

PROTEOMICA E BIOMARCATORI NELLA DIAGNOSTICA CLINICA E NEL FOLLOW-UP DI MALATTIA

Marco Crescenzi, Serena Camerini

Servizio Grandi Strumentazioni e Core Facilities, Istituto Superiore di Sanità, Roma

Introduzione

Il concetto di biomarcatore è sorprendentemente ampio. Il *Biomarkers Definitions Working Group* dei *National Institutes of Health* statunitensi lo definisce una caratteristica misurabile oggettivamente e utilizzabile come indicatore di processi biologici normali, di stati patologici o di risposta a interventi terapeutici (1). Nel concetto sono quindi comprese molecole quali anticorpi, metaboliti, proteine, glicidi, misurati nelle comuni analisi cliniche, ma anche dati strumentali, come le risultanze di un elettrocardiogramma o di una radiografia, o segni fisici di malattia, per esempio un fegato aumentato di volume o un soffio cardiaco. E la stessa definizione appena enunciata potrebbe essere ulteriormente allargata. In questo articolo, tuttavia, ci limitiamo a considerare i biomarcatori biochimici, con particolare riguardo a quelli proteici.

I biomarcatori possono avere diverse finalità. Quelli diagnostici sono di grande utilità nell'identificazione di una patologia; i marcatori prognostici aiutano a prevedere il decorso o l'esito di un'affezione, mentre i predittori di risposta alla terapia intervengono nell'identificazione dei pazienti che possono giovare di una determinata azione terapeutica.

L'utilità di un biomarcatore viene misurata mediante vari indici, stabiliti attraverso la valutazione statistica di ripetute misurazioni in appropriate popolazioni. I parametri più utili e più comunemente utilizzati sono i seguenti:

- Accuratezza: grado di corrispondenza fra il valore misurato, che contiene inevitabilmente errori, e quello “vero”;
- Precisione: la distanza fra diverse misurazioni di una stessa quantità nota (dispersione);
- Riproducibilità: la dispersione ottenuta compiendo le stesse misurazioni, per esempio, con strumenti e operatori differenti, in laboratori diversi, e/o in un intervallo temporale relativamente lungo;
- Robustezza: capacità di un metodo di misura di fornire risultati simili, resistendo a modifiche delle condizioni in cui la misura viene effettuata;
- Sensibilità: probabilità di un test, per esempio di un biomarcatore diagnostico, di indicare un risultato positivo nei pazienti colpiti dalla malattia. Si calcola dividendo il numero di pazienti positivi al test per quello di tutti gli individui affetti dalla patologia in esame;
- Specificità: reciprocamente alla sensibilità, è la probabilità di un test di indicare un risultato negativo nei soggetti indenni dalla patologia. Si calcola dividendo il numero di persone non affette negative al test per quello di tutti i soggetti non affetti.

I due ultimi indici, sensibilità e specificità, sono sempre relativi alla popolazione considerata e in particolare alla prevalenza della malattia (cioè la percentuale di individui affetti in un dato momento). Sensibilità e specificità possono essere combinati tra loro per calcolare la predittività positiva e negativa di un determinato biomarcatore.

Immaginiamo di misurare un biomarcatore diagnostico per una patologia della tarda età in una popolazione di individui compresi fra i 50 e gli 80 anni di età e di trovare che il risultato del test è anomalo (troppo alto o troppo basso) nel 97% dei pazienti affetti dalla malattia; supponiamo

inoltre che nel 95% dei soggetti non colpiti dalla patologia in esame il test risulti negativo (entro i valori normali). Questo significa che, “all’interno della popolazione considerata”, il biomarcatore ha una sensibilità del 97% e una precisione del 95%. Questi dati, da soli, non possono dirci quale sia la probabilità che un soggetto positivo al test sia realmente affetto dalla malattia (predittività positiva) o, al contrario, che una persona negativa ne sia davvero esente (predittività negativa). Per calcolare questi indici è necessario conoscere le caratteristiche della popolazione di riferimento.

Supponiamo quindi che, nel nostro caso, la popolazione comprenda un milione di individui e che fra questi la prevalenza della malattia sia dell’1%. Possiamo allora compilare una piccola tabella in cui saranno comprese quattro categorie: malati positivi al test, malati negativi, non affetti positivi e non affetti negativi. Le prime due categorie assommano a 10.000 persone (1% di 1.000.000), di cui il 97% (9.700) positive al test e il 3% (300) negative. Le persone non affette, invece, saranno 990.000 (1.000.000-10.000), 95% delle quali negative al test (940.500) e 5% positive (49.500) (Tabella 1).

Tabella 1. Calcolo delle predittività positiva e negativa di un biomarcatore diagnostico

Biomarcatore	Malati	Non affetti
positivo	9.700	49.500
negativo	300	940.500

Il valore di predittività positiva è dato da:

$$\frac{\text{n. di malati positivi per il biomarcatore}}{\text{n. di tutti i positivi per il biomarcatore}} = \frac{9.700}{9.700 + 49.500} = 16,4\%$$

Quindi un soggetto positivo per il biomarcatore ha solo il 16,4% di probabilità di essere malato.

Analogamente, il valore di predittività negativa è dato da:

$$\frac{\text{n. di non affetti negativi per il biomarcatore}}{\text{n. di tutti i negativi per il biomarcatore}} = \frac{940.500}{300 + 940.500} = 99,97\%$$

Perciò la negatività per il biomarcatore esclude quasi totalmente la patologia.

Poiché dipendono da sensibilità e specificità, anche gli indici di predittività positiva e negativa variano con la prevalenza della malattia nella popolazione di riferimento. Se avessimo considerato l’intera popolazione di una certa area geografica, invece della sola fascia in età avanzata, la prevalenza della patologia (che avevamo supposto essere tipica della tarda età) sarebbe stata inferiore e gli indici sarebbero variati di conseguenza.

Biomarcatori proteici

Come già accennato, la natura biochimica del biomarcatore può essere molto varia. Può trattarsi di un carboidrato, di un lipide, di un acido nucleico, di una piccola molecola o di altre specie ancora. Le proteine sono molto rappresentate fra i biomarcatori e di queste ci occuperemo in modo particolare.

Una prima domanda è in quale materiale biologico vada valutato il biomarcatore. In alcuni casi la misurazione viene effettuata su cellule o tessuti, per esempio nel caso di marcatori tumorali ricercati in materiale biotico. Ovunque possibile, però, si preferisce analizzare un fluido biologico, particolarmente il sangue, che può essere raccolto in modo semplice, economico e poco invasivo. Le proteine presenti nel siero (la parte liquida, non cellulare del sangue) possono avere diverse provenienze. In certi casi può trattarsi di proteine direttamente secrete nel sangue (ad esempio un fattore della coagulazione); in altri le proteine sono secrete solo localmente (ad esempio citochine in un tessuto infiammato), e solo secondariamente vengono trasportate nella corrente ematica, dove possono essere rilevate. Ancora in altri casi la proteina biomarcatore viene rilasciata da cellule danneggiate o morte, indicando quindi un danno o una patologia, e poi trascinata nel sangue. Si può osservare che il sangue potrebbe contenere i marcatori in concentrazione minore, in confronto ai tessuti/organi interessati dalla patologia. Ma è altrettanto vero che, in assenza di indicazioni riguardo la sede affetta dalla patologia (è questo il caso dello screening precoce per tumori) il sangue, come altri fluidi biologici accessibili (urina, saliva), diventa prezioso per la ricerca di marcatori.

Quando si ricerca un nuovo biomarcatore proteico, ci si può trovare in due diverse condizioni: in alcuni casi la sorgente del biomarcatore è nota, almeno in via presuntiva, come quando si vuole utilizzare una proteina per evidenziare un danno tissutale. In queste circostanze le migliori candidate sono proteine che siano abbondanti nel tessuto di interesse e possibilmente solo in esso.

Un'eventualità frequente è invece quella in cui si ricerchi per esempio un marcatore diagnostico per una patologia, ma non si abbiano indizi sulla sua identità. In questi casi si effettuano spesso ricerche "a tappeto", senza conoscere in anticipo le caratteristiche del potenziale biomarcatore, ma cercando una proteina la cui concentrazione vari in funzione della presenza o dell'assenza della patologia di interesse. È necessario ricorrere allora ad un metodo capace di identificare e misurare il maggior numero possibile di proteine nel materiale biologico di scelta. Negli ultimi venti anni la tecnologia maggiormente utilizzata per questo scopo è stata la spettrometria di massa (*Mass Spectrometry*, MS) delle proteine. Questa straordinaria tecnologia, tuttora in fase di veloce evoluzione, permette di misurare con estrema accuratezza il peso molecolare di minutissime quantità di proteine e dei loro frammenti. Sulla base di questi dati è possibile identificare la proteina allo studio e determinarne la concentrazione nel campione di origine. In pratica, nella maggior parte dei casi, le proteine vengono sottoposte ad una digestione enzimatica, che le riduce ad un insieme di peptidi; sono questi ultimi ad essere analizzati in MS.

Gli spettrometri di massa, anche se molto diversi fra loro, sono accomunati dalla stessa struttura generale, che comprende tre sezioni (Figura 1) (2). La prima, la sorgente, ionizza le molecole allo studio, i peptidi nel nostro caso, trasformandoli in ioni molecolari carichi positivamente in fase gassosa. Nella seconda sezione dello strumento, l'analizzatore, i peptidi vengono separati da potenti campi elettromagnetici sulla base del loro peso molecolare. In questa sede i peptidi possono essere fatti collidere con un gas, producendo un insieme di frammenti che vengono a loro volta separati in base al peso. La terza componente dello strumento è il rivelatore, che produce un segnale elettronico in funzione della quantità di ioni da cui viene raggiunto.

Gli spettrometri di massa identificano una proteina determinando il peso molecolare dei peptidi e dei loro frammenti prodotti nell'analizzatore. Questi dati sono utilizzati da appositi *software*, che li confrontano con banche dati di proteine e determinano quale sia l'unica proteina compatibile con i diversi peptidi e i loro frammenti. E non è tutto qui: i moderni strumenti sono in grado di effettuare queste operazioni non su una molecola alla volta, ma su miscele molto complesse. In questo modo un'analisi di poche ore può dare informazioni su centinaia o migliaia di proteine diverse. In questi casi si parla di strategie "proteomiche".

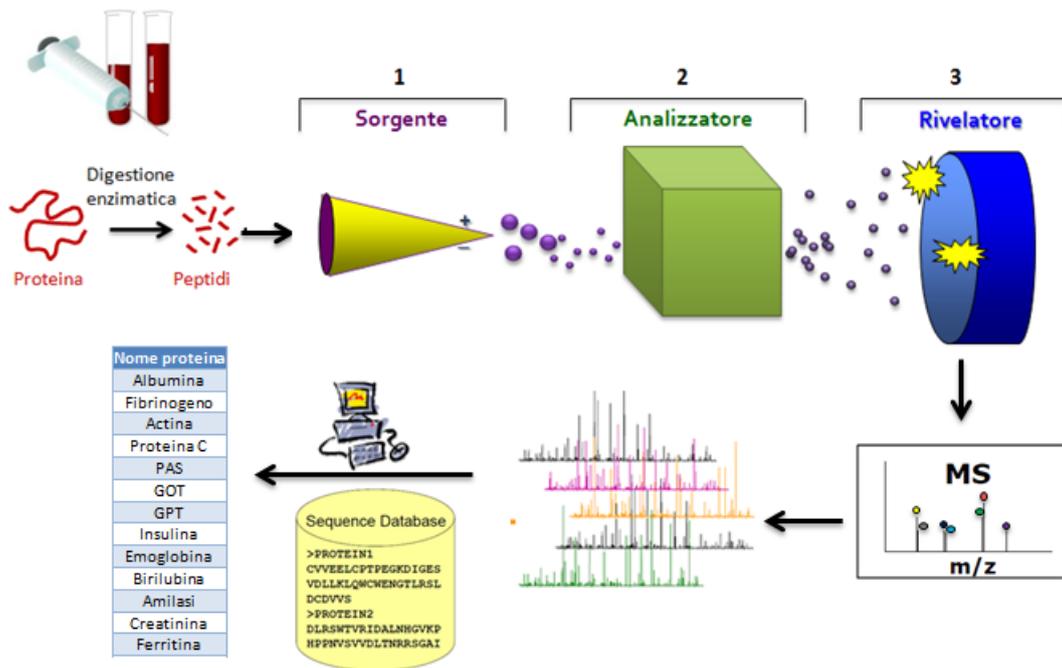


Figura 1. Processo per l'analisi di proteine mediante spettrometria di massa

Date determinate condizioni, gli spettrometri di massa sono anche in grado di misurare in quale quantità ciascuna proteina sia presente nel campione analizzato. In questo modo è possibile cercare correlazioni fra uno stato patologico e variazioni di concentrazione di una o più proteine nel siero. È da notare che per estrarre questa informazione diagnostica non è essenziale conoscere la funzione delle proteine coinvolte e neppure il motivo che ne fa variare le concentrazioni ematiche nel corso della patologia presa in esame. Può quindi accadere che un biomarcatore identificato nel modo descritto entri nell'uso clinico prima che si conoscano i suoi rapporti con la patologia per la quale serve da indicatore.

Nonostante la potenza dei moderni spettrometri di massa, la ricerca di biomarcatori proteici incontra formidabili ostacoli. Uno dei più importanti è l'enorme variabilità nella concentrazione delle proteine del siero. Il sangue è un collettore di informazioni da ogni distretto dell'organismo, perché le proteine secrete o rilasciate da qualunque cellula o tessuto finiscono prima o poi nel torrente circolatorio. Ne discende l'importanza di cercare potenziali biomarcatori proprio nel sangue. Ma se alterazioni macroscopiche si riflettono in grandi variazioni della composizione del siero (si pensi all'insufficienza epatica, che determina una cospicua riduzione della proteina sierica più abbondante, l'albumina), patologie che interessano un ristretto numero di cellule (ad esempio un tumore in uno stadio molto precoce) possono produrre cambiamenti infinitesimali. La difficoltà risiede proprio nel dover ricercare cambiamenti, che possono essere grandi o piccolissimi, della concentrazione di proteine ignote a priori. Fra le proteine più e meno abbondanti del siero intercorrono dieci ordini di grandezza. Chiedere a uno spettrometro di massa di analizzare le une e le altre equivale a pretendere che una bilancia pesi con uguale precisione oggetti di un milligrammo e di dieci tonnellate.

A queste difficoltà è possibile in parte ovviare preparando i campioni con metodiche volte ad arricchire le proteine poco abbondanti, prima di procedere con l'analisi mediante MS. La MS

prevede lunghe procedure di laboratorio e l'uso di costosi strumenti, che non ne facilitano l'utilizzo diretto nella diagnostica clinica. Tuttavia la MS si rivela estremamente utile nella fase di identificazione (*discovery*) dei biomarcatori proteici. Ad oggi, già numerosi biomarcatori sono stati individuati con strategie proteomiche e il loro successivo utilizzo in specifici test diagnostici, di più semplice e rapida esecuzione, è entrato nella pratica clinica (3-5). Tabelle che elencano quelli già in uso e altri in corso di approvazione sono state pubblicate (4).

Conclusioni

Il progresso tecnologico sta portando allo sviluppo di spettrometri di massa sempre più veloci, sensibili e accurati, strumenti preziosi per l'analisi sia qualitativa che quantitativa delle proteine anche in campioni complessi quali i fluidi biologici. L'utilizzo della spettrometria di massa sta dunque diventando sempre più frequente e anche più efficiente nella ricerca di nuovi biomarcatori (2).

Bibliografia

1. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics* 2001;69(3):89-95.
2. Crutchfield CA, Thomas SN, Sokoll LJ, Chan DW. Advances in mass spectrometry-based clinical biomarker discovery. *Clinical Proteomics* 2016;13:1-12.
3. Füzéry AK, Levin J, Chan MM, Chan DW. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin Proteomics* 2013;10(1):13.
4. Parker C E, Borchers CH. Mass spectrometry based biomarker discovery, verification, and validation - Quality assurance and control of protein biomarker assays. *Mol Oncol* 2014 Jun;8(4):840-58.
5. Verrills NM. Clinical proteomics: present and future prospects. *Clin Biochem Rev* 2006;27(2):99-116.