: Epilessia mioclonica progressiva

Epilessia mioclonica progressiva, tipo 6

BT: Epilessia mioclonica progressiva

Epilessia mioclonica progressiva, tipo Unverricht-Lundborg

BT: Epilessia mioclonica progressiva

Ermafroditismo Vero

RT: Pseudoermafroditismo

SC: E Malformazioni congenite

Farber, malattia di

BT: Malattia cutanea genetica rara

Fibrosi Cistica

BT: Malattia respiratoria genetica

SC: H Malattie genetiche

Filippi, sindrome di

BT: Anomalie congenite multiple

SC: Anomalie dello sviluppo durante l'embriogenesi

Focomelia, tipo Schinzel

BT: Sindromi con difetti di riduzione degli arti

Gardner, malattia di

BT: Malattia con possibile degenerazione neoplastica associata
a segni oculari

Gastroschisi

BT: Malformazione non sindromica del diaframma e della parete
addominale

GCPS

USE: Greig, sindrome di

Gerodermia osteodisplastica

BT: Cutis Laxa

Glaucoma congenito di Peters

USE: Anomalia di Peters

Glaucoma secondario disgenetico

NT: Anomalia di migrazione delle cellule della cresta neurale
con glaucoma

Granulomatosi lipofagica intestinale

USE: Whipple, malattia di

Greig, sindrome di

UF: GCPS

SC: Malformazioni congenite

Hansen, malattia di

UF: Lebbra

BT: Malattie infettive croniche

SC: Malattie infettive

Hartnup, sindrome di

UF: Aminoaciduria, tipo Hartnup

Malattia di Hartnup

BT: Malattia del trasporto degli aminoacidi

Hirschprung, malattia di

BT: Malformazione intestinale non sindromica

Incontinentia pigmenti

BT: Sindrome da displasia ectodermica

Infezione congenita da enterovirus

BT: Embriofetopatia infettiva

Infezione congenita da virus herpes

BT: Embriofetopatia infettiva

Insufficienza ipofisaria

NT: Deficit acquisito degli ormoni pituitari, escluso il
craniofaringioma

Iperaldosteronismi primitivi

BT: Malattia surrenale rara

NT: Iperaldosteronismo primitivo raro correggibile
chirurgicamente

Iperaldosteronismo primitivo raro non correggibile
chirurgicamente

Iperaldosteronismo primitivo raro correggibile chirurgicamente

BT: Iperaldosteronismi primitivi

Iperaldosteronismo primitivo raro non correggibile chirurgicamente

BT: Iperaldosteronismi primitivi

Ipogonadismo ipogonadotropo congenito

BT: Anomalia della pubertà e/o del ciclo mestruale di origine
genetica

NT: Ipogonadismo ipogonadotropo congenito isolato

Ipogonadismo ipogonadotropo congenito isolato

BT: Ipogonadismo ipogonadotropo congenito

NT: Kallmann, sindrome di

Iridogoniodisgenesia

BT: Anomalia di migrazione delle cellule della cresta neurale
con glaucoma

NT: Anomalia di Axenfeld-Rieger

Ittiosi

NT: Ittiosi acquisita

Ittiosi ereditaria

Ittiosi acquisita

BT: Ittiosi

Ittiosi ereditaria

BT: Ittiosi

Jackson-Weiss, sindrome di

RT: anomalie dei piedi

Craniosinostosi

Joubert, sindrome di

UF: CPD IV

Disturbo cerebelloparenchimale IV

Sindrome di Joubert classica

Sindrome di Joubert pura

Sindrome di Joubert, tipo A

Sindrome di Joubert-Boltshauser

BT: Malformazione cerebellare genetica

SC: Malattie genetiche

Kallmann, sindrome di

BT: Ipogonadismo ipogonadotropo congenito isolato

SC: Malattie genetiche

Kearns-Sayre, sindrome di

BT: Difetto della fosforilazione ossidativa mitocondriale da
anomalie del DNA mitocondriale

Difetto della fosforilazione ossidativa mitocondriale da
delezione singola a grande scala del DNA mitocondriale

Klippel-Feil, sindrome di

BT: Disostosi con prevalente coinvolgimento costale e vertebrale

Lebbra

USE: Hansen, malattia di

Leigh, malattia di

BT: Mitocondriopatia con malattia oculare

NT: Deficit di citocromo ossidasi di Saguenay-Lac-St. Jean

Sindrome di Leigh con cardiomiopatia

Sindrome di Leigh con leucodistrofia

Sindrome di Leigh con sindrome nefrosica

Leprecaunismo

UF: Sindrome di Donohue

BT: Sindrome da insulino-resistenza

Leucodistrofie

BT: Malattia neurologica genetica

NT: Adrenoleucodistrofia legata all'X

Adrenoleucodistrofia neonatale

Lipodistrofia intestinale

USE: Whipple, malattia di

Lipodistrofia totale

BT: Malattia endocrina genetica

SC: Malattie genetiche

Lissencefalia

BT: Sindrome genetica con malformazione del sistema nervoso
centrale come segno cardinale

NT: Altra sindrome con lissencefalia come segno cardinale

Lissencefalia classica

Lissencefalia con ipoplasia cerebellare

Lissencefalia da mutazioni di TUBA1A

Lissencefalia, tipo 2

Lissencefalia, tipo 3

Microlissencefalia

SC: Malattie genetiche

Lissencefalia classica

BT: Lissencefalia

Lissencefalia con ipoplasia cerebellare

BT: Lissencefalia

Lissencefalia da mutazioni di TUBA1A

BT: Lissencefalia

Lissencefalia, tipo 2

BT: Lissencefalia

Lissencefalia, tipo 3

BT: Lissencefalia

Lyme, malattia di

UF: Borreliosi di Lyme

BT: Malattia infettiva batterica

SC: Malattie infettive

Maculopatia primaria o secondaria non classificata

NT: Malattia cutanea genetica rara

Malattia autoimmune bollosa cutanea

NT: Epidermolisi bollosa acquisita

Malattia con possibile degenerazione neoplastica associata a segni oculari

NT: Gardner, malattia di

Malattia congenita cistica del Fegato

USE: Malattia del fegato policistico

Malattia cutanea genetica rara

BT: Maculopatia primaria o secondaria non classificata

NT: alopecia

Farber, malattia di

SC: Malattie genetiche

Malattia cutanea rara

NT: Malattia del tessuto sottocutaneo

Malattia del fegato policistico

UF: MALATTIA CONGENITA CISTICA DEL FEGATO

SN: Patologia ereditaria caratterizzata dalla formazione di
multiple cisti nel fegato che, nel 50% dei casi, si
sviluppano anche nei reni.

Le cisti epatiche sono di dimensione variabile, da pochi
millimetri ad alcuni 15 centimetri di diametro.

Malattia del tessuto sottocutaneo

BT: Malattia cutanea rara

NT: Adiposi dolorosa

Malattia del trasporto degli aminoacidi

BT: Difetti congeniti del metabolismo

Difetto del metabolismo degli aminoacidi o di altri acidi
organici

NT: Hartnup, sindrome di

SC: Malattie genetiche

Malattia dello sviluppo sessuale di origine endocrina

NT: Disturbo dello sviluppo sessuale 46,XY

Malattia di Cacchi-Ricci

USE: Rene con midollare a spugna

Malattia di Hartnup

USE: Hartnup, sindrome di

Malattia di Lafora

BT: Epilessia mioclonica progressiva

Malattia distonica

NT: Distonia di torsione idiopatica

Malattia ematologica genetica

NT: Emoglobinuria parossistica notturna

Malattia endocrina genetica

NT: Lipodistrofia totale

Malattia genetica dei nervi periferici

NT: Disautonomia familiare

Malattia genetica rara

NT: Anomalie ossee dei cingoli

Malattia granulomatosa cronica

Malattia infettiva batterica

NT: Lyme, malattia di

Malattia infiammatoria e autoimmune con epilessia

NT: Behcet, malattia di

Malattia lisosomiale

NT: Mucopolipidosi

Malattia metabolica associata a segni oculari

NT: Mitocondriopatia con malattia oculare

Malattia mitocondriale associata a neuropatia periferica

NT: Alpers, sindrome di

Malattia neurologica genetica

NT: Coffin-Siris, sindrome di

Leucodistrofie

Malattia oculare genetica

NT: Cataratta rara

Neuropatia ottica

Malattia respiratoria genetica

NT: Fibrosi Cistica

Malattia surrenale rara

NT: Iperaldosteronismi primitivi

Malattia teratologica

NT: Embriofetopatia infettiva

Malattia vitreo-retinica

NT: Retinoblastoma

Malattie infettive croniche

NT: Hansen, malattia di

Malformazione anorettale

NT: Ano imperforato

Malformazione cerebellare genetica

BT: Malformazione genetica della fossa posteriore

NT: Joubert, sindrome di

Malformazione cerebrale genetica

NT: Oloprosencefalia

Malformazione del canale neuroenterico, del midollo spinale e della colonna

BT: Anomalie di chiusura del tubo neurale

NT: Arnold-chiari, sindrome di

Sequenza da regressione caudale

Malformazione della parete addominale o diaframmatica

NT: Malformazione non sindromica del diaframma e della parete
addominale

Malformazione di Arnold-Chiari, tipo 1

USE: Arnold-chiari, sindrome di

Malformazione di Chiari, tipo 1

USE: Arnold-chiari, sindrome di

Malformazione di Chiari, tipo I

USE: Arnold-chiari, sindrome di

Malformazione genetica della fossa posteriore

BT: Malformazione non sindromica genetica del sistema nervoso
centrale

NT: Malformazione cerebellare genetica

Malformazione intestinale

NT: Atresia del piccolo intestino

Malformazione intestinale non sindromica

Malformazione intestinale non sindromica

BT: Malformazione intestinale

NT: Atresia del piccolo intestino

Hirschprung, malattia di

Malformazione non sindromica del diaframma e della parete addominale

BT: Malformazione della parete addominale o diaframmatica

NT: Gastroschisi

SC: Anomalie dello sviluppo durante l'embriogenesi

Malformazione non sindromica genetica del sistema nervoso centrale

NT: Malformazione genetica della fossa posteriore

Malformazione sindromica degli arti

NT: Poland, sindrome di

Sindromi con difetti di riduzione degli arti

Malformazione vascolare

BT: Angioma o difetto vascolare

NT: Malformazione venosa

Malformazione venosa

BT: Malformazione vascolare

NT: Bean, sindrome di

Malformazioni degli organi: fegato, tratto biliare, pancreas, milza

NT: Malformazioni viscerali non sindromiche

Malformazioni viscerali non sindromiche

BT: Malformazioni degli organi: fegato, tratto biliare,
pancreas, milza

NT: Atresia biliare

Micosi rara

SC: Malattie infettive

Microcefalia

UF: Sindrome di Teebi-Kaurah

BT: Anomalie congenite multiple senza deficit cognitivo

Microlissencefalia

BT: Lissencefalia

Mitocondriopatia con malattia oculare

BT: Malattia metabolica associata a segni oculari

NT: Leigh, malattia di

Morning glory, anomalia di

BT: Neuropatia ottica

MSK

USE: Rene con midollare a spugna

Mucopolipidosi

BT: Malattia lisosomiale

NT: Mucopolipidosi, tipo 2

Mucopolipidosi, tipo 3

Mucopolipidosi, tipo 4

SC: Malformazioni congenite

Mucopolipidosi, tipo 2

BT: Mucopolipidosi

Mucopolipidosi, tipo 3

BT: Mucopolipidosi

Mucopolipidosi, tipo 4

BT: Mucopolipidosi

Nefroblastoma

UF: Tumore di Wilms

SC: Malattie genetiche

Nefropatia secondaria a malattia metabolica da accumulo o altra

NT: Deficit di alfa-1-antitripsina

Neurodegenerazione con accumulo cerebrale di ferro

BT: Aceruloplasminemia

SC: Malattia neurologiche

Neurolipomatosi

USE: Adiposi dolorosa

Neuropatia autonoma e sensoriale ereditaria, tipo 3

USE: Disautonomia familiare

Neuropatia ottica

BT: Malattia oculare genetica

NT: Morning glory, anomalia di

Infezione congenita da virus di Epstein-Barr

BT: Embriofetopatia infettiva

Oloprosencefalia

BT: Malformazione cerebrale genetica

NT: Oloprosencefalia alobare

Oloprosencefalia lobare

Oloprosencefalia microforma

Oloprosencefalia semilobare

Oloprosencefalia, variante interemisferica della linea

mediana

SC: Malattie genetiche

Oloprosencefalia alobare

BT: Oloprosencefalia

Oloprosencefalia lobare

BT: Oloprosencefalia

Oloprosencefalia microforma

BT: Oloprosencefalia

Oloprosencefalia semilobare

BT: Oloprosencefalia

Oloprosencefalia, variante interemisferica della linea mediana

BT: Oloprosencefalia

Pallister W, sindrome di

UF: Sindrome di Pallister W

BT: Sindrome della schisi orofacciale

SC: Anomalie dello sviluppo durante l'embriogenesi

Persistenza della membrana pupillare

SN: Rara anomalia congenita caratterizzata da una incompleta involuzione della tunica vasculosa lentis . Il mancato riassorbimento della membrana pupillare (membrana di Wackendorf) , che avviene verso il sesto mese di vita intrauterina , comporta la persistenza di residui , il cui significato patologico dipende dalla loro posizione ed estensione . dalla persistenza di residui del sistema del cristallino che di regola viene riassorbito durante lo sviluppo embrionale.

Pigmentazione cutanea

USE: Conkite-Canada, malattia di

Poichiloderma ereditaria

NT: Werner, sindrome di

Poland, sindrome di

BT: Malformazione sindromica degli arti

SC: L Anomalie dello sviluppo durante l'embriogenesi

Poliposi adenomatosa coloretale

USE: Poliposi adenomatosa familiare

Poliposi adenomatosa familiare

UF: Poliposi adenomatosa coloretale

Poliposi familiare del colon

BT: Poliposi familiare

Poliposi familiare

NT: Poliposi adenomatosa familiare

Poliposi familiare del colon

USE: Poliposi adenomatosa familiare

Poliposi gastrointestinale

USE: Conkite-Canada, malattia di

Progeria dell'adulto

USE: Werner, sindrome di

Pseudoermafroditismo

RT: Ermafroditismo Vero

SC: E Malformazioni congenite

Reifenstein, sindrome di

UF: Sindrome da resistenza parziale agli androgeni

BT: Sindrome da insensibilità agli androgeni

SN: La sindrome di Reifenstein fa parte di un gruppo di

malattie in cui il corpo non è in grado di rispondere adeguatamente agli ormoni sessuali maschili (androgeni), che includono il testosterone.

Quando un bambino di sesso maschile cresce nel ventre materno, gli ormoni maschili (androgeni) vengono prodotti.

Questo porta allo sviluppo degli organi sessuali maschili.

In questa condizione vi è un cambiamento nel gene che aiuta

il corpo a riconoscere ed utilizzare correttamente gli

ormoni maschili. Questo porta a problemi con lo sviluppo

degli organi sessuali maschili. Alla nascita, il bambino può avere genitali ambigui, che conducono alla confusione sul sesso del bambino.

Rene con midollare a spugna

UF: Ectasia canalicolare precaliceale

Malattia di Cacchi-Ricci

MSK

SC: Malformazioni congenite

SN: Malattia cistica della porzione midollare delle piramidi renali, asintomatica se non complicata da infezioni, calcoli od ostruzione; va distinta dalla Malattia policistica congenita dei reni.

Retinoblastoma

BT: Malattia vitreo-retinica

SC: Malattie genetiche

Sequenza da regressione caudale

UF: Displasia caudale

Sindrome da agenesia sacrale

Sindrome da regressione sacrale

BT: Malformazione del canale neuroenterico, del midollo spinale e della colonna

Sindrome da agenesia sacrale

USE: Sequenza da regressione caudale

Sindrome da displasia ectodermica

NT: Incontinentia pigmenti

Sindrome da insensibilità agli androgeni

BT: Disturbo dello sviluppo sessuale 46,XY

NT: Reifenstein, sindrome di

SC: Malattie endocrine

Sindrome da insulino-resistenza

NT: Leprecaunismo

Sindrome da regressione sacrale

USE: Sequenza da regressione caudale

Sindrome da resistenza parziale agli androgeni

USE: Reifenstein, sindrome di

Sindrome da rosolia congenita

BT: Embriofetopatia infettiva

Sindrome del Nevo blu

USE: Bean, sindrome di

Sindrome del nucleo rosso superiore

USE: Chediak-higashi, malattia di

Sindrome della cornea fragile

BT: Ehlers danlos, sindrome di

Sindrome della schisi orofacciale

NT: Pallister W, sindrome di

Sindrome di Apert

BT: Acrocefalosindattilia

Sindrome di Cantu

BT: Cutis Laxa

Sindrome di Donohue

USE: Leprecaunismo

Sindrome di Ehlers-Danlos con eterotopia periventricolare

BT: Ehlers danlos, sindrome di

Sindrome di Ehlers-Danlos, tipo 10

BT: Ehlers danlos, sindrome di

Sindrome di Ehlers-Danlos, tipo 11

BT: Ehlers danlos, sindrome di

Sindrome di Ehlers-Danlos, tipo 5

BT: Ehlers danlos, sindrome di

Sindrome di Ehlers-Danlos/osteogenesi imperfetta

BT: Ehlers danlos, sindrome di

Sindrome di Joubert classica

USE: Joubert, sindrome di

Sindrome di Joubert pura

USE: Joubert, sindrome di

Sindrome di Joubert, tipo A

USE: Joubert, sindrome di

Sindrome di Joubert-Boltshauser

USE: Joubert, sindrome di

Sindrome di Leigh con cardiomiopatia

BT: Leigh, malattia di

Sindrome di Leigh con leucodistrofia

BT: Leigh, malattia di

Sindrome di Leigh con sindrome nefrosica

BT: Leigh, malattia di

Sindrome di Pallister W

USE: Pallister W, sindrome di

Sindrome di Pfeiffer

BT: Acrocefalosindattilia

Sindrome di Plummer-Vinson

Sindrome di Riley-Day

USE: Disautonomia familiare

Sindrome di Saethre-Chotzen

BT: Acrocefalosindattilia

Sindrome di Teebi-Kaurah

USE: Microcefalia

Sindrome di Zinsser-Engman-Cole

USE: Discheratosi congenita

Sindrome dismorfica che interessa le ossa

NT: Displasia craniometafisaria

Sindrome isolata di Klippel-Feil

Sindrome dismorfica con amartosi

NT: Discheratosi congenita

Sindrome dismorfica non condrodiplasica con anomalie scheletriche

NT: Sindrome isolata di Klippel-Feil

Sindrome epilettica con esordio nell'adolescenza

NT: Epilessia mioclonica progressiva

Sindrome genetica con malformazione del sistema nervoso centrale con segno cardinale

NT: Lissencefalia

Sindrome isolata di Klippel-Feil

BT: Sindrome dismorfica che interessa le ossa

Sindrome dismorfica non condrodisplasica con anomalie
scheletriche

Sindrome MACS

BT: Cutis Laxa

Sindrome SCARF

BT: Cutis Laxa

Sindromi con difetti di riduzione degli arti

BT: Malformazione sindromica degli arti

NT: Adams-Oliver, sindrome di

Focomelia, tipo Schinzel

Sprengel, anomalia di

BT: Anomalie ossee dei cingoli

Sprue non tropicale secondaria

USE: Whipple, malattia di

Toxoplasmosi congenita

BT: Embriofetopatia infettiva

Tumore di Wilms

USE: Nefroblastoma

Vasculite crioglobulinemica

USE: Crioglobulinemia mista

Werner, sindrome di

UF: Progeria dell'adulto

WS

BT: Poichilodermia ereditaria

SN: La sindrome di Werner è una rara sindrome ereditaria caratterizzata da invecchiamento precoce a esordio nella terza decade di vita e da alcuni segni clinici cardinali: cataratta bilaterale, bassa statura, incanutimento e diradamento dei capelli, difetti cutanei caratteristici ed esordio precoce di altri disturbi associati all'avanzare dell'età.

Whipple, malattia di

UF: Granulomatosi lipofagica intestinale

Lipodistrofia intestinale

Sprue non tropicale secondaria

SC: Malattie infettive

WS

USE: Werner, sindrome di

Xeroderma pigmentoso

UF: XP

XP

USE: Xeroderma pigmentoso