

# Gli intervalli di riferimento

D. Giavarina

Laboratorio di Chimica clinica ed Ematologia, Ospedale San Bortolo, Vicenza

## Riassunto

Gli intervalli di riferimento sono lo strumento maggiormente utilizzato nelle decisioni mediche. Essi sono utilizzati come linee guida, quando non ci sono mezzi migliori per distinguere la condizione di salute dalla malattia. Molto spesso, negli screening di popolazione per una data patologia, si dovrebbe utilizzare un intervallo di riferimento basato sul 95% centrale della popolazione di riferimento ma, se lo scopo degli intervalli di riferimento è quello di aiutare il clinico nell'interpretare i risultati, in popolazioni ospedalizzate dovremmo utilizzare valori diversi da quelli della popolazione sana.

Gli standard e le linee guida del CLSI e dell'IFCC sono utili e corrette se utilizzate nelle condizioni definite, ossia con un campione veramente sano e con un minimo di 120 valori. Quando non si possono rispettare le raccomandazioni del CLSI e dell'IFCC, si dovrebbero utilizzare le "normalizzazioni" dei dati e l'analisi degli *outliers*. Si possono utilizzare anche altri metodi (i metodi robusti o i metodi indiretti) per ottenere una stima corretta dei valori di riferimento.

Tuttavia, non tutti i problemi presenti nella definizione di un livello decisionale possono essere risolti con una procedura per il calcolo degli intervalli di riferimento, anche se corretta e robusta. Occorre considerare altri approcci basati sulla clinica; la sfida per i prossimi anni potrebbe essere quella di abbandonare gli intervalli di riferimento, piuttosto che quella di migliorarne la qualità.

## Summary

### "Reference intervals"

The reference interval is the most widely used medical decision-making tool. Reference intervals are used as guidelines when there are no better methods of distinguishing healthy from diseased states. Most often when screening populations for potential disease, the 95% reference interval is commonly used but, if the purpose of reference intervals is to help the clinician interpret the data, in a hospitalized population, values different from the traditional healthy population may be used. The CLSI and IFCC guidelines are useful under their stated conditions, i.e., a truly healthy sample with a minimum of 120 values. When CLSI and IFCC guidelines are not met, then one must use transformations and outlier detection, and consider the robust method or the indirect methods for computing reference intervals.

Nevertheless, not all the problems in the definition of a decision level can be achieved with a good, correct and robust procedure for the reference interval computations. Different approaches must be considered, based on clinical finding; the challenge for the next years could be to abandon reporting of normal ranges, rather than to improve their quality.

---

Nel 1994 Solberg scriveva: "lo scopo dei valori di riferimento, e dei limiti di riferimento o degli intervalli derivati da loro, è di fornire la base per interpretare i risultati di laboratorio ottenuti durante le valutazioni cliniche".

Gli intervalli di riferimento sono lo strumento più largamente utilizzato nelle decisioni mediche. Svolgo-

no una funzione centrale nel determinare se un soggetto sia o meno sano. Sono usati come linee guida, quando non ci sono metodi migliori per distinguere lo stato di malattia dallo stato di salute.

Nonostante si siano spesi fiumi di parole per sostenere che gli intervalli di riferimento non sono "valori normali", di fatto essi sono, per loro natura, livelli

decisionali e quando sono combinati con altri dati clinici, aiutano a definire lo stato clinico del paziente.

Verso la fine degli anni '80, due organizzazioni professionali, la Federazione Internazionale di Chimica Clinica (IFCC) e l'Istituto per gli Standard Clinici e di Laboratorio (CLSI; precedentemente Comitato Nazionale per gli Standard del Laboratorio Clinico, NC-CLS) elaborarono linee guida per la determinazione degli intervalli di riferimento, e queste linee guida sono tutt'oggi generalmente accettate. Tuttavia, queste linee guida non risolvono completamente i problemi pratici che i laboratori incontrano nel tentativo di trovare i valori di riferimento per ciascun analita.

Solamente istituti con ampie risorse possono avere la possibilità di stabilire gli intervalli di riferimento usando i criteri della IFCC o del CLSI. Di fatto, molti ospedali usano i valori forniti dai produttori, dalla letteratura e da colleghi, e solamente in rari casi con la verifica della trasferibilità approvato dal CLSI<sup>1</sup>, descritta oltre. Anche per i laboratori più grandi e con maggiori possibilità economiche e organizzative, molti punti rimangono complessi e difficili da raggiungere. Questi includono l'elevato numero di soggetti necessari per ottenere intervalli di riferimento, quando sono impiegati metodi non parametrici; come riconoscere e valutare gli outliers; come definire gli intervalli di riferimento per i pazienti ospedalizzati; ecc. La Tabella I riassume i principali problemi che si possono incontrare nella definizione di un intervallo di riferimento e alcune soluzioni possibili, discusse in questo articolo.

## Alcuni esempi pratici

Gli esempi sotto riportati intendono dimostrare

come molti problemi decisionali non possano essere risolti solamente con metodi di calcolo degli intervalli di riferimento corretti e robusti.

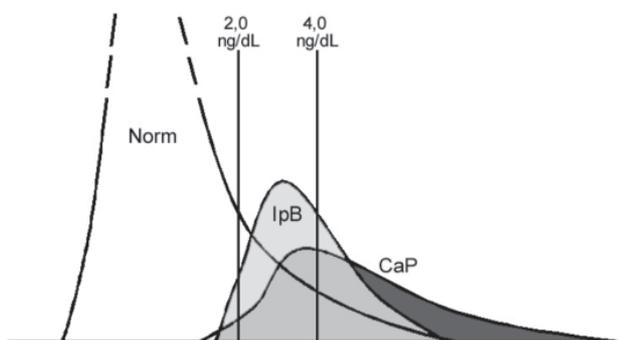
## 1. gli intervalli di riferimento per il TSH

Nel corso di un'appropriate valutazione clinica delle malattie tiroidee, i medici e i laboratoristi utilizzano la misura dei livelli di TSH, basandosi generalmente su attuali intervalli di riferimento di 0,45-4,0 o 0,45-4,5 mUI/L. Quando l'Accademia Nazionale di Biochimica Clinica (NACB) pubblicò la sua linea guida nel 2002 sulla diagnosi ed il monitoraggio della malattia tiroidea, gli editori L.M. Demers e C.A. Spencer<sup>2</sup> aprirono e sostennero un dibattito sulla opportunità di ridurre il limite superiore del TSH da 4,5 a 2,5 mUI/L. Questa opinione era basata sui dati di due studi longitudinali (un sottogruppo del terzo National Health and Nutrition Examination Survey (NHANES III) e lo studio Whickham, retrospettivo di 20 anni), che trovarono come un'alta percentuale di soggetti con livelli di TSH superiori a 2,5 mUI/L avevano un aumentato rischio di sviluppare un ipotiroidismo.

La determinazione degli intervalli di riferimento per il TSH presenta numerosi problemi legati alla ripartizione per età, sesso, razza e alla esclusione dei soggetti con patologia subclinica. I livelli dell'ormone sono più alti durante l'infanzia e nell'età avanzate; inoltre ci sono differenze legate al sesso, essendo questo ormone più alto nelle femmine. Esistono anche differenze razziali, con livelli maggiori nei soggetti bianchi o latinoamericani, più bassi nella popolazione nera. Ma il punto sostanziale della difficoltà di una corretta definizione del limite superiore dell'intervallo di riferimento risiede nella

**Tabella I.** Problemi irrisolti nella definizione degli intervalli di riferimento e possibili soluzioni.

<i>Problema</i>	<i>Soluzioni possibili</i>
Disorientamento nell'uso dei metodi per la definizione degli intervalli di riferimento	<ul style="list-style-type: none"> <li>– Fare riferimento alle linee guida IFCC e CLSI</li> <li>– Utilizzare validi software</li> </ul>
Difficoltà nella ripartizione, esclusione	<ul style="list-style-type: none"> <li>– Considerare che il raffronto dei dati del paziente dovrebbe essere il più possibile simile alla popolazione</li> <li>– Ove possibile, utilizzare anche una ripartizione a posteriori</li> </ul>
Difficoltà ad ottenere 120 soggetti sani <ul style="list-style-type: none"> <li>– Costi</li> <li>– Tempi</li> <li>– Tipologia dei pazienti (es. Pediatrici)</li> </ul>	<ul style="list-style-type: none"> <li>– Uso di metodi robusti che danno peso inferiore ai valori estremi</li> <li>– Uso di trasformazione dei dati in distribuzione normale</li> <li>– Utilizzo della trasferibilità</li> <li>– Approccio della regressione</li> <li>– Ampliare il campione di riferimento</li> <li>– Utilizzo dei metodi indiretti</li> </ul>
Difficoltà a trasformare i dati	<ul style="list-style-type: none"> <li>– Analisi dei dati aberranti</li> <li>– Utilizzo di statistiche non parametriche</li> </ul>
Difficoltà a identificare i dati aberranti	<ul style="list-style-type: none"> <li>– Riferimento alle linee guida CLSI</li> <li>– Metodo di Dixon</li> <li>– Utilizzo di programmi software</li> </ul>
Verifica degli intervalli per cambio metodo	<ul style="list-style-type: none"> <li>– Confronto dei metodi</li> <li>– Utilizzo della trasferibilità</li> </ul>



**Figura 1.** Rappresentazione schematica della distribuzione delle concentrazioni di PSA nelle popolazioni di riferimento (Norm), con ipertrofia prostatica benigna (IpB) e con cancro alla prostata (CaP), con indicazione di due livelli di cut-off.

possibile presenza di un ipotiroidismo subclinico asintomatico, che porterebbe ad includere il soggetto nella popolazione di riferimento. Ad esempio, recentemente Kratsch e collaboratori, selezionando i soggetti anche per assenza di anticorpi anti-tiroidei e per ecografia tiroidea negativa, ha ottenuto  $0,4 - 3,77$  mUI/L<sup>3</sup>; due studi multicentrici condotti pochi anni fa in Spagna e nel Regno Unito ottenevano  $0,43 - 3,69$  e  $0,48 - 3,63$  mUI/L<sup>4,5</sup>. Anche il (NHANES) III, per una popolazione tra i 20 e i 60 anni, propone  $3,56$  and  $3,82$  mUI/L<sup>6</sup>. Quest'ultimo studio dimostra molto bene come le differenze tra gli attuali limiti utilizzati ed i nuovi limiti proposti siano determinate dalla corretta applicazione di criteri di selezione della popolazione di riferimento, considerando non solo l'assenza di malattia ma anche l'esclusione per alcuni fattori di rischio (gravidanza, assunzione di estrogeni, androgeni, litio, la presenza di anticorpi anti-tiroidei o di altri segni biochimici di evidenza di ipo o i ipertiroidismo).

## 2. Gli intervalli di riferimento per il PSA

Il limite di riferimento superiore per un marcatore tumorale potrebbe avere un significato nello screening di questa patologia. Gran parte dei marcatori tumorali hanno ampiamente dimostrato di non avere caratteristiche di sensibilità e specificità sufficienti per poter essere utilizzati come test di screening. L'antigene prostatico specifico (PSA) invece, pur con caratteristiche simili, se non peggiori, non ha la stessa abiura. La task force dei servizi di prevenzione degli Stati Uniti nelle sue linee guida conclude che non ci sono evidenze per essere pro o contro uno screening di routine mediante la misura del PSA o l'esplorazione rettale, (raccomandazione di grado II)<sup>7</sup>. Se non c'è accordo sull'opportunità di fare lo screening, c'è chi sostiene che il classico cut-off di  $4,0$  ng/dL causi una certa perdita di tumori in fase iniziale, stimata nel 10-20%. Sembra anzi esserci un continuo di rischio ad ogni valore di PSA, per cui c'è chi propone cut-off più bassi, di  $3,0$  ng/dL<sup>8</sup>, di  $2,6$  ng/dL<sup>9</sup>, ecc.

Gli intervalli di riferimento possono essere di aiuto? Ha significato l'analisi della distribuzione dei livelli di

PSA nella popolazione normale?

Purtroppo, in questo caso ci troviamo di fronte ad un duplice problema: esiste un'ampia sovrapposizione tra condizioni "normali" e patologiche; non siamo in presenza di una discriminazione tra due condizioni (sano/cancro) ma di almeno tre: sano, ipertrofia benigna, cancro (Figura 1). In queste condizioni, la determinazione di un valore di cut-off non avrà alcun aiuto dalla metodologia per la definizione degli intervalli di riferimento.

## 3. Il limite decisionale per la troponina

Definire un limite per un costituente virtualmente assente nella condizione di salute presenta problemi particolari. Nel caso della troponina si è adottato un duplice criterio: prima si è definito di considerare come limite decisionale per danno miocardico il 99° percentile della popolazione "normale" (questo è tipicamente un approccio da intervallo di riferimento); poi, dato che la precisione dei metodi è molto bassa ai valori trovati, si è raccomandato di avere almeno il 10% di Coefficiente di Variazione (CV) a quel valore di cut-off<sup>10</sup>. Di fatto, nessun sistema analitico ha il 10% di CV ai livelli del 99° percentile trovato<sup>11</sup>, forse perché nella popolazione normale la troponina è completamente assente e quello che si misura è solo un "rumore" analitico determinato dai limiti di sensibilità e di precisione del metodo. Allora, il limite decisionale non fa riferimento alla popolazione normale, ma alle capacità del sistema analitico di essere sensibile e preciso.

## Gli standard CLSI e IFCC

Le prima definizione moderna della metodologia per ottenere validi intervalli di riferimento fu inizialmente definita dalla IFCC<sup>12-17</sup>. Il CLSI ha incorporato questi standard nelle proprie linee guida. In estrema sintesi<sup>18</sup>, la metodologia presenta la definizione di un individuo di riferimento, che deve essere compreso in una popolazione di riferimento, dalla quale è selezionato un campione di riferimento, sul quale verranno eseguite le determinazioni per ottenere dei valori di riferimento, da cui calcolare una distribuzione di riferimento, dalla quale si calcolano i limiti di riferimento che definiscono l'intervallo di riferimento; questo intervallo di riferimento permette la comparazione con un dato osservato (Tabella II).

Il CLSI raccomanda di considerare un minimo di 120 risultati accettabili per calcolare gli intervalli di riferimento, usando un metodo non parametrico. Questo per raggiungere una sufficiente significatività statistica. Si preferiscono i metodi non parametrici perché la distribuzione di frequenza dei livelli dei costituenti del sangue raramente è "normale" o Gaussiana. Una distribuzione normale sarebbe in effetti preferibile se si ha la certezza di una distribuzione "normale" dei dati, perché permette stime con una numerosità campionaria inferiore ed anche di riconoscere con maggiore fa-

**Tabella II.** Relazione tra i diversi termini definiti dai documenti IFCC.

AZIONE		<i>Fase di costruzione di un Intervallo di riferimento</i>	AZIONE
Paziente Valore osservato (risultato)	<i>Comparazione</i>	<b>Individuo di riferimento</b> Compreso nella	<i>Definizione</i>
		<b>Popolazione di riferimento</b> Da cui si seleziona	<i>Selezione</i>
		<b>Gruppo campione di riferimento</b> sul quale sono determinati i	<i>Determinazione</i>
		<b>Valori di riferimento</b> dai quali si può osservare una	<i>Descrizione</i>
		<b>Distribuzione di riferimento</b> dalla quale sono calcolati i	<i>Calcolo</i>
		<b>Limiti di riferimento</b> che definiscono gli	<i>Definizione</i>
		<b>Intervalli di riferimento</b>	

cilità i dati aberranti. Per questa ragione sono stati sviluppati modelli matematici per trasformare le distribuzioni “non normali” in Gaussiane, come la trasformazione logaritmica oppure la trasformazione descritta da Box e Cox<sup>19</sup>.

Occorre inoltre definire cosa è un risultato accettabile e come determinare la significatività statistica. Ottenere risultati accettabili dalla popolazione di riferimento significa riconoscere i valori aberranti o outliers. Il CLSI riconosce che un dato aberrante tra i dati raccolti è una possibilità reale. Tuttavia, fintantoché non si è in grado di capire perché quel dato è aberrante (per esempio un errore nell’analisi), deve essere fatto ogni tentativo per mantenere quel dato anziché cancellarlo. Il punto di vista dello standard è quindi quello di verificare se il dato è accettabile; il CLSI raccomanda l’uso del metodo di Dixon per la verifica dei dati aberranti, in special modo se gli intervalli sono determinati con procedura non parametrica. Una spiegazione semplice del test di Dixon è la seguente: sia R il range di valori (massimo - minimo) e sia D la differenza assoluta tra il valore più estremo (grande o piccolo) ed il successivo valore estremo (il penultimo o il secondo). Se il rapporto D/R supera il valore di 1/3, allora il valore estremo in questione deve essere cancellato<sup>20</sup>. Lo standard pone l’accento sul fatto che se ci sono due o tre risultati aberranti dallo stesso lato questo potrebbe causare il fallimento del metodo, poiché il valore “meno estremo” maschera “il più estremo”. La raccomandazione è di testare il valore aberrante meno estremo come se esso fosse il solo aberrante. Se il test D/R rigetta questo penultimo estremo, allora anche l’ultimo estremo dev’essere ugualmente rigettato<sup>21</sup>. Altri metodi più complessi sono proposti per il riconoscimento dei valori aberranti. Sistemi più robusti possono infatti essere alla base di procedure che riconoscano le popolazione di riferimento sulla sola base della distribuzione di valori<sup>22</sup>.

Una “sufficiente significatività statistica” indica che i limiti trovati corrispondono ai limiti “veri” della po-

polazione di riferimento e che non siano invece falsati da eventuali errori di campionamento. Se si raccolgono 120 dati dal campione di riferimento e si ordinano dal valore più piccolo al valore più grande, i limiti di riferimento riferiti al 95% centrale sono presentati dal valore in posizione quattro (il limite inferiore) e dal valore in posizione 117 (limite superiore). Una numerosità di 120 campioni garantisce una confidenza (al 90%) di questi limiti dal valore in posizione uno al valore in posizione sette per il limite inferiore e dal valore in posizione 114 al valore in posizione 120 per il limite superiore.

Ottenere 120 risultati accettabili per la costruzione di un campione di riferimento è molto spesso difficile, per questa ragione lo standard ammette procedure alternative, quali la trasferibilità dei valori di riferimento da un laboratorio all’altro o dal produttore al laboratorio. La possibilità di trasferire gli intervalli di riferimento va verificata mediante l’esecuzione di 20 test su un campione di riferimento; se tutti i risultati ottenuti stanno all’interno dell’intervallo di riferimento proposto o solo un risultato è al di fuori, l’intervallo è trasferibile. Se più di un risultato è al di fuori dell’intervallo, si possono eseguire altri 20 campioni; se questi 20 campioni sono tutti all’interno dell’intervallo di riferimento al meno di uno, l’intervallo è trasferibile; se ancora una volta più di un risultato è al di fuori dell’intervallo di riferimento l’intervallo proposto dovrà essere rigettato.

Identificare una popolazione di riferimento significa stabilire i criteri di esclusione. Oltre al primo criterio essenziale che è quello dell’assenza della malattia indagata dal test in oggetto, altri criteri generali possono essere di esclusione, come ad esempio l’alcolismo, una malattia recente, l’allattamento, l’obesità, la tossicodipendenza, il fumo, trasfusioni e donazione di sangue recenti, terapie farmacologiche in atto, la gravidanza.

Esistono ulteriori problemi relativi alla popolazione di riferimento, il maggiore dei quali è la verifica della sua omogeneità. Le analisi di distribuzione di frequen-

za dei dati raccolti possono evidenziare sub-popolazioni, con la presenza di più mode della curva di distribuzione. In alternativa, considerazioni di natura epidemiologica possono indurre a verificare se esistano differenze negli intervalli di riferimento per gruppi chiaramente separabili. Nel primo caso sono necessarie numerose osservazioni, in genere oltre 2000, per poter identificare i criteri di ripartizione o di stratificazione e l'esclusione. Questo processo è detto *a posteriori*.

Nel secondo caso si applicano criteri di ripartizione *a priori*. Classici criteri di ripartizione sono: la razza, l'età, il sesso, la fase del ciclo mestruale, la settimana di gravidanza, il fumo, l'esercizio fisico.

Se per riconoscere diversi fattori di esclusione e di ripartizione sono necessarie numerose osservazioni, molto spesso la possibilità di ottenere campioni di popolazione di riferimento numericamente sufficienti non è raggiungibile, per ragioni di costo, tempo, disponibilità di soggetti idonei. L'esempio della popolazione pediatrica è, da questo punto di vista, emblematico. In presenza di campioni piccoli o di classi troppo numerose si possono utilizzare ulteriori sistemi matematici e statistici per stimare gli intervalli di riferimento. Uno di questi è il metodo delle stime robuste; si tratta di sistemi statistici che pesano diversamente i pochi dati disponibili dando maggior rilevanza ai valori centrali e minore ai valori estremi<sup>23</sup>. Quando invece le classi di ripartizione sono troppo numerose, può essere utilizzato l'approccio della regressione per stimare gli intervalli di riferimento. Semplificando, se si ottengono i limiti di riferimento per due o tre valori di una classe di ripartizione rappresentata da una variabile continua (ad esempio età, peso, altezza, ecc.) si possono stimare gli intervalli per altri valori della stessa classe mediante una regressione tra il limite di riferimento e la variabile continua. Se un dato costituente aumenta nella popolazione in maniera abbastanza costante con l'età per anno o per decade, allora si può prevedere anche lo spostamento del suo intervallo di riferimento<sup>24</sup>.

Come visto prima, se si ha la certezza della distribuzione Gaussiana, possono bastare campioni molto meno numerosi. Nel caso di distribuzioni non normali una via può essere quella della trasformazione alla normalità. Un ulteriore approccio al problema dei campioni piccoli è quello di includere sia campioni di soggetti sani che di individui "non sani". Questo può essere un compromesso accettabile, quando esistono diversi livelli dello stato di salute (ad esempio, eccellente, molto buono, buono, modesto, scarso). Può risultare ragionevole includere non solo la prima classe ma anche la seconda della formazione della popolazione di riferimento; l'applicazione di metodi robusti può permettere di pesare i valori estremi in misura minore rispetto ai valori centrali.

L'utilizzo dei dati di soggetti non "di riferimento", espone a numerosi rischi di errore nella stima dell'intervallo di riferimento, mentre rappresenta un approccio interessante quando è possibile utilizzare campioni

molto numerosi. È questo il caso delle banche dati dei laboratori, dove sono registrati centinaia di migliaia di dati. Partendo dal presupposto che, soprattutto per la popolazione non ricoverata, gran parte di questi dati derivano da accertamenti di screening e si riferiscono a individui senza patologia correlata, sono stati sviluppati alcuni metodi, chiamati indiretti, per la stima dell'intervallo di riferimento.

Questi metodi, detti anche oggettivi, si basano sull'assunto che la maggior parte dei risultati prodotti dal laboratorio si riferisce a soggetti non affetti da malattia. Tra questi metodi possiamo citare la procedura di Bhattacharya<sup>25</sup>, il metodo di Martin<sup>26</sup> ed il metodo di Kairisto<sup>27,28</sup>. I tre metodi hanno performance simili tra loro<sup>29</sup> e comparabili con i documenti IFCC quando l'asimmetria e la curtosi delle distribuzioni di frequenza non sono molto ampie. I primi due presuppongono che la distribuzione di frequenza dei risultati sia di tipo gaussiano e comunque operano per normalizzare le distribuzioni di frequenza. La terza procedura ipotizza che la moda della distribuzione di frequenza di tutti i dati sia la stessa della eventuale popolazione di riferimento. In questa procedura, ciascun lato della distribuzione di frequenza è considerato come mezza distribuzione gaussiana, con la stessa moda e la stessa frequenza di moda, ma con differente deviazione standard. Tale metodo ha il vantaggio di essere incluso in un programma statistico per Windows<sup>30</sup>.

I metodi citati mostrano numerosi vantaggi: la procedura è semplice ed economica; i limiti sono derivati da dati ottenuti nelle stesse condizioni preanalitiche, con le stesse modalità analitiche e nella stessa popolazione in cui saranno utilizzati; è possibile, quando necessario, ripartire facilmente i risultati dei pazienti sulla base di variazioni biologiche (età, sesso, gravidanza, ecc.).

Una nostra verifica degli intervalli di riferimento del TSH, conseguente alle discussioni in atto in questi tempi, ha dato prova della validità di questi metodi: applicato ad una popolazione di 15389 femmine e 3862 maschi che negli ultimi due anni avevano eseguito il dosaggio dei livelli di TSH in associazione con la misura degli anticorpi antitiroidei, il computo secondo Kairisto ha determinato un limite superiore dell'intervallo di riferimento pari a 3,7 mUI/L, di fatto identico a quello proposto dal NHANES III<sup>31</sup>.

Gli intervalli di riferimento rimangono il sistema maggiormente utilizzato nella categorizzazione della condizione clinica dei pazienti. Essi rappresentano comunque una espressione matematico statistica della distribuzione dei valori in una popolazione di riferimento e non sono sempre e comunque identificabili con limiti decisionali. Quando chiesero a David Sackett, uno dei più importanti medici che praticano la medicina basata sull'evidenza, cosa si aspettasse dalla patologia clinica basata sull'evidenza, egli espresse tre desideri: discutere della malattia del paziente con il collega; avere un'evidenza disponibile che supporti la validità, l'importanza e l'utilità clinica dei test biochimici; smettere

di utilizzare i range di normalità<sup>32</sup>.

Quando si valuta un risultato analitico, i termini di confronto sono molto spesso più di uno: i valori della popolazione di riferimento (i classici intervalli di riferimento), i valori ideali e desiderabili, i valori precedenti di quel soggetto, i valori osservati in diverse patologie, il significato di quel valore integrato con altri parametri e misurazioni. Ad esempio, il livello di colesterolo è confrontato con livelli ideali e con la valutazione del rischio complessivo; la concentrazione di creatinina in soggetti con insufficienza renale è confrontata con diversi stadi della progressione della malattia (confronto con livelli in condizioni patologiche diverse), ma un rapido incremento (confronto con i valori precedenti) può indicare l'evoluzione verso l'insufficienza acuta; la valutazione delle concentrazioni dell'emoglobina per una terapia trasfusionale fa riferimento a diverse condizioni patologiche, alle condizioni attuali (chirurgia, trauma, emorragia in atto, ecc).

Il punto essenziale di questa discussione è che ciò che serve nel referto, a fianco di un risultato, non è una classificazione di "normalità" ma uno o più livelli decisionali. Gli intervalli di riferimento rappresentano un approccio alla definizione di livelli decisionali ma altre considerazioni, basate soprattutto sull'effetto finale sul paziente e quindi su considerazioni di tipo clinico terapeutico, possono essere maggiormente efficaci nell'utilizzo dei test di laboratorio.

## Bibliografia

- National Committee for Clinical Laboratory Standards. How to define and determine reference intervals in the clinical laboratory; Approved Guideline. NCCLS document C28-A2. Wayne, PA: NCCLS, 2000; Vol. 20, n. 13.
- Demers LM, Spencer CA. Laboratory medicine practice guidelines: laboratory support for the diagnosis and monitoring of thyroid disease. *Clin Endocrinol* 2003; 58:138-40.
- Kratzsch J, Fiedler GM, Leichtle A, Brügel M, Buchbinder S, Otto L, et al. New reference intervals for thyrotropin and thyroid hormones based on National Academy of Clinical Biochemistry criteria and regular ultrasonography of the thyroid. *Clin Chem* 2005; 51:1480-6.
- Ferre-Masferrer M, Fuentes-Arderiu X, Goma-Llongueras M, Aluma-Trullas A, Aramendi-Ramos M, Castano-Vidriales JL, et al. Regional reference values for some quantities measured with the Advia Centaur analyser: a model of cooperation between the in vitro diagnostic industry and clinical laboratories. *Clin Chem Lab Med* 2001; 39:166-9.
- Andrew CE, Hanning I, McBain AM, Moody D, Price A. A model for a multicentre approach to the derivation of reference intervals for thyroid hormones and testosterone for laboratories using identical analysers. *Clin Chem Lab Med* 2000; 38:1013-9.
- Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA, et al. Serum TSH, T4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab* 2002; 87:489-99.
- U.S. Preventive Services Task Force. Screening for prostate cancer: recommendation and rationale. *Ann Intern Med* 2002; 137:915-6.
- Thompson IM, Ankerst DP, Chi C, Lucia MS, Goodman PJ, Crowley JJ, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. *JAMA* 2005; 294:66-70.
- Nadler RB, Loeb S, Roehl KA, Antenor JA, Eggner S, Catalona WJ. Use of 2.6 ng/ml prostate specific antigen prompt for biopsy in men older than 60 years. *J Urol* 2005; 174:2154-7.
- Alpert J, Thygesen K, for the Joint European Society of Cardiology/American College of Cardiology Committee. Myocardial infarction redefined—a consensus document of the Joint European Society of Cardiology/American College of Cardiology Committee for the Redefinition of Myocardial Infarction. *J Am Coll Cardiol* 2000; 36:959-69.
- Panteghini M, Pagani F, Yeo KT, Apple FS, Christenson RH, Dati F, et al. Committee on Standardization of Markers of Cardiac Damage of the IFCC. Evaluation of imprecision for cardiac troponin assays at low-range concentrations. *Clin Chem* 2004; 50:327-32.
- Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved Recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *J Clin Chem Clin Biochem* 1987; 25:337-42.
- PetitClerc C, Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 2. Selection of individuals for the production of reference values. *J Clin Chem Clin Biochem* 1987; 25:639-44.
- Solberg HE, PetitClerc C. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. *J Clin Chem Clin Biochem* 1988; 26:593-8.
- Solberg HE, Stamm D. Approved recommendation on the theory of reference values. Part 4. Control of analytical variation in the production, transfer, and application of reference values. *Eur J Clin Chem Clin Biochem* 1991; 29:531-5.
- Solberg HE. Approved recommendations (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *J Clin Chem Clin Biochem* 1987; 25:645-56.
- Dybkaer R, Solberg HE. Approved recommendations (1987) on the theory of reference values. Part 6. Presentation of observed values related to reference values. *J Clin Chem Clin Biochem* 1987; 25:657-62.
- Giavarina D, Dorizzi RM, Guerra G. Linee guida per la produzione di intervalli di riferimento. *Riv Med Lab - IJLM* 2001; 2:99-105.
- Box G, Cox D. An analysis of transformation. *J R Stat Soc* 1964; B26:211-52.

20. Dixon W. Processing data for outliers. *Biometrics* 1953; 9:74-89.
21. Linnet K. Two-stage transformation systems for normalization of reference distributions evaluated. *Clin Chem* 1987; 33:381-6.
22. Horn PS, Pesce AJ. Outliers. In: *Reference Intervals. A user's guide*. Washington: AACCC press 2005. p. 29-36.
23. Horn P. A biweight prediction interval for random samples. *J Am Stat Assoc* 1988; 83:249-56.
24. Horn PS, Pesce AJ. Regression approach. In: *Reference Intervals. A user's guide*. Washington: AACCC press 2005. p. 71-75.
25. Bhattacharya CG. A simple method of resolution of a distribution into gaussian components. *Biometrics* 1967; 23:115-35.
26. Martin HF, Hologgitas JV, Driscoll J, Fanger H, Gudzinowicz BJ, Barnett R. Reference values based on population accessible to hospitals. In: Gräsbeck R, Alström T, eds. *Reference values in laboratory medicine*. Chichester: John Wiley & Sons Ltd., 1981; 233-62.
27. Kouri T, Kairisto V, Virtanen A, Uusipaikka E, Rajamaki A, Finneman H, et al. Reference intervals developed from data for hospitalized patients: computerized method based on combination of laboratory and diagnostic data. *Clin Chem* 1994; 40:2209-15.
28. Kairisto V, Hanninen KP, Leino A, Pulkki K, Peltola O, Nanto V, et al. Generation of reference values for cardiac enzymes from hospital admission laboratory data. *Eur J Clin Chem Clin Biochem* 1994; 32:789-96.
29. Ferre-Masferrer M, Fuentes-Arderiu X, Puchal-Ane R. Indirect reference limits estimated from patients' results by three mathematical procedures. *Clin Chim Acta* 1999; 279:97-105.
30. Kairisto V, Poola A. Software for illustrative presentation of basic clinical characteristics of laboratory tests: GraphROC for Windows. *Scand J Clin Lab Invest* 1995; 55:43-60.
31. Giavarina D, Dorizzi RM, Soffiati G. Indirect Methods for Reference Intervals Based on Current Data. *Clin Chem* 2006; 52:335-7.
32. RA Moore. On the need for evidence-based clinical biochemistry. *Evidence-based Medicine* 1998; 3:7-8.