

che ha dimostrato la capacità di discriminare pazienti con diversa attività di malattia. I nostri risultati suggeriscono l'impiego routinario nella pratica clinica di questo indice di facile e rapido utilizzo.

Diversi sono stati inoltre i momenti di confronto con altri colleghi di tutta Europa e per questo ringrazio il Gruppo LES che mi ha dato l'opportunità di esserci, di partecipare, di conoscere, di imparare e di approfondire. Come spesso si dice, la medicina non è una scienza esatta, ma una cosa certa che mi è stata insegnata è che un bravo medico è prima di tutto un uomo di scienza, ossia un uomo che studia, che approfondisce e che ha quella costante curiosità che lo spinge a voler sapere sempre di più.

**Laura Massaro**  
laura.massaro84@gmail.com

---

**RICERCA:  
SCADENZA BANDO 28 NOVEMBRE 2009**

---

**Finanziamento di due progetti per la ricerca e cura del LES dell'importo di 30.000 euro**

Sandra D'Alfonso, Nadia Barizzone  
Dip. di Scienze della Salute, Univ. del Piemonte Orientale A. Avogadro, Novara<sup>1</sup>

*Progetto pubblicato su ICARO 60, Maggio 2010*  
*Relazione intermedia pubblicata su ICARO 63, Maggio 201*  
*Relazione conclusiva*

**“Ricerca di geni coinvolti nella suscettibilità al Lupus Eritematoso Sistemico e valutazione delle loro possibili implicazioni cliniche attraverso un algoritmo per la stima del rischio genetico” – Resp. Prof.ssa Sandra D'Alfonso**

**Elenco delle Unità di Ricerca:**

- 1) Centro di Riferimento per le Malattie Autoimmuni Sistemiche, Università di Milano & Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, Milano Dipartimento Medicina Interna - *Coordinatore Scientifico: Prof.ssa. R. Scorza*
- 2) Sezione di Clinica Medica, Università Politecnica delle Marche & Ospedali Riuniti, Ancona - *Coordinatore Scientifico: Prof.ssa M.G. Danieli*
- 3) Azienda Ospedaliera San Camillo Forlanini, Roma, UOC di Reumatologia - *Coordinatore Scientifico: Prof. G.D. Sebastiani*
- 4) Struttura Semplice di Diagnosi e Cura delle Malattie Autoimmuni Sistemiche Università di Siena, Reumatologia - *Coordinatore Scientifico: Prof. M. Galeazzi*
- 5) Univ. Vita-Salute San Raffaele, Milano - *Coordinatore Scientifico: Prof.ssa M.G. Sabbadini*
- 6) Il Università di Napoli, UO di Reumatologia, *Coordinatore Scientifico: Prof. G. La Montagna*

---

<sup>1</sup> Corrispondenza: Sandra D'Alfonso, via Solaroli 17, 28100 Novara - Telefono numero: 0321 660684 - Fax numero: 0321 620421 - E-mail: dalfonso@med.unipmn.it

**Introduzione:**

A oggi i diversi studi di associazione genetica, basati sull'approccio del gene candidato o di tipo GWA (Genome Wide Association), hanno identificato almeno 30 geni la cui associazione con il lupus eritematoso sistemico (LES) è stata confermata in due o più studi indipendenti ed è pertanto ritenuta attendibile. Il nostro gruppo di ricerca ha contribuito all'identificazione di molti di questi geni sia nell'ambito di progetti italiani che grazie alla collaborazione con i principali gruppi di ricerca internazionali che si occupano della genetica del LES.

L'identificazione di questi geni è sicuramente assai importante per comprendere i meccanismi patogenetici alla base dello sviluppo della malattia. Tuttavia queste osservazioni fino a questo momento hanno avuto uno scarso impatto nella pratica clinica. La presenza di ciascuno dei singoli alleli di suscettibilità identificati conferisce solamente un modesto rischio rispetto alla sua assenza per cui preso singolarmente nessuno dei loci di suscettibilità al LES ad oggi identificato può avere un valore predittivo-diagnostico.

Tuttavia è possibile che la valutazione dell'effetto combinato dei diversi geni, mediante l'analisi simultanea delle diverse varianti associate, possa aumentarne notevolmente il potere predittivo. Questo approccio è stato applicato con successo in altre patologie complesse quali malattie metaboliche cardiovascolari e sclerosi multipla. I risultati finora sono promettenti: considerare simultaneamente i diversi fattori genetici aumenta significativamente il potere predittivo della malattia sia rispetto all'utilizzo dei singoli geni sia rispetto all'analisi dei soli fattori di rischio non genetici (Yang 2003, Yang 2005, Kathiresan 2008, Lyssenko 2008, De Jager 2009). In questo studio abbiamo utilizzato questo metodo per il LES.

**Risultati**

Allo scopo di definire un algoritmo che permetta di calcolare l'indice di rischio genetico (weighted genetic risk score - wGRS) per il LES valutando l'effetto combinato dei diversi geni associati abbiamo selezionato marcatori di tipo SNP (polimorfismi di un singolo nucleotide) situati nei geni associati alla malattia sulla base di dati della letteratura replicati in più studi diversi. Per ogni gene abbiamo selezionato i marcatori più fortemente associati con la suscettibilità al LES.

L'algoritmo è stato calcolato come descritto da De Jager (2009), correggendo ciascun marcatore associato per il suo "peso", calcolato sulla base degli Odd Ratio (OR) riportati in letteratura su ampie casistiche (alcune migliaia d'individui) totalmente indipendenti dalla nostra, e replicate almeno in una seconda casistica di pazienti e controlli.

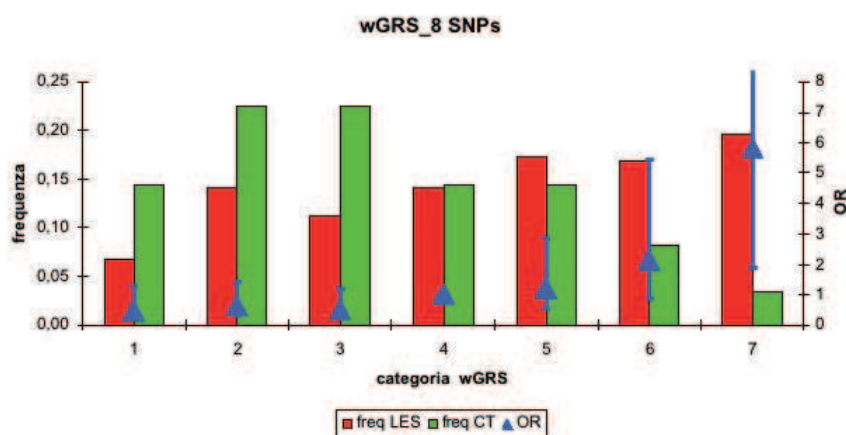
Nella prima fase del nostro studio (descritta nella prima relazione) abbiamo eseguito un'analisi preliminare utilizzando 15 marcatori, localizzati in 8 diversi loci di suscettibilità. Nell'anno successivo abbiamo in un primo tempo rifinito le modalità di analisi, decidendo di limitarci ad utilizzare esclusivamente un polimorfismo per ciascun locus genetico. In tal modo tutti i marcatori inclusi nell'algoritmo sono completamente indipendenti l'uno dall'altro, e questo permette di eliminare possibili effetti spuri dovuti alla presenza di Linkage Disequilibrium tra le varianti genetiche utilizzate. Inoltre abbiamo ricalcolato gli OR per i polimorfismi selezionati sulla base dei nuovi studi di associazione pubblicati nel frattempo allo scopo di definire con più precisione i "pesi" dei singoli polimorfismi.

Abbiamo pertanto calcolato un nuovo algoritmo utilizzando 8 marcatori indipendenti su una casistica di 178 casi e 173 controlli. Il valore medio di wGRS era significativamente più elevato ( $p < 0,0001$ , test di Student) nei pazienti (2,26) rispetto ai sani (1,84): questo dato

forniva una prima indicazione che l'insieme di questi marcatori è associato alla suscettibilità alla malattia anche nella nostra casistica. Per valutare più in dettaglio come i valori di wGRS così calcolati siano diversamente distribuiti nei pazienti e nei controlli sani, abbiamo suddiviso l'intera casistica in 7 categorie progressive di wGRS (da valori bassi a elevati). Come si può osservare dagli istogrammi nella **fig. 1**, la frequenza dei pazienti è nettamente inferiore a quella dei controlli nelle categorie con bassi valori di wGRS (1-3), ed aumenta progressivamente rispetto alla frequenza dei sani nelle categorie con valori di wGRS più elevato. Abbiamo valutato l'associazione tra i valori di wGRS e il rischio di malattia confrontando le varie categorie con quella mediana (cat. 4) che abbiamo considerato come riferimento, calcolando gli OR e le relative significatività statistiche. I triangoli blu nella **fig. 1** rappresentano l'andamento dell'OR nelle 7 categorie rispetto alla categoria mediana: come si può osservare l'OR presenta un aumento progressivo raggiungendo il valore più elevato per la categoria con wGRS massimo (cat. 7 vs. cat. 4: OR=5,83,  $p=0,0000023$ ). Un risultato simile si otteneva dal confronto tra le categorie estreme (cat. 6+7 vs. cat. 1+2: OR=5,62  $p=0,0000001$ ). Tali valori di rischio risultavano chiaramente superiori rispetto al rischio conferito da ciascuno dei singoli marcatori inclusi nel calcolo del wGRS (OR compresi tra 1.2 e 2.3) confermando quindi la maggiore informatività risultante dal considerare il contributo complessivo di diversi marcatori genetici.

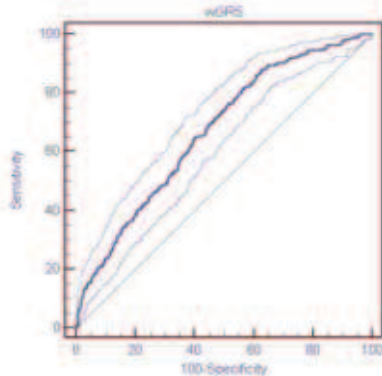
Per valutare il potere predittivo di questo algoritmo preliminare abbiamo tracciato una curva ROC (Receiver Operating Characteristic), che mette in relazione la sensibilità di una metodica con l'inverso della sua specificità, ed abbiamo calcolato l'area al di sotto della curva (AUC). L'AUC può essere ritenuta la probabilità con cui il nostro algoritmo distinguerebbe i pazienti dai controlli sani. Come si può osservare dalla **fig. 2**, la curva ROC si discosta significativamente ( $p<0,0001$ ) dalla diagonale (valore nullo: AUC=0,500 corri-

**Fig. 1** Distribuzione dei pazienti LES e controlli (CT) nelle varie categorie di rischio calcolate sulla base dei valori dell' algoritmo wGRS ottenuto con 8 SNPs

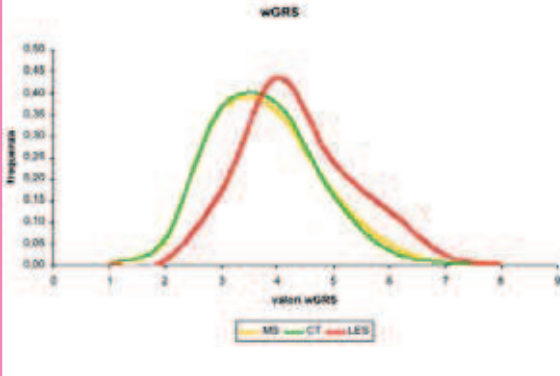


**Nota alla Fig. 1:** Gli istogrammi rappresentano le frequenze di LES e CT nelle 7 categorie del wGRS (dalla categoria 1: rischio minimo, alla 7: rischio massimo). I triangoli blu rappresentano i valori di OR relativi al confronto tra le varie categorie con quella di riferimento (categoria 4).

**Fig. 2** Curva ROC per l'algoritmo wGRS calcolato con 8 SNPs. La diagonale rappresenta il riferimento (assenza di potere predittivo). E' rappresentata la curva ROC, a tratto più marcato, e i suoi intervalli di confidenza



**Fig. 3** Grafico delle frequenze del wGRS in pazienti LES e SM e controlli sani, algoritmo con 17 SNPs.



spondente ad una probabilità del 50% di classificare correttamente come paziente o controllo un individuo e quindi ad una mancanza totale di potere predittivo), anche se i valori delle AUC non sono molto elevati (AUC=0,639, corrispondente ad una probabilità del 63,9% di classificare correttamente come paziente o controllo un individuo qualsiasi). Nella fase successiva, ampliando il numero di polimorfismi e la casistica di pazienti e controlli, abbiamo progettato un nuovo algoritmo, basato su 17 marcatori indipendenti, comprendenti quelli già utilizzati nel primo wGRS, e lo abbiamo testato e calcolato su una seconda casistica di 278 pazienti e 1541 controlli, osservando una distribuzione di casi e controlli nelle 7 categorie di rischio simile a quella già osservata per l'algoritmo con soli 8 SNPs (cat. 6+7 vs. cat. 1+2: OR=6,66,  $p < 1 \times 10^{-7}$ ). Abbiamo valutato anche in questo caso le curve ROC: l'introduzione di nuovi SNPs migliora il potere predittivo dell'algoritmo: AUC=0,689 (17 SNPs) vs. AUC=0,639 (8 SNPs), anche se la differenza non è statisticamente significativa ( $p=0,13$ ).

I valori di AUC da noi ottenuti sono nell'ordine di grandezza di quelli comunemente osservati con questo tipo di studi per le malattie multifattoriali.

Infine, per testare la specificità del nostro algoritmo, abbiamo deciso di analizzarlo anche su un'altra patologia autoimmune clinicamente distinta dal LES, la sclerosi multipla (SM), analizzando 737 pazienti affetti da SM disponibili grazie ad un progetto parallelo del nostro gruppo. Come si può osservare dalla **fig. 3**, che rappresenta l'andamento del wGRS nelle 3 casistiche con una rappresentazione in forma di curva analoga a quella a istogramma della **fig. 1**, la curva relativa ai pazienti SM è praticamente sovrapponibile a quella dei controlli sani. Inoltre le curve ROC relative ai pazienti SM non si discostano dalla diagonale (AUC=0,512,  $p=0,34$ ) e questo indica un'assenza totale di potere predittivo. Questo dato ci dimostra che l'algoritmo da noi disegnato è specifico per il LES.

*Quale può essere l'utilità pratica di questo algoritmo?*

Nella sua forma attuale l'algoritmo non può avere una valenza diagnostica in quanto il suo potere predittivo è troppo basso (68.9%) per essere utilizzato come test diagnostico: il miglior compromesso tra i migliori risultati di sensibilità e specificità si ottiene con un wGRS superiore a 3,45 che tuttavia consente di raggiungere una sensibilità e specificità di circa il 61,5%. Tuttavia questo algoritmo potrebbe essere utile per individuare individui a rischio di sviluppare la malattia eventualmente selezionabili per essere sottoposti ad interventi preventivi. Analizzando con maggior dettaglio (di cui forniamo di seguito solo un riassunto) la sensibilità e specificità di questo test si può osservare che selezionando individui con valori molto alti di wGRS ( $> 4,98$ ) avremo una specificità elevata (96%) anche se una sensibilità bassa in quanto solo il 15% dei malati analizzati in questo studio presenta questo valore e quindi si può ipotizzare che in una futura situazione (es. un nuovo protocollo clinico sperimentale) solo questa percentuale di individui verrebbe catalogata come individuo ad alto rischio e quindi selezionato per interventi successivi. Questo tipo di selezione potrebbe essere utile ad esempio qualora fossero disponibili interventi preventivi impegnativi o rischiosi da adottare quindi solo per individui con un'alta probabilità di ammalarsi: analizzare con wGRS, un test a costo modesto, individui *potenzialmente* a rischio di malattia (es. parenti di un paziente) permetterebbe di selezionare le persone *realmente* a rischio elevato di sviluppare la malattia, riservando quindi solo a queste ultime un possibile trattamento preventivo che potrebbe essere costoso o comportare notevoli effetti collaterali.

### Conclusioni e prospettive future

In questo studio è stato costruito un algoritmo di rischio genetico per il LES comprendente marcatori genetici che si sono rivelati specifici per il LES. L'algoritmo ha dimostrato un discreto potere predittivo, analogo a quello osservato per altre malattie multifattoriali e nonostante, almeno nella presente forma, non possa essere utilizzabile per diagnosi, potrebbe essere utilizzato, eventualmente integrandolo con altre informazioni cliniche o ambientali, per identificare individui ad elevato rischio di sviluppare il LES in opportuni e controllati protocolli clinici.

L'aggiunta di marcatori genetici aumenta il potere predittivo ma solo moderatamente: da una predittività del 63.9% (AUC=0,639) con l'algoritmo costruito con 8 marcatori passiamo a un valore del 68.9% (AUC=0,689) con un numero superiore al doppio di marcatori (17 SNPs). Questo moderato aumento potrebbe essere motivato dalla natura della malattia, multigenica e multifattoriale, per cui in ogni paziente potremmo osservare l'influenza solo di un sottogruppo di tutti i geni di suscettibilità identificati nella totalità dei pazienti, e dal peso piuttosto modesto dei singoli fattori genetici finora identificati costituiti generalmente da varianti genetiche frequenti nella popolazione (frequenza  $>5-10\%$ ). A questo proposito potrebbe essere interessante valutare nell'immediato futuro il peso delle varianti rare e l'integrazione di queste nell'algoritmo. L'analisi delle varianti rare in tutto il genoma costituirà la prossima sfida della genetica delle malattie multifattoriali. In questo progetto abbiamo iniziato a valutare il peso delle varianti rare risequenziando alcuni singoli geni candidati. I risultati, sottomessi per pubblicazione hanno riguardato due geni *TREX1*, che codifica per una esonucleasi coinvolta nella risposta agli stress ossidativi e nell'apoptosi e *PRFI* che codifica per perforina. I nostri risultati hanno identificato 13 varianti rare nel gene *PRFI* (7 missenso e 6 sinonime), evidenziate ciascuna in 1-6 pazienti affetti da LES tra i 145 sequenziati per l'intero gene, la cui frequenza complessivamente risultava

significativamente maggiore nei pazienti rispetto alla popolazione di controllo (0.12 vs 0.084  $p=0.0038$ ,  $OR=1.89$ ,  $95\%CI=1.22-2.94$ ). Abbiamo inoltre osservato che i pazienti portatori di due di queste varianti missense mostrano un ridotto numero di cellule che esprimono marcatori delle cellule NK (CD16 e perforina) rispetto a individui negativi per queste mutazioni.

Al contrario non confermiamo il ruolo di varianti rare di *TREX1* nella suscettibilità al LES dopo aver sequenziato completamente il gene in 210 pazienti.

Riassumendo, si è messo a punto un algoritmo predittivo con risultati promettenti che per poter essere utilizzato nella pratica clinica necessita di ulteriori sviluppi, quali l'integrazioni con varianti genetiche rare, con fattori di esposizione ambientale e con caratteristiche cliniche oltre che ulteriori test di specificità con altre malattie autoimmuni con caratteristiche cliniche più simili al LES.

Un algoritmo integrato in grado di stimare il rischio del LES ha importanti ricadute in quanto permette di predire il rischio di malattia e quindi di identificare individui ad alto rischio di sviluppare la malattia. In particolare, essi potranno essere individui con alcuni segni clinici di malattia che non soddisfano ancora i criteri diagnostici oppure i parenti di primo grado dei pazienti che hanno una probabilità di circa 30 volte di sviluppare il LES rispetto alla popolazione generale. A questo proposito stiamo completando uno studio in cui analizziamo il wGRS in 42 famiglie con uno o due pazienti LES e parenti di primo grado sani. Una volta identificati, gli individui ad alto rischio di sviluppare la malattia potranno essere: a) seguiti clinicamente in maniera accurata; b) sottoposti a trattamenti ai primi sintomi; c) sottoposti a misure preventive es. eliminazione di fattori di rischio non genetici quali esposizione a fattori ambientali.

PARTECIPATE ALLA STESURA DI ICARO.  
INVIATECI LE VOSTRE STORIE E VOSTRI SUGGERIMENTI.

SE DESIDERATE CHE LE VOSTRE GENERALITÀ,  
IL VOSTRO INDIRIZZO, NUMERO DI TELEFONO  
E RECAPITO E-MAIL SIANO PUBBLICATE SU ICARO,  
È NECESSARIO CHE CI AUTORIZZIATE ESPRESSAMENTE,  
CON DICHIARAZIONE RESA IN TAL SENSO IN CALCE ALLA  
VOSTRA LETTERA.

