

Università degli Studi di Catanzaro "Magna Graecia"



Facoltà di Medicina e Chirurgia

*Scuola di Specializzazione in
Medicina Interna
"P.L. Mattioli"*

TESI DI SPECIALIZZAZIONE

***Definizione utilità e limiti del registro calabrese delle
iperlipidemie***

Relatore

Prof. Arturo Pujia

Candidato

Dr. Vittorio Emanuele

Matricola 68292

Introduzione

In questa breve trattazione ci poniamo l'obiettivo di definire in che modo ed entro quali limiti possa essere utile l'istituzione di un registro, in Calabria, delle iperlipidemie, con particolare riferimento alle forme familiari di dislipidemie. Recentemente tutti i campi lavorativi sono stati massivamente invasi dall'uso di quella che, nel suo insieme, viene definita *Information Technology* o *IT*. La medicina e la ricerca scientifica più in generale, ovviamente, non hanno fatto eccezione, anzi, hanno cominciato fin dall'inizio a servirsi largamente di queste nuove tecnologie. Nessuno, oggi, penserebbe di raccogliere dati in forma cartacea o, meglio, solo in forma cartacea, oltre che per motivi di conservazione, soprattutto perché l'informatizzazione delle banche dati permette tempi di consultazione e accesso infinitamente più piccoli di quelli che la carta da sempre è in grado di offrire. Attualmente, nella comunità scientifica esistono già diverse realtà che si avvalgono dello strumento del registro per altri ambiti (per esempio registro oncologico o di malattie genetiche); curiosamente, non ne esiste uno per le dislipidemie, che hanno un valore, in negativo purtroppo, sul piano clinico, enormemente più rilevante degli altri ambiti di cui sopra. Vi sono, inoltre, motivi precisi, che analizzeremo in dettaglio in seguito, per cui si debba cominciare con la Calabria.

Definizione

Cerchiamo, innanzitutto, di spiegare perché l'interesse è caduto proprio sulle dislipidemie. Nei paesi occidentali, il maggior numero di morti è dovuto a eventi morbosi legati all'arteriosclerosi. Quest'ultima riconosce diversi fattori di rischio, alcuni non modificabili (il sesso, l'età), alcuni modificabili (ipertensione, ipercolesterolemica *etc.*), altri ambientali (fumo, dieta). Uno dei più importanti fattori di rischio per l'arteriosclerosi, nonché momento patogenetico iniziale, risulta essere l'ipercolesterolemia e più esattamente l'aumento del colesterolo LDL. A sua volta quest'ultimo può essere espressione di una miriade di patologie (familiari e non familiari) genericamente inquadrata come dislipidemie. Risulta evidente, quindi, come vi sia notevole attenzione da parte della comunità scientifica nel tentare di diagnosticare e curare queste malattie. Molti di questi difetti non hanno cause genetiche note e risentono fortemente dell'ambiente. Sul piano teorico una delle prospettive più allettanti sarebbe quella di utilizzare le informazioni provenienti dal progetto genoma, per potere identificare quali sono i geni coinvolti nelle alterazioni del metabolismo lipidico^{1,2,3}. Le tecniche classiche hanno come presupposto quello di correlare l'espressione del fenotipo con la presenza di *markers* sul genoma mediante studi di linkage. Questi studi hanno permesso l'identificazione di alcune dislipidemie (per esempio Ipercolesterolemia Autosomica Recessiva⁴). Sono, in genere, molto utili quando il numero di geni coinvolto nella patologia è limitato.

Risultano inutili e, peggio ancora, a volte fuorvianti, invece, nella patologia in cui i geni coinvolti sono molti o in cui la loro stessa espressione è fortemente regolata da interazioni gene-ambiente (per esempio Iperlipidemia Familiare Combinata⁵). A questo punto entra in gioco il registro che, quindi, potrebbe essere definito come uno strumento per la raccolta sistematica e strutturata di informazioni correlabili con la patologia dislipidica. Le informazioni dovrebbero provenire da tutti gli ambiti possibili (clinici, genetici, ambientali, geografici, *etc.*) così da poter essere in grado di correlare, in fase di analisi, in maniera corretta tutte le informazioni in possesso⁶.

L'acquisizione dei dati

Risulta evidente che il primo passo per la definizione di un registro sanitario è la raccolta dei dati⁷. Indipendentemente dal supporto primario (per esempio la carta), non è in discussione che l'archiviazione e gli strumenti di analisi debbano essere di tipo informatico. Vi sono diverse tecniche per la raccolta dei dati che non saranno oggetto di questa trattazione^{8,9}. E' importante, però, dire che esistono dei canoni universali ed indipendenti dall'infrastruttura informatica utilizzata. La bioinformatica, insieme ad altro, si occupa specificatamente delle tecniche di raccolta e strutturazione delle informazioni bio-mediche^{10,11,12}. La caratteristica più importante, in questo ambito, è quella dell'universalità. Una banca dati è molto più utile se i dati in essa contenuti possono essere condivisi e/o mischiati con altre banche dati. Alla luce di questo, la loro quantità passa in secondo ordine. Un'altra caratteristica importante dei database è la loro relazionalità. I dati raccolti devono essere correlati tra loro e la natura della relazione deve essere studiata in maniera tale da poter associare elementi diversi di uno stesso record nel minor tempo possibile. Quello che deve essere sempre conservato, quindi, è il flusso logico-informativo. Questo inizia con l'inserimento dei dati o *Input*, continua con il processamento, la memorizzazione, e la distribuzione, che avviene mediante l'ultima fase del processo informativo, cioè la trasmissione. L'implementazione di questi processi nel loro insieme ha permesso alla moderna medicina la strutturazione di sistemi

informativi eccezionali, senza i quali oggi nessun medico e nessun ricercatore potrebbero mai lavorare o studiare. Pensiamo, per esempio, alla differenza che esiste fra un laureando di oggi e uno di soli 10-15 anni fa. Il primo, per effettuare una ricerca bibliografica in merito ad un argomento, impiega un tempo compreso fra 0,1 e 5 secondi, necessario per la risoluzione della *query* (domanda che viene effettuata al database e che si basa su operatori logici universali booleani) da parte del motore di ricerca del database e per l'invio delle informazioni al computer terminale da cui l'interrogazione viene effettuata (per esempio internet). Fino a 10-15 anni fa, invece, la ricerca veniva effettuata manualmente sugli *Index Medicus* o pubblicazioni simili e poteva, quindi, impiegare anche giorni. Questo esempio è emblematico: lo stesso livello di sviluppo (e in alcuni casi anche maggiore) ha riguardato tutti i campi della ricerca scientifica, con ripercussioni sul piano della produttività che è inutile sottolineare.

Processo fondamentale, subito successivo all'inserimento, come abbiamo visto nel flusso informativo, è la memorizzazione. E' impressionante la velocità con cui cresce la capienza dei dispositivi per la memorizzazione di massa e anche degli algoritmi informativi che permettono alla macchina di funzionare. Tale velocità correla perfettamente con quella di crescita della capacità di calcolo dei processori, descritta nella famosa legge di Moore (il numero di transistor in un *microchip* raddoppia ogni 18 mesi). In questo momento nella comunità scientifica internazionale esistono pochissimi esempi (pubblici) di

sistemi informativi di questo tipo (database delle mutazioni per il gene del recettore per le LDL, gestito dalla *College University* di Londra, e sistema MEDPAD con scopo simile al precedente, ma in fase di costituzione presso la *University of Utah*, USA). Questi esempi, comunque, non sono orientati allo studio clinico, ma solo ad una funzione di supporto per gli studi di biologia molecolare (per esempio raccolta di *primer* utilizzati per la ricerca delle mutazioni). Per potere essere utile anche sul piano clinico (per l'assistenza e l'epidemiologia), un database di questo tipo dovrebbe contenere informazioni di correlazione genotipo-fenotipo al variare del tipo di popolazione. Sicuramente è utile collezionare tutte le mutazioni di un recettore che producono una malattia, per sapere in che modo una mutazione può essere evidenziata nei pazienti, ma è sicuramente più utile, per esempio, conoscere in che modo queste mutazioni si distribuiscono nella popolazione dal punto di vista geografico. In questo modo un paziente con diagnosi fenotipica di FH¹³ (Ipercolesterolemia Familiare), che proviene da una regione in cui il 90% delle mutazioni sono del tipo G528D (per esempio la zona di Palermo, in Sicilia), permetterà uno screening molecolare rapido e nella maggior parte dei casi poco costoso^{14,15}. Affinché questo possa avvenire, però, è necessario che vi sia la possibilità di raccogliere, strutturare, conservare e rendere accessibili informazioni di questo tipo (oggi in Italia non esiste un sistema di questo tipo che, invece, è ben sviluppato in Inghilterra e soprattutto in Olanda¹⁶, paese pioniere in questo campo). La necessità di sistemi simili è quanto mai indispensabile soprattutto nel centro-sud della nostra

penisola, in particolar modo in Calabria e Sardegna. In queste due regioni, infatti, si verificano condizioni rarissime sul pianeta, che consentono di classificare queste zone geografiche come “isole genetiche”. In altre parole, la bassissima immigrazione, abbinata con l’elevata emigrazione e l’elevata prevalenza e incidenza di matrimoni fra consanguinei (anche cugini di I grado), fa sì che la mescolanza genetica sia molto bassa e che, di conseguenza, la prevalenza di malattie genetiche sia elevata. Questo effetto, di recente, ha consentito di utilizzare dati in gran parte provenienti proprio dalla Sardegna per l’identificazione di un nuovo tipo di iperlipidemia familiare (Ipercolesterolemia Familiare Recessiva⁴), che, per il tipo di trasmissione, difficilmente si sarebbe potuta identificare in zone con mescolanza genetica più elevata. Ciò nonostante, però, ad oggi siamo in grado di stimare solo approssimativamente la prevalenza delle diverse dislipidemie in Sardegna, mentre nulla o poco si sa sulle popolazioni calabresi. In queste regioni, per complicare e abbellire il quadro dal punto di vista scientifico, è presente una bassa incidenza di malattie cardiovascolari rispetto ai paesi del nord Europa (per esempio rispetto all’Olanda, dove pure vi è una discreta mappatura delle dislipidemie almeno per quanto riguarda le ipercolesterolemie), per esempio, a significare che non esiste la mutazione di un recettore (LDL) ma il paziente in cui questa mutazione si verifica¹⁷. Proprio il principale gruppo di ricerca olandese in merito ha recentemente pubblicato proprio dati di questo tipo, in cui si dimostra come, a parità di mutazioni, la mortalità anche per le sole malattie cardiovascolari varia

al variare delle famiglie e come, anche in presenza di gravi forme di ipercolesterolemia, alcune famiglie esprimono una longevità inaspettata¹⁶. Risulta evidente, quindi, l'utilità che un sistema organico di raccolta di questi dati clinici e biochimici potrebbe avere. Informazioni sul fenotipo lipidico espresso a livello biochimico (ipercolesterolemia isolata piuttosto che mista con ipertrigliceridemia), sul fenotipo clinico (eventi cardiovascolari, xantomi, aterosclerosi carotidea, *etc.*), sul tipo di familiarità se presente sulle caratteristiche genetiche (mutazioni, satelliti ed altri marcatori fisici genetici), sul tipo di risposta alla terapia (sensibilità al trattamento, *non responders*, farmacogenetica, *etc.*) sono indispensabili per disegnare un quadro d'insieme su queste patologie in una data popolazione. Ottenere un'informazione di questo tipo è estremamente importante, perché questi dati possono servire per tantissime applicazioni, dalla stima di prevalenza e incidenza di malattie alle scelte farmacoeconomiche ed epidemiologiche di assistenza sanitaria.

Linee progettuali per la raccolta strutturata dei dati

Il record, nell'ipotetico registro per le dislipidemie, è ovviamente il paziente affetto da dislipidemia. Le informazioni inerenti ogni record (in gergo si definiscono "campi del record"), poi, possono essere divise in tre grosse sezioni. La prima dovrebbe riguardare i dati anagrafici del paziente. All'interno di questa sezione devono essere inseriti anche i dati riguardanti le informazioni anamnestiche familiari e personali. La seconda parte dovrebbe essere quella responsabile della raccolta dei dati fenotipici. L'ultima sezione, infine, si dovrebbe occupare della raccolta dei dati genetici. Si può discutere sulla politica di strutturazione dei dati. Per i fini del registro, infatti, potrebbe essere sufficiente l'archiviazione del caso inteso come record composto da pochi campi che comprendono la diagnosi, il luogo di origine del paziente ed eventualmente gli eventi clinici ad esso correlati (per esempio IMA, ICTUS, MORTE). La strutturazione di una banca dati più complessa, però, sarebbe auspicabile per gli evidenti vantaggi che questa è in grado di offrire rispetto alla semplice raccolta dei casi delle varie patologie. Per completare le scelte politiche non dovrebbe essere trascurata l'autorizzazione per l'accesso all'archivio del registro. Alcune di queste patologie, infatti, non sono facilmente diagnosticabili come altre (tumori, AIDS) e richiederebbero la possibilità di creare dei centri intermedi specialistici. I dati di un paziente visitato dal medico di medicina generale, per esempio, prima di essere inseriti nel registro, potrebbero essere inviati ad un

centro intermedio, in cui vengono revisionati. La creazione di un gruppo di lavoro, dove il registro viene gestito fisicamente, potrebbe essere una soluzione sostitutiva o additiva. Se si scegliesse il semplice inserimento del caso (inteso come tipo diagnosi ed eventi correlati), ovviamente, i sistemi di controllo sarebbero estremamente più snelli. Su queste scelte, evidentemente, gravano anche questioni di *budget*. All'aumentare di quest'ultimo possono crescere le risorse umane e fisiche dedicabili alla creazione del registro. Indubbia, comunque, è la sua utilità sul versante della ricerca scientifica e su quello dell'assistenza pubblica. Numerosi esempi dimostrano che il registro è una formula di studio e analisi vincente: il registro dell'AIDS e quello, più noto, dei tumori. Sempre di più, soprattutto per quello che riguarda le scelte per i finanziamenti e per il dislocamento delle risorse sanitarie, le informazioni provenienti dai registri risultano utili e insostituibili.

La raccolta dei dati anagrafici

Da questo punto in poi, quindi, non si farà netta distinzione fra registro e sistema informativo, in quanto nella realtà organizzativa le due cose di fatto non sono divise¹⁸. Il registro, dunque, dovrebbe essere costituito, prima di tutto, da una “sezione” in cui vengono acquisite le informazioni anagrafiche. La raccolta dei dati, ovviamente, deve avvenire in maniera anonima o criptata e comunque sempre nel rispetto delle normative vigenti in materia di *privacy*. Questo vuol dire che il paziente deve essere informato e, quindi, fornire il suo consenso sul motivo per il quale i dati vengono raccolti, sugli scopi che la raccolta si pone e su chi è personalmente responsabile della procedura di raccolta e analisi di questi dati. La stesura di un documento di consenso informato, che il paziente visionerà e firmerà prima della raccolta dei dati, risulta, dunque, necessaria. Oltre alle informazioni scontate che dovrebbero essere raccolte (nome, cognome, *etc.*), particolare attenzione dovrebbe essere rivolta al luogo di nascita, alla presenza di eventuale parentela fra i genitori e alla raccolta di informazioni cliniche sui parenti. Questo è un punto di fondamentale importanza per lo studio di tutte le patologie di tipo familiare. In questa sezione dovrebbero essere, quindi, raccolte tutte le informazioni necessarie per la successiva (o contemporanea) stesura di un albero genealogico. Diversi algoritmi e sistemi di programmazione consentono anche di generare in maniera automatica, a partire da queste informazioni, un albero genealogico di tipo grafico, gradevole dal

punto di vista estetico ma, soprattutto, utile per definire la presenza e il tipo di trasmissione di un'eventuale patologia familiare¹⁹. Le informazioni in tal modo acquisite, ovviamente, sono insostituibili per la diagnosi della malattia, oltre che per la terapia e la prognosi del paziente. A tal proposito, e sempre per agevolare lo studio dell'albero genealogico, è bene identificare il soggetto anche nell'ambito della stessa famiglia (familiare affetto, sano o probando). Insieme all'anamnesi familiare bisogna raccogliere anche quella del paziente. In questo ambito particolare interesse deve essere dato alle informazioni relative allo stile di vita dell'individuo ma, soprattutto, agli eventi (nello specifico quelli cardiovascolari o anche quelli correlati con le altre iperlipidemie, come per esempio le pancreatiti nelle iperchilomicronemie) per il futuro calcolo della loro incidenza e prevalenza.

La raccolta dei dati fenotipici

Per dati fenotipici intendiamo la raccolta dei dati riguardanti il modo in cui le informazioni genetiche si esprimono sull'essere umano²⁰. In altre parole, si identificano quelli che Mendel chiamava i caratteri, la maggior parte dei quali, però, è espressione di un corredo di geni. Ognuno di questi partecipa in un determinato modo, affinché il carattere sia conservato nella specie. Questi dati rivestono un valore del tutto peculiare, in quanto caratterizzano la malattia a livello di genetica di popolazione. Diverse razze, infatti, rispondono diversamente anche se in ambienti simili. Quando per carattere intendiamo i valori di colesterolo plasmatico, per esempio, la cosa si complica ulteriormente. I livelli plasmatici dei lipidi, infatti, sono regolati da una quantità di attività enzimatiche, trasportatori di membrana e proteine con varie funzioni endo ed esocellulari. Spesso l'espressione anche di una sola di queste proteine è mediata o regolata da più geni e complica il quadro d'insieme in maniera piuttosto evidente. I valori di colesterolemia, come quelli della trigliceridemia, del colesterolo HDL, delle apolipoproteine, sono, ovviamente, essenziali nell'iter diagnostico di una dislipidemia, ma assolutamente da associare ad un insieme di altri dati fenotipici. Per quello che riguarda le forme familiari, cui il registro dovrebbe evidentemente rivolgere maggiore attenzione, di particolare importanza sono anche gli xantomi. Sono questi degli accumuli di lipidi, che si possono diversamente presentare e associare; la loro presenza, nelle diverse

varianti, risulta spesso patognomonica di un determinato disordine lipidico. Non dovrebbero, poi, mancare tutte le informazioni utili per le diagnosi differenziali con le dislipidemie secondarie e i dati riguardanti gli esami strumentali, volti a studiare, soprattutto, l'albero vascolare del soggetto. Tra gli elementi da raccogliere, ai fini della diagnosi differenziale, sono senz'altro gli indici di funzionalità epatica, renale, tiroidea, essendo proprio il fegato, i reni e la tiroide gli organi maggiormente responsabili di disordini che secondariamente producono dislipidemia. Sebbene di utilità indiscutibile, queste ultime informazioni potrebbero non essere di primaria rilevanza nell'ipotesi dell'istituzione di un registro dei soli casi. Tra gli esami strumentali, invece, non dovrebbe mancare un Eco-Doppler carotideo. E' noto, infatti, che la maggior parte dei pazienti dislipidemici morirà con eventi cardiovascolari di cui l'aterosclerosi è responsabile. Indagini mirate all'identificazione precoce e, soprattutto, preclinica della malattia aterosclerotica, quindi, saranno prioritarie. Il distretto vascolare carotideo, infatti, essendo estremamente superficiale e contemporaneamente importante sul piano delle dimensioni, si presta facilmente ad approfondite valutazioni flussimetriche ed ecografiche. Con quest'ultima metodica, inoltre, si è in grado di rilevare un marker surrogato di aterosclerosi, lo spessore medio intimale, che è un sensibile predittore di eventi cardiovascolari. Anche l'indice caviglia-braccio (o indice di Winsor) appartiene a quest'ultima categoria di indici ben correlati con gli eventi e, pertanto, neanche questo dovrebbe essere mai trascurato in un paziente dislipidemico. Non bisogna

dimenticare, infine, tutte le misure biometriche, tra cui l'indice di massa corporea e la circonferenza della vita, la cui straordinaria correlazione con eventi e patologie è paragonabile solo alla semplicità con la quale essi possono essere reperiti dal medico in fase di esame obiettivo. Riassumendo, per quello che riguarda la raccolta dei dati fenotipici, questi possono essere suddivisi in tre tipi. Il primo è costituito dalle informazioni di tipo obiettivo e, cioè, quelle derivanti dall'esame fisico del paziente (xantomi, indice di massa corporea, *etc.*). Il secondo è costituito dalle informazioni di tipo strumentale, che vengono acquisite dopo l'esame fisico del paziente, di cui sono un diretto completamento. Tra questi esami, come già sottolineato, annoveriamo in primis l'esame Eco-Doppler e l'indice di Winsor. Vi sono, infine, le informazioni relative ai dati biochimici del paziente (colesterolemia, trigliceridemia) e tutte le informazioni che derivano dall'analisi bio-chimica del sangue e che non sono di tipo genetico (apolipoproteine, *etc.*). Molti dei dati sopra analizzati possono essere influenzati da terapie farmacologiche concomitanti che, ovviamente, devono essere anche acquisite in fase di raccolta dei dati stessi. Siamo fin qui riusciti a strutturare la metodologia di raccolta delle informazioni anagrafiche del paziente, di quelle anamnestiche, personali e familiari, di quelle relative all'esame obiettivo del paziente, alle indagini strumentali e, infine, a quelle ematochimiche. Dopo avere analizzato il prodotto dell'espressione genica (enzimi, proteine e, quindi, il fenotipo) e avere raccolto dati sull'ambiente esterno, in cui l'espressione ha

avuto luogo (per esempio abitudini alimentari e voluttuarie), passiamo ora alla metodologia per la raccolta dei dati genetici.

La raccolta dei dati genotipici

Con la conclusione del progetto genoma la comunità scientifica internazionale ha a disposizione da pochi mesi il codice genetico dell'uomo. Questo ha prodotto l'archiviazione di molti Gbit²¹ di dati di diverso tipo. Pur trattandosi di un problema prettamente biologico (il sequenziamento e soprattutto l'interpretazione dei dati), infatti, la più importante sfida tecnica che gli scienziati si sono trovati a dover affrontare è stato di tipo informatico²². Il problema era stato evidenziato già dalle prime sequenze nucleotidiche, ma la necessità di risolverlo si è presentata solo con l'accumularsi dei dati, cui si è aggiunta l'eterogeneità delle informazioni e la complessità delle loro relazioni. Diversi tipi di soluzione, che non sono oggetto di questa trattazione, hanno consentito, comunque, di ottenere dei database di tipo relazionale, veloci, facili da consultare e soprattutto molto ben correlati tra loro. Questi sistemi informativi consentono a chiunque (la maggior parte, infatti, sono liberamente accessibili) di cercare sequenze nucleotidiche e proteiche, di capire se si ci trova in un tratto di genoma codificante, promotore o altro; contengono, insomma, un'immensa mole di informazioni il cui fine ultimo dovrebbe essere quello di essere messo in relazione con le caratteristiche fenotipiche. In altre parole, la sequenza nucleotidica in sé è un'informazione di tipo linguistico – matematico (sequenza di quattro tipi diversi di caratteri), che però contiene un numero enorme di altri dati (per esempio geni e loro regolazione espressiva) variamente

codificati (per esempio codice genetico). Queste informazioni sono la base per la produzione di proteine e, di conseguenza, di cellule, tessuti e organi. La sfida della medicina moderna è quella di mettere in relazione i dati genetici con quelli clinici allo scopo di prevenire, diagnosticare ed infine curare le malattie. L'approccio al problema è complesso per diverse ragioni. In primo luogo, l'analisi di questi dati richiede un'integrazione di conoscenze (matematiche, informatiche, biologiche e mediche) che non trova riscontro nelle figure professionali formate dalle nostre università. In secondo luogo, ancora oggi, le tecniche mediante cui questi studi si svolgono sono poche e in continua evoluzione. In questo mare di difficoltà, tuttavia, esistono dei punti fermi, metodiche e tecniche che, se ben condotte, danno enormi soddisfazioni. Quest'ultime sono facili da ottenere quando ci si trova di fronte a malattie monogeniche, in cui l'assioma "un gene una malattia" è valido. In questi casi (per esempio Ipercolesterolemia Familiare) un'alterazione nota in un gene (recettore delle LDL) produce un fenotipo molto preciso (aumento plasmatico delle LDL), che a sua volta causa clinica e patologia facilmente identificabili (infarto del miocardio in età giovanile). Rari sono, però, i casi in cui l'assioma di cui sopra si può applicare, dal momento che, per lo più, si tratta di malattie plurifattoriali e multigeniche, in cui il fenotipo (per esempio infarto del miocardio) è la risultante di un insieme di alterazioni genetiche (diversi polimorfismi) ed un insieme di alterazioni ambientali (fumo) e/o malattie concomitanti (diabete). La presenza di una sola delle precedenti alterazioni in un

singolo individuo non è sufficiente a produrre il fenotipo patologico. Una delle funzioni del registro potrebbe essere proprio quella di creare database di correlazione genotipo-fenotipo da cui trarre spunto per studi di associazione genomica. Quest'ultimo potrebbe essere il vero valore aggiunto scientifico di un curato registro delle dislipidemie.

Utilità del registro

Cerchiamo, ora, di analizzare in dettaglio i vantaggi che una raccolta di dati di questo tipo potrebbe offrire. Diversi, in primo luogo, sono gli ambiti di applicazione. L'*output* principale di un sistema simile sarebbe sicuramente di tipo epidemiologico. Questo significa che si potrebbero stimare con esattezza prevalenza e incidenza delle patologie dislipidemiche. Questo, però, non sarebbe che l'inizio. Sempre per rimanere in ambito epidemiologico, si potrebbero collezionare informazioni sulla frequenza delle diverse dislipidemie nei diversi strati della popolazione (divisa per età piuttosto che per sesso, per esempio) o delle regioni, *etc.* Tutte le analisi di questo tipo sarebbero possibili. L'utilità immediata di queste informazioni sta nel fatto di poter prendere decisioni di sanità pubblica e anche di farmacoeconomia. E' evidente, infatti, che se dovesse risultare dai dati del registro che le dislipidemie sono particolarmente frequenti in una regione italiana, per esempio, in quella regione le decisioni su come dispensare farmaci ipolipidemizzanti e prestazioni sanitarie correlate non potrebbero essere uguali alle decisioni prese in un'ipotetica regione dove non esistono queste malattie. Questa estremizzazione, ovviamente, ci è utile per intravedere un'applicazione in situazioni intermedie di utilizzo delle informazioni del registro sulle dislipidemie. Un altro campo di applicazione indipendente dall'epidemiologia è la genetica. La possibilità di poter associare determinati genotipi con varianti fenotipiche permetterebbe agli scienziati il

disegno di studi *ad hoc* per valutare questa associazione e stabilire, in tal modo, meccanismi patogenetici per quelle patologie dislipidemiche che ancora ne sono prive. Di questo tipo di approccio potrebbe giovare una frequente dislipidemia familiare, l'Iperlipidemia Familiare Combinata. Questa patologia sarebbe responsabile di un infarto su quattro in Italia e, tuttora, non riconosce un'eziopatogenesi precisa, pur presentando, sul piano genetico, una trasmissione di tipo autosomico dominante. Troppo lungo, e anche inutile, sarebbe descrivere un approccio di studio di questo tipo per tutte le dislipidemie, cambiando, nella pratica, solo i termini delle associazioni. Oltre all'epidemiologo e al genetista, comunque, le informazioni ricadrebbero, come già sottolineato, anche al clinico. Avere un'idea precisa di prevalenza e incidenza modifica profondamente l'iter diagnostico di un clinico; avere un'idea di come una determinata patologia risponde all'ambiente semplifica la scelta terapeutica e i consigli di stile di vita che il clinico fornisce al paziente; avere un'idea del modo con cui probabilmente quel dato difetto genetico reagirà al trattamento farmacologico, infine, ha una utilità che si commenta da sé.

Limiti del registro

Come tutti gli strumenti, anche il registro non sarebbe privo di difetti e limiti, che sono, comunque, largamente inferiori per numero e importanza ai vantaggi che se ne avrebbero. Prima di tutto, vi sono i limiti intrinseci all'istituzione. La creazione di un registro presuppone la dedizione di persone e mezzi con il relativo costo. Tenendo presente questo risvolto, tutt'altro che trascurabile, possiamo suddividere i limiti e difetti del registro in tre tipi. Vi sono i limiti legati alla strutturazione (impossibilità di inserire un numero altissimo di campi per ogni record), quelli connessi alla compilazione del registro (errori di inserimento), quelli derivanti dall'uso (errori di interpretazione, di gestione, di sicurezza, di uso improprio). E' evidente che l'inserimento di accortezze in ognuna di queste fasi è in grado di ridurre al minimo gli errori e rendere il sistema sicuro e accurato.

Conclusioni

Diversi sono oggi i campi che si avvalgono dell'uso di un registro a cui segnalare e in cui conservare i dati provenienti dall'ambito sanitario e non solo. Indubbiamente, l'esperienza maturata in questi anni grazie all'utilizzo dei registri esistenti (soprattutto quello sui tumori e sull'AIDS) consente di affermare con sicurezza che l'istituzione di un registro permette un indubbio avanzamento sul piano delle conoscenze scientifiche ed epidemiologiche. Queste si traducono, in tempi medio – lunghi, in pubblici vantaggi per le conoscenze che si acquisiscono su aspetti fondamentali delle patologie, consentendo contemporaneamente una migliore e più razionale assistenza sanitaria e un più rapido progresso delle conoscenze scientifiche. E' pertanto auspicabile l'istituzione di un registro con le caratteristiche sopraelencate per le dislipidemie che, nel loro ruolo di fattore di rischio fondamentale per l'aterosclerosi, sono sicuramente responsabili di uno dei più importanti problemi sanitari di tutti i Paesi occidentali. Si ritiene, inoltre, che in Calabria un tale registro potrebbe avere particolarmente successo a causa delle particolari caratteristiche genetiche di cui la sua popolazione è portatrice.

BIBLIOGRAFIA

1. Functional Genomics. *Insight. Nature* 2000;405:819.

2. Human Genomics and Medicina. *Insight. Nature* 2004;429:439.

3. Proteomics. *Insight. Nature* 2003;422:191.

4. Christine Kim Garcia, Kenneth Wilund, Marcello Arca, Giovanni Zuliani,⁴ Renato Fellin,⁴ Mario Maioli, Sebastiano Calandra, Stefano Bertolini, Fausto Cossu, Nick Grishin, Robert Barnes, Jonathan C. Cohen, Helen H. Hobbs. Autosomal Recessive Hypercholesterolemia Caused by Mutations in a Putative LDL Receptor Adaptor Protein. *SCIENCE* 2001;292:1394.

5. A.D. Sniderman, J. Ribalta, and M. Castro Cabezas. How should FCHL be defined and how should we think about its metabolic bases? *Nutr Metab Cardiovasc Dis* 2001;11:259-273.

6. A. Camussi et al. *Metodi Statistici per la Spermitantazione Biologica*. Zanichelli. 2002

7. SUSHMITA MITRA, TINKUACHARYA. *Data Mining Multimedia, Soft Computing, and Bioinformatics*. JOHN WILEY & SONS, INC. 2003.

8. Cynthia Gibas. *Developing Bioinformatics Computer Skills*. O'Reilly Inc. 2001.
9. Bryan Bergeron. *Bioinformatics Computing*. Prentice Hall PTR. 2002.
10. Arthur M. Lesk. *Introduction to Bioinformatics*. OXFORD UNIVERSITY PRESS 2002.
11. G. Valle et al. *Introduzione alla Bioinformatica*. Zanichelli. 2003.
12. A. Tramontano. *Bioinformatica*. 2002.
13. Dalya Marks et al. A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia. *Atherosclerosis* 2003;168:14.
14. Marks D. et al. Screening of hypercholesterolemia versus case finding for familial hypercholesterolemia: a systematic review and cost-effectiveness analysis. *HTA* 2000;4:29.
15. Dalya Marks et al. Cost effectiveness analysis of different approaches of screening for familial hypercholesterolaemia. *BMJ* 2002;324:1303.
16. Marina A.W. et al. Low-Density Lipoprotein Receptor Gene Mutations and

Cardiovascular Risk in a Large Genetic Cascade Screening Population. *Circulation*. 2002;106:3031-3036.

17. Mariko Harada-Shiba, Atsuko Takagi, Yoshihiro Miyamoto, Motoo Tsushima, Yasuyuki Ikeda, Shinji Yokoyama, And Akira Yamamoto. Clinical Features and Genetic Analysis of Autosomal Recessive Hypercholesterolemia. *J Clin Endocrinol Metab* 2003;88:2541–2547.

18. Andreas D. Baxevanis, B.F. Francis Ouellette. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, Second Edition. 2001 John Wiley & Sons, Inc.

19. James Tisdall. *Beginning Perl for Bioinformatics*. O'Reilly Inc. 2001.

20. Human Genome. Special. *Nature* 2001;409:813.

21. D. Mount. *Bioinformatics: sequence and genome analysis*. CSHL Press. 2003.

22. Computational Biology. Insight. *Nature* 2002;420:205.