La stratificazione di rischio

Filippo Ottani

Unità di ricerca cardiovascolare, Fondazione cardiologica "MZ Sacco", Forlì - Italia

E'noto a tutti i medici che è impossibile ricordare tutti i dati riguardanti un paziente in medicina. Da ciò il vantaggio tratto dagli operatori sanitari legato all'uso di tecnologie informatiche che consentono l'impiego di linee guida per la pratica clinica quotidiana.

L'impiego di questi tipi di tecnologie rende anche utilizzabile, in modo relativamente semplice, sistemi di punteggio del rischio (*risk score systems*, RSS) assai complessi ed accurati. Il problema risiede nel definire gli standard per la creazione di questi RSS.

Primo, per essere applicabili alla pratica clinica, la coorte di derivazione dell'RSS deve rappresentare un quadro generale della popolazione. Secondo, gli RSS devono garantire predizioni accurate. Tale accuratezza è tipicamente descritta dal valore di *C statistic* o *C index*, dove la *C* sta per *Concordance*. Tale indice definisce quanto adeguatamente un modello o una regola di predizione è in grado di discriminare tra i pazienti che hanno o non hanno un evento. Per un end-point binario, il *C-statistic* rappresenta la proporzione di tutte le paia di pazienti, uno con e uno senza l'evento, in cui il paziente con l'evento ha una probabilità predetta più elevata di avere l'evento. In buona sostanza, il *C-statistic* è la misura di quanto efficacemente l'RSS stratifica i pazienti in base al loro grado di rischio previsto. Un modello che discrimina efficacemente i pazienti per l'85% delle volte dovrebbe avere un *C-statistic* di 0,85, mentre un modello che fornisce una previsione di rischio del tutto casuale, ovvero l'equivalente di tirare una moneta, dovrebbe avere un *C-statistic* di 0,50. Infine, un modello che discriminasse perfettamente i pazienti con e senza evento dovrebbe avere un *C-statistic* di 1.0. Pertanto un valore di *C-statistic* <0,6 viene considerato sostanzialmente privo di valore clinico, tra 0,6 e 0,7 un valore limitato, tra 0,7 e 0,8 un modesto valore, mentre solo valori >0,8 hanno una reale utilità clinica. E' degno di nota il fatto che gli RSS possono presentare una associazione statisticamente significativa con l'outcome clinico, ma la relazione può non essere sufficientemente discriminante da permettere al clinico di distinguere accuratamente i pazienti che hanno e che non hanno l'evento in questione.

Inoltre, il *C-statistic* è usualmente più elevato quando valuta l'accuratezza predittiva nella coorte di derivazione rispetto alla coorte di validazione.

E' importante anche definire l'evento per cui gli RSS sono stati sviluppati. Nei pazienti con sindrome coronarica acuta senza sopraslivellamento persistente del tratto ST-T (NSTEMI), ad esempio, sono stati sviluppati modelli predittivi di rischio di morte, (re)IMA, combinato di morte, (re)IMA non fatale e severa ischemia ricorrente.

Nelle NSTEMI, RSS derivati da coorti vaste e non fortemente selezionate sono quelli dello studio PURSUIT e dello studio GUSTO IIb. Entrambi contengono variabili predittive comuni, quali l'età avanzata, il sesso femminile, la frequenza cardiaca superiore a 100 battiti per minuto, la pressione sistolica ridotta (<100 mmHg) e la severità degli attacchi anginosi nelle ore/giorni precedenti. L'RSS derivato dal GUSTO IIb contiene invece delle variabili caratteristiche della popolazione studiata, quali l'ipertensione arteriosa, la vasculopatia periferica e l'abitudine al fumo. In entrambi i modelli la presenza di elevazione degli enzimi cardiaci (inteso come CK totale e/o CKMB) interagisce con i predittori sopraelencati nel selezionare una sottopopolazione a più alto rischio. Tuttavia il più famoso degli RSS nell'ambito dell NSTEMI è rappresentato dal TIMI risk score, derivato dalla popolazione dello studio TIMI-11b che ha arruolato pazienti piuttosto selezionati, escludendo i pazienti con segni di insufficienza ventricolare sinistra.

L'importanza della scelta dell'*endpoint* per lo sviluppo del modello di RSS è testimoniata proprio dalla capacità di "percepire" il grado di rischio di questi tre modelli. L'abilità di predire la mortalità è eccellente per il PURSUIT (*C-statistic* 0,81) e per il GUSTO IIb (*C-statistic* 0,84), mentre è stata solo discreta per il TIMI risk score (*C-statistic* 0,78 per il gruppo randomizzato ad eparina non frazionata e 0,72 per il gruppo che ha ricevuto Clexane). Già la capacità di predire *l'endpoint* composito di morte e (re)IMA non-fatale è stata inferiore in tutti i trials (ad esempio per il PURSUI un *C-statistic* di 0,67), mentre ancora inferiore ed ai limiti dell'utilità clinica è stata la capacità predittiva del triplo *endpoint* composito che includesse anche l'ischemia ricorrente e/o la rivascolarizzazione urgente. Ciò a testimoniare le difficoltà e la sfida inerenti la capacità di predire gli eventi ischemici non fatali.

Il più recente RSS sviluppato deriva dal registro internazionale GRACE condotto in 94 ospedali di differente grandezza con un ampio *range* di disponibilità di tecniche di cardiologia interventistica. Accanto ai classici predittori l'RSS del GRACE include anche alcuni nuovi potenti indicatori come l'insufficienza renale e la presenza/assenza di rivascolarizzazione intraospedaliera. Il *C-statistic* per un follow up a 6 mesi della

Corrispondenza a: Dr. Filippo Ottani, Fondazione cardiologica "MZ Sacco", Forlì- Italia.
E-mail: ottanif@omf.dsnet.it

mortalità è risultato di 0,81 nella coorte di derivazione (15.000 pazienti) e 0,75 nella corte di validazione (7.000 pazienti) e mediamente di 0,77 sui 22.000 pazienti valutati.

Al contrario degli studi precedentemente citati, lo studio GRACE è un registro che include pazienti senza la restrizione esercitata dai criteri di inclusione/esclusione classici dei trials clinici. La popolazione derivata, pertanto, riflette ampiamente il paziente del mondo reale. Nello studio TIMI 11b ad esempio i criteri di esclusione prevedevano i pazienti con rivascolarizzazione pianificata entro 24 ore o i pazienti con insufficienza ventricolare sinistra. Inoltre, in quest'ultimo studio, a causa del trattamento, nel braccio attivo con Clexane, venivano esclusi i pazienti con elevata probabilità a priori di sanguinamento; ciò potrebbe estendersi ai molti pazienti con insufficienza renale. Lo studio PURSUIT ha escluso i pazienti con insufficienza renale. Pur con queste variazioni importanti i dati crudi di *C-statistic* non appaiono molto differenti tra i vari RSS, almeno per quanto riguarda l'evento più "pesante" ovvero la morte.

Tuttavia un recente studio di de Araujo Gonçalvels *et al*. ha comparato, in una coorte di 460 pazienti non selezionati con NSTEMI, la capacità discriminativa di 3 differenti RRS (PURSUIT, TIMI e GRACE risk score) per pazienti NSTEMI nel predire la morte/(re)IMA non fatale a 30 giorno ed 1 anno.

A 30 giorni l'RSS del PURSUIT ha mostrato un *C-statistic* di 0,615 (95% *Confidence Interval* (CI): 0,569-0,660), ed il GRACE di 0,672 (95% CI: 0,627-0,714), mentre assai modesta è stata la prestazione esibita dal TIMI risk score 0,551 (95% CI: 0,504-0,597). Il GRACE è risultato statisticamente superiore rispetto al TIMI (delta 0,121; p=0,054). Ad 1 anno di follow-up la performance del GRACE è ulteriormente migliorata (*C-statistic* 0,715; 95% CI: 0,672-0,756) mentre sostanzialmente sovrapponibili sono rimaste le performance degli RSS del PURSUIT (*C-statistic* 0,630; 95% CI: 0,584-0,674) e del TIMI (*C-statistic* 0,585; 95% CI: 0,539-0,631). La superiore prestazione del GRACE è divenuta quindi anche formalmente statisticamente superiore rispetto al TIMI (delta 0,130; p=0,004).

Gli RSS sono importanti per informare dello stato della loro salute tanto i pazienti stessi quanto i parenti, ma sono anche importanti se il loro valore è in grado di guidare scelte terapeutiche appropriate.

de Araujo Gonçalvels *et al*. hanno trovato che il GRACE ed il PURSUIT RSS interagivano, in modo positivo, con l'esecuzione di rivascolarizzazione. Questo nuovo dato osservazionale è estremamente interessante poiché la rivascolarizzazione è eseguita spesso in pazienti che vengono classificati a basso rischio, esponendo maggiormente questi ultimi ai rischi intrinseci della metodica rispetto ai i pazienti ad alto rischio, i quali probabilmente beneficiano in modo assai più marcato della rivascolarizzazione percutanea.

In conclusione, la valutazione dinamica del rischio condotta in modo seriale e continuo può identificare i pazienti che necessitano di terapia più aggressive. Valutazione seriate dell'NT-pro BNP nelle prime 48-72 ore ha recentemente dimostrato la capacità di aiutare a stratificare ulteriormente i pazienti con NSTEMI, infatti coloro che presentano valori in rapido calo rispetto al valore basale mostrano una prognosi migliore.

Non c'è al momento attuale una adeguata valutazione di come nella realtà clinica venga stratificato il rischio. IL TIMI risk score è di gran lunga il più usato degli RSS, anche in Italia, come suggerisce la recente analisi del BLITZ-2. E' tuttavia evidente che i modelli di RSS vanno ancora rifiniti attentamente, magari inserendo una valutazione biochimica multiparametrica.

L'attività di valutazione del rischio è sicuramente fondamentale, ma, come in altre aree della medicina, la sfida più grande per i clinici è utilizzare correttamente i mezzi già in loro possesso.

Bibliografia

1. de Araujo Gonçalves P, Ferreira J, Aguiar C, Seabra-Gomes R. TIMI, PURSUIT, and GRACE risk score: sustained prognostic value and interaction with revascularization in NSTE-ACS. Eur Heart J 2005; 26:865-72.
2. Eagle KA, Lim MJ, Dabbous OH, Pieper KS, Goldberg RJ, Van de Werf F, *et al*. A validated prediction model for all forms of acute coronary syndromes. JAMA 2004; 291:2727-33.
3. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982; 247:2543-6.