

Schneider Chafen JJ, Newberry SJ, Riedl MA, Bravata DM, Maglione M, Suttorp MJ, Sundaram V, Paige NM, Towfigh A, Hulley BJ, Shekelle PG.
Diagnosing and Managing Common Food Allergies: A Systematic Review *JAMA* 2010;303(18):1848-1856
<http://jama.ama-assn.org/cgi/content/full/303/18/1848>

Nell'anno 2007 i National Institutes of Allergy and Infectious Diseases (NIAID), l'American Academy of Allergy Asthma and Immunology (AAAAI) ed il Food Allergy and Anaphylaxis Network (FAAN), sotto i diretti auspici del governo degli Stati Uniti d'America, stabilirono che fosse venuto il momento per la produzione Linee Guida cliniche, evidence-based, sull'Allergia Alimentare (AA). Per far ciò fu deciso di coinvolgere l'Evidence-based Practice Center della South California, Palo Alto, nella produzione di una Revisione Sistemática onnicomprensiva, sugli argomenti definizione, prevalenza, diagnosi, trattamento e prevenzione dell'AA. L'EPC della South California è uno dei circa dodici centri EPC sparsi negli Stati Uniti e in Canada, designati dall'Agenzia per la ricerca e la qualità in campo sanitario [Agency for Healthcare Research and Quality (AHRQ)]; esso fa parte della RAND Corporation, un'istituzione no-profit, di origini post-belliche (il periodo era quello della guerra fredda) che dalla sua nascita si è occupata di fornire ai vari settori tecnologici dello stato americano, sanità compresa, gli strumenti adatti a migliorare le proprie linee programmatiche e le proprie pratiche decisionali (< ... health, education, national security, international affairs, law and business, the environment, and more >).

Tutto ciò RAND porta avanti grazie al lavoro di numerosi staff di ricerca e analisi composti dalle migliori, così dicono, intelligenze e professionalità a livello mondiale.

La Revisione Sistemática (RS) che andiamo ora ad analizzare *brevemente* (1) non è altro che il riassunto, estremamente sintetico, di un voluminoso Evidence Report costituito da 416 pagine, 12 figure, 55 tavole e 12 appendici metodologiche (2), prodotto dai medesimi autori presso l'EPC di Palo Alto. Esso è anche accessibile gratuitamente in rete, per chi avesse voglia di verificare in prima persona le modalità di lavoro dell'équipe californiana.

Non deve stupire dunque il fatto che fra gli autori del lavoro sia possibile trovare un solo *content expert* (Marc Riedl), mentre il resto del panel risulta costituito da epidemiologi, statistici, esperti in EBM, policy makers, policy analysts, programmatori e assistenti dello staff. E' indubbio che, nel confezionare la RS si sia adottato un sistema altamente tecnicistico, a scapito probabilmente della stessa esaustività della trattazione: sono stati infatti esclusi, come vedremo, importanti e nuovi filoni di ricerca scientifica e di pratica clinica in ambito allergologico per il solo fatto che la letteratura ad essi inerente non rispecchiava i rigidi criteri di inclusione previsti dal panel nella fase organizzativa della revisione. Tali criteri erano infatti orientati alla selezione dei soli lavori in grado di rispondere alle Key Questions preliminarmente individuate, per ognuno degli ambiti della RS (diagnosi, terapia e prevenzione, in particolare).

Per valutare la qualità globale della RS mi sono servito di AMSTAR (3), uno strumento validato per la valutazione di Revisioni Sistemáticas. Qualche approssimazione è stata necessaria per il fatto che Schneider Chafen e colleghi hanno revisionato argomenti disparati, dalla prevalenza alla diagnosi, dalla terapia alla prevenzione, per i quali sono differenti i metodi di selezione dei lavori e di accorpamento dei risultati. Nelle 11 domande chiave di cui è costituito AMSTAR, la RS di Schneider Chafen ha collezionato il "SI" (le possibili risposte ai vari items erano: sì, no, non si può rispondere, non applicabile) in almeno 8 casi, come è possibile vedere nella seguente Tabella.

AMSTAR criteria (applicati alla RS di Schneider Chafen et al. 2010)

1. Was an 'a priori' design provided?	SI'
2. Was there duplicate study selection and data extraction?	SI'
3. Was a comprehensive literature search performed?	SI'

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	NO
5. Was a list of studies (included and excluded) provided?	SI'
6. Were the characteristics of the included studies provided?	SI'
7. Was the scientific quality of the included studies assessed and documented?	SI'
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	SI'
9. Were the methods used to combine the findings of studies appropriate?	SI'
10. Was the likelihood of publication bias assessed?	Non si può rispondere
11. Was the conflict of interest stated?	NO

Note critiche

A. Le limitazioni del disegno iniziale della RS erano piuttosto rigide (*item 1*): per lo studio della prevalenza potevano essere inclusi esclusivamente lavori basati su campioni di popolazioni; per gli studi diagnostici soltanto gli studi prospettici su ampie popolazioni in cui il test in oggetto fosse comparato con il *gold-standard* (Test di Provocazione Orale) e dai quali fossero desumibili sufficienti dati per il calcolo dei parametri di sensibilità e specificità; per gli studi di intervento, preventivo o terapeutico che fosse, venivano inclusi soltanto i trials controllati, anche se non RCT, e le revisioni sistematiche.

B. Furono discutibilmente esclusi a priori due importanti databases della letteratura primaria, Embase e Cinhal (*item 3*), così come furono esclusi lavori scritti in lingua diversa dall'inglese. La limitazione temporale a partire dal 1988, e fino al Settembre 2009, è probabilmente giustificata.

C. Per la valutazione della qualità degli RCT sono stati considerati ancora validi i criteri di Jadad, che non comprendono, ad esempio, la verifica del mascheramento delle liste di randomizzazione! (*item 7*). Proseguendo però l'analisi dell'Evidence Report, risulta che le prove di efficacia ricavate dai lavori inclusi venivano classificate anche secondo il più moderno sistema del GRADE (Grading quality of evidence and strength of recommendations). Questo sistema opera una classificazione qualitativa dei livelli di evidenza basandosi non tanto sulle caratteristiche formali del singolo studio da cui la prova nasce, ma piuttosto basandosi sulla maggiore o minore qualità dell'evidenza per ogni singolo outcome, in modo relativamente indipendente dalla "robustezza" dell'intero disegno dello studio. Le squisite potenzialità di questa metodica hanno poi finito per riversarsi sulle raccomandazioni contenute nelle Linee Guida ACAAI-NIAID, di cui la RS californiana costituisce la *evidence base* (4). In queste Linee Guida infatti, ancora in via di pubblicazione ma già consultabili sotto forma di draft, la forza delle varie raccomandazioni risulta direttamente proporzionale e collegata ad un livello di evidenza, specifico per quella singola Key Question, alla quale la raccomandazione fornisce la risposta. Si tratta, non vi è dubbio, di un punto a favore della qualità di dette Linee Guida.

D. La stragrande maggioranza dei risultati non è stata aggregata (*item 9*), tant'è che di ogni lavoro incluso nell'analisi finale è stato presentato un riassunto in forma discorsiva che rende difficile focalizzare i risultati e capire le differenze fra quelli di uno studio e quelli di un altro studio. Infatti nella RS non è presente nessuna metanalisi, soltanto alcune Receiver Operating Curves, o ROC (**vedi Box 1**), sull'efficacia diagnostica degli SPT e delle IgEs confrontata con l'efficacia del TPO.

E. Per quanto ho potuto verificare, in nessuna delle 416 pagine del Working Report, e così nemmeno nell'articolo originale di JAMA e nel suo supplemento in rete, sono esplicitati i potenziali conflitti di interesse (*item 11*). L'appartenenza alla RAND costituisce sufficiente documentazione a questo riguardo?.

Luci ed ombre dunque ... come in ogni Revisione Sistemática.

BOX 1

Curva ROC (Receiver Operating Curve): è un insieme di punti, iscritti su di un sistema cartesiano (in modo empirico oppure attraverso una funzione matematica, detta modello). La curva ottenuta permette di descrivere la variazione delle

caratteristiche intrinseche di un test diagnostico (sensibilità, specificità, rapporti di verosimiglianza) al variare della soglia di positività del test stesso, ovvero del cut-off individuato. Per ogni cut-off si individua un punto sul quadrante delimitato dai due assi cartesiani: le coordinate di ognuno di questi punti sono date dal tasso di veri positivi ($y =$ sensibilità) e dal tasso di falsi positivi ($x = 1 -$ specificità).

L'area posta al di sotto della curva ROC esprime il *potere diagnostico* del test. La peggior curva possibile è quella che giace tutta sulla diagonale principale (potere informativo del test nullo, area sotto la curva = 0,50). Il test è tanto migliore quanto più la ROC è spostata verso nord-ovest; il test perfetto ha una ROC che coincide con i due lati (quello a sinistra e quello in alto: potere informativo massimo, assenza di errori, area sotto la curva = 1).

La pendenza della curva in un punto ci informa sulla dinamica dei rapporti fra sensibilità e specificità per quel cut-off: se la pendenza è elevata significa che c'è un buon incremento della sensibilità a fronte di una piccola perdita di specificità.

Riassunto e qualche commento sui risultati principali della RS

Dopo aver individuato 12378 lavori attraverso strategie di ricerca multiple su 4 databases (PubMed, Cochrane Database of Systematic Reviews, Cochrane Database of Abstracts of Reviews of Effects, and Cochrane Central Register of Controlled Trials), gli Autori hanno analizzato il full-text in 1216 casi, includendo nella loro revisione soltanto 182 studi. I risultati pubblicati deriveranno poi dai 72 studi sugli alimenti di maggior interesse, ovvero latte, uova, arachidi, noci, pesce, molluschi e crostacei (1 metanalisi su incidenza e prevalenza, 18 studi sulla diagnosi, 28 studi e 4 metanalisi sulla terapia, 21 studi di prevenzione).

A. I risultati relativi alla **prevalenza** sono inficiati da una grande variabilità nella definizione diagnostica di Allergia Alimentare all'interno dei vari lavori inclusi nelle analisi. Grandi risultano perciò le differenze nei risultati a seconda che si fossero tenute in conto le diagnosi *self-made* oppure le diagnosi ottenute attraverso metodiche di laboratorio. Chi fosse interessato potrà leggerli nel lavoro originale.

B. Nell'ambito della **diagnosi**, Schneider Chafen e colleghi hanno analizzato le performance di SPT (13 studi), IgEs (11 studi) e Patch-test (8 studi), nel confronto con un test di scatenamento orale. Gli autori, con l'ausilio di metodi statistici molto complessi, sono stati in grado di accorpate i dati per la costruzione delle curve ROC solamente nei casi di SPT, IgEs e Patch-test per il latte e nei casi di SPT e IgEs per l'uovo. *Le aree sotto la curva ROC* oscillano fra 0.78 e 0.87 (fuorchè quella dei Patch-test per il latte, che risulta di molto inferiore). Un potere diagnostico da ritenersi modesto. Inoltre, le ROC costruite per comparare l'efficacia diagnostica degli SPT nei confronti dell'efficacia diagnostica delle IgEs (**vedi Figura**), hanno dimostrato che non vi sono differenze significative fra i due test nella diagnosi di allergia verso l'insieme di tutti i tipi di alimenti, così come verso l'uovo e il latte presi singolarmente.

Figure 5: ROC curve for SPT and sIgE for Hen's egg

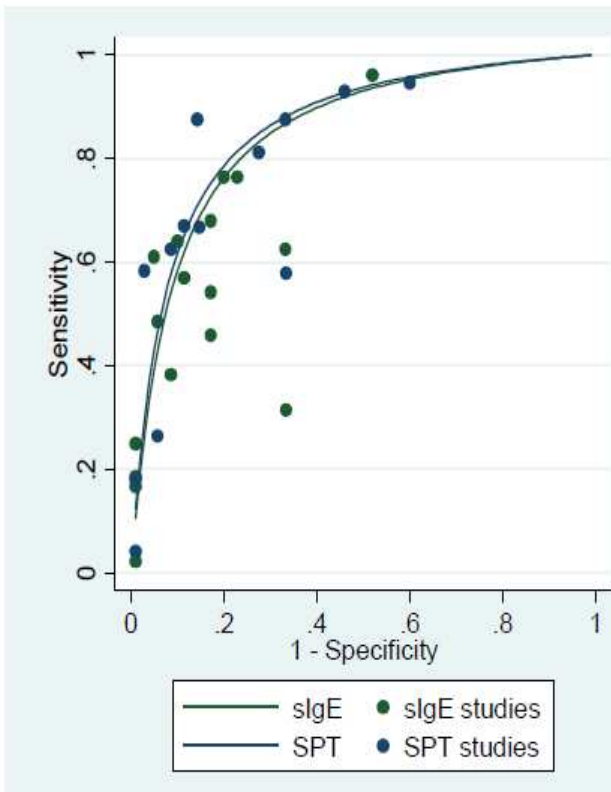
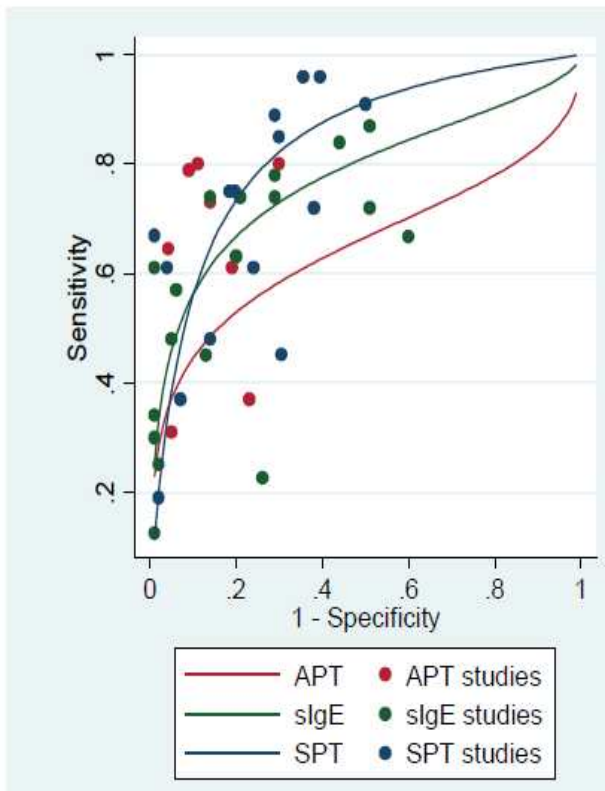


Figure 4: ROC curve for SPT, sIgE, APT for Cow's Milk



Gli autori stressano il concetto della scarsa riproducibilità delle metodiche diagnostiche, causata dalla scarsa standardizzazione dei materiali con cui vengono eseguiti i test cutanei e dell'altrettanto

scarsa standardizzazione dei criteri per definire la positività dei test; non fa eccezione in questo lo stesso Test di Provocazione Orale, un Gold Standard ... in realtà poco standardizzato ... oltre tutto criticato per il fatto stesso di essere considerato tale. Schneider Chafen ricorda che individui positivi al challenge sono poi risultati in grado di tollerare il cibo incriminato in differenti condizioni ambientali e/o di preparazione dell'alimento.

Singolare, e molto criticabile a mio avviso, appare il modo con cui gli autori, dopo aver ricordato che nella RS potevano essere inclusi soltanto i lavori di comparazione con il Gold Standard, liquidano una metodica diagnostica di provata scientificità quale la Component Resolved Diagnosis (CRD): essi la citano infatti, senza alcun termine di differenziazione, nel calderone dei cosiddetti "altri test per la diagnosi di allergia alimentare", insieme al Test di provocazione/neutralizzazione, alla Kinesiologia applicata, all'Analisi del capello, al Test leucocitotossico etc. (?).

C. Nel campo della **terapia** i revisori californiani riportano risultati alquanto contraddittori, basati su studi metodologicamente deboli o molto deboli:

- è stato ritenuto valido un solo studio, del 1993, in cui una Dieta di Eliminazione risultò efficace nel migliorare il decorso di pazienti con Dermatite Atopica (DA).
- Solo cinque studi, validi, sull'Immunoterapia Orale, dimostratisi efficaci nel desensibilizzare pazienti affetti da Allergia Alimentare ma, come ricordano più volte gli autori, tali interventi non sono ancora stati studiati per quanto riguarda la vera tolleranza, cioè la possibilità che un paziente tolleri un alimento per il quale è stato desensibilizzato, anche dopo un lungo periodo di assenza di quell'alimento dalla sua dieta. Anche la sicurezza della SOTI non è ritenuta sufficientemente dimostrata. [*Trovo davvero poco comprensibile la mancata individuazione, fra i trials potenzialmente eleggibili, del lavoro di Longo e colleghi sulla desensibilizzazione di pazienti con sintomi gravi anche per piccole dosi di alimento (JACI 2008 Feb;121(2):343-7) : esso infatti, oltre a non far parte, pur trattandosi di un RCT, del gruppo dei lavori inclusi nell'analisi finale, non compare in nessuna delle lunghe liste di articoli esclusi con le più diverse motivazioni*].
- Importanti limitazioni di metodo impediscono di trarre conclusioni dagli 8 lavori analizzati sull'efficacia dei latti alternativi al latte di mucca.
- Un lavoro sull'efficacia terapeutica di una dieta ipoallergenica alla madre nutrice è stato selezionato in modo discutibile in quanto aveva come outcome le "coliche gassose", non certo un indicatore d'esito di sicura eziologia allergica.
- Per quanto riguarda la terapia farmacologica, sono inclusi nell'analisi, tre studi, fra cui due italiani (Cavagni 1989 e Businco 1990), che hanno dimostrato l'efficacia di Timomodulina e Cromoglicato sodico associate a diete di eliminazione, sull'andamento della DA in soggetti allergici. Due studi invece hanno dimostrato l'efficacia dell'anticorpo Anti-IgE TNX-901 nel consentire l'ingestione di maggiori quantità di arachidi.
- Due lavori hanno dimostrato l'inefficacia dei Probiotici su DA e sanguinamento rettale in prematuri ... anche in questi casi sembra trattarsi di due outcomes non comunemente attribuibili a causa allergica.
- Un trial educativo ha dimostrato invece la possibilità di ridurre intensità e frequenza delle reazioni acute in soggetti allergici alle arachidi.

D. **Prevenzione** dell'Allergia Alimentare e della Dermatite Atopica.

- Troppo debole la qualità degli unici due lavori sull'efficacia dell'introduzione ritardata dei cibi solidi per poter trarre conclusioni, a dispetto dei risultati positivi
- Una Revisione Sistemica Cochrane abbastanza recente ha negato l'efficacia, nella prevenzione della DA, di diete ristrette alle madri in gravidanza o in allattamento, a differenza di quanto era risultato da due più vecchi RCT sull'argomento.

- Anche le prove di efficacia sul prolungamento dell'allattamento esclusivo al seno nel prevenire lo sviluppo di eczema nei primi anni di vita sono state giudicate conflittuali (tre trial da un'unica popolazione di lattanti seguita prospetticamente).
- L'utilizzo di latti estensivamente o parzialmente idrolizzati sembra essere efficace, in soggetti ad alto rischio di AA, a prevenire lo sviluppo di allergia al latte vaccino; irrilevante invece la capacità di ottenere lo stesso risultato da parte dei latti di soia. Gli Autori ritengono però che nei vari lavori originali fossero scarsamente definiti sia l'intervento (dubbi sulla definizione della profondità dell'idrolisi) sia le caratteristiche della popolazione (dubbi sulla definizione di rischio elevato).
- L'efficacia preventiva di alcuni Probiotici deriva principalmente da trials nei quali non è possibile distinguere l'influenza contemporanea di altri fattori quali l'allattamento al seno e l'allattamento con formule speciali.

Ragionando sull'intero corpus di evidenze da loro ottenuto, Schneider Chafen e colleghi sentenziano che si tende a produrre troppi lavori e che questi prendono in esame una troppo estesa quantità di outcomes; ma il problema più diffuso e importante sembra essere la mancanza di un criterio uniforme per la diagnosi Allergia Alimentare.

Commenti finali

La lettura dell'articolo di recente pubblicato su JAMA è un esercizio riduttivo, in quanto le informazioni in esso presentate non sono che un troppo piccolo, conciso e schematico riassunto di un enorme pubblicazione reperibile online. D'altro canto, la lettura dell'Evidence Report della RAND Organization è lunga, difficile per chi ha poca dimestichezza con le tecniche e con il linguaggio specifico della Evidence Based Medicine, e molto dispersiva alla resa dei conti, ma è questo in realtà l'unico modo per comprendere e apprezzare al meglio il grande lavoro fatto dai revisori. Sono presenti anche in questa RS alcuni elementi di debolezza che ho cercato volta a volta di chiarire. Si tratta tuttavia di un'impresa titanica, portata avanti sulla scorta di un livello di competenze tecniche estremamente elevato. Per verificare appieno l'efficacia di tale modalità di verifica sistematica dei dati scientifici provenienti dalla letteratura non resterà che valutare analiticamente le Linee Guida, in corso di pubblicazione definitiva, di cui la RS di Schneider Chafen costituisce la base conoscitiva. Ma questo ... alla prossima puntata ...

Bibliografia

1. Schneider Chafen JJ, Newberry SJ, Riedl MA, Bravata DM, Maglione M, Suttorp MJ, Sundaram V, Paige NM, Towfigh A, Hulley BJ, Shekelle PG. Diagnosing and Managing Common Food Allergies: A Systematic Review *JAMA* 2010;303(18):1848-1856
2. Shekelle P, Maglione M, Riedl M et al. Food Allergy: Evidence Report. <http://www.rand.org/health/centers/epc/>. Accessed June 6, 2010
3. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR. *BMC Med Res Methodol.* 2007;7:10.