

# La Rivista Italiana della Medicina di Laboratorio Italian Journal of Laboratory Medicine

Linee guida per il confronto di procedure di esami di laboratorio: utilizzo delle indicazioni di CLSI EP09-A3 e EP31-A-IR. Guidelines for the comparison of laboratory examination procedures: use of indications of CLSI EP09-A3 and EP31-A-IR.  
--Manuscript Draft--

<b>Manuscript Number:</b>	RIME-D-15-00033
<b>Full Title:</b>	Linee guida per il confronto di procedure di esami di laboratorio: utilizzo delle indicazioni di CLSI EP09-A3 e EP31-A-IR. Guidelines for the comparison of laboratory examination procedures: use of indications of CLSI EP09-A3 and EP31-A-IR.
<b>Article Type:</b>	Review (Rassegna)
<b>Section/Category:</b>	Basic Science Section
<b>Keywords:</b>	methods comparison; statistics; accreditation; ISO; CLSI. comparazione metodi; statistica; accreditamento; CLSI.
<b>Corresponding Author:</b>	Marco Pradella, M.D.  Castelfranco veneto, Treviso ITALY
<b>Corresponding Author Secondary Information:</b>	
<b>Corresponding Author's Institution:</b>	
<b>Corresponding Author's Secondary Institution:</b>	
<b>First Author:</b>	Marco Pradella, M.D.
<b>First Author Secondary Information:</b>	
<b>Order of Authors:</b>	Marco Pradella, M.D.  Bruno Mario Cesana, Prof.
<b>Order of Authors Secondary Information:</b>	
<b>Funding Information:</b>	
<b>Abstract:</b>	<p>Summary</p> <p>Background. Accreditation standards (ISO 15189 and ISO 22870) require both the medical laboratory and the manufacturer of systems and reagents to compare methods that provide results for the same examination. CLSI guideline EP09 and EP31 provide some operational tools and statistics.</p> <p>Methods . The text of the CLSI guidelines EP09 and EP31 has been critically analyzed and linked to key literature references. Was also considered the guideline CLSI EP30 for the particular case of commutable calibrators.</p> <p>Results . The guidelines provide very different indications to the medical laboratory and the manufacturer of diagnostic systems. Fundamental is a graphical representation of the results . The tables for interpretation of CLSI EP31 appear quite complex .</p> <p>Conclusions . The comparison of methods of examinations is one of the main managerial commitments of medical laboratory, to which are devoted considerable human resources and mainly cultural. CLSI guidelines EP09 and EP31 provide different tools for the laboratory and for the manufacturer of diagnostic systems, sometimes complex.</p> <p>Premesse. I requisiti degli standard di accreditamento (ISO 15189 e ISO 22870) obbligano sia il laboratorio medico che il produttore di sistemi e reagenti a comparare i metodi che forniscono risultati per lo stesso esame. Le linee guida CLSI EP09 ed EP31 forniscono alcuni strumenti operativi e statistici.</p> <p>Metodi. Il testo delle linee guida CLSI EP09 ed EP31 è stato analizzato criticamente e collegato ai principali riferimenti di letteratura. E' stata considerata anche la linea guida</p>

CLSI EP30 per il caso particolare di taratura con calibratori commutabili.

Risultati. Le linee guida forniscono indicazioni molto diverse per il laboratorio medico e per il produttore di sistemi diagnostici. Fondamentale è la rappresentazione grafica dei risultati. Le tabelle interpretative di CLSI EP31 si presentano abbastanza complesse.

Conclusioni. L'attività di confronto di metodi è uno dei principali impegni gestionali del laboratorio medico, a cui vanno dedicate notevoli risorse umane e soprattutto culturali. Le linee guida CLSI EP09 ed EP31 forniscono strumenti, talvolta complessi, diversi per il laboratorio e per il produttore di sistemi diagnostici.

*Conflict of Interest Disclosure Form*

È politica *La Rivista Italiana della Medicina di Laboratorio-Italian Journal of Laboratory Medicine* garantire l'equilibrio, l'indipendenza, l'obiettività e il rigore scientifico dei suoi contenuti. Tutti gli autori sono tenuti a esplicitare ai lettori un conflitto reale o apparente di interessi che possono avere un rapporto diretto con il loro articolo.

Questo riguarda i rapporti con le aziende farmaceutiche, i produttori di dispositivi biomedicali o altre società i cui prodotti o servizi possono essere correlati all'argomento dell'articolo o alla sponsorizzazione dello studio descritto.

Non si vuole assolutamente contrastare la pubblicazione di articoli da parte di autori con un potenziale conflitto di interessi. L'esplicitazione di quest'ultimo infatti è necessario esclusivamente ai lettori, che avranno così gli strumenti per potersi formare un proprio giudizio e stabilire se il conflitto di interessi abbia o meno portato a una possibile distorsione sia nell'esposizione sia nelle conclusioni presentate.

Si prega il *corresponding author* di compilare e inviare il modulo per l'*Editor-in-Chief* per conto di tutti gli autori elencati di seguito.

It is the policy of *La Rivista Italiana della Medicina di Laboratorio-Italian Journal of Laboratory Medicine* to ensure balance, independence, objectivity, and scientific rigor in the Journal. All authors are expected to disclose to the readers any real or apparent conflict(s) of interest that may have a direct bearing on the subject matter of the article. This pertains to relationships with pharmaceutical companies, biomedical device manufacturers or other corporation whose products or services may be related to the subject matter of the article or who have sponsored the study.

The intent of the policy is not to prevent authors with a potential conflict of interest from publication. It is merely intended that any potential conflict should be identified openly so that the readers may form their own judgements about the article with the full disclosure of the facts. It is for the readers to determine whether the authors' outside interest may reflect a possible bias in either the exposition of the conclusions presented.

The corresponding author will complete and submit this form to the Editor-in-Chief on behalf of all authors listed below.

*Rivista/Journal* [La Rivista Italiana della Medicina di Laboratorio - Italian Journal of Laboratory Medicine](#)

*Titolo dell'Articolo/Article Title*

[Linee guida per il confronto di procedure di esami di laboratorio: utilizzo delle indicazioni di CLSI EP09-A3 e EP31-A-IR](#)

*Autori/Authors*

[Marco Pradella, Bruno Mario Cesana](#)

Per cortesia tenga presente che il conflitto di interessi verrà pubblicato su ogni articolo.

Please note that a conflict of interest statement is published with each paper.

Per cortesia dichiarare qui di seguito l'eventuale conflitto di interesse

If any conflict exists, please define hereafter:

*(Se non è presente alcun conflitto scrivere "Nessuno", altrimenti descriva gli accordi/interessi finanziari con una o più organizzazioni che potrebbero essere percepiti come reali o apparenti conflitti di interesse in relazione ai contenuti del suo articolo.)*

*(If none, "None" or describe financial interest/arrangement with one or more organizations that could be perceived as a real or apparent conflict of interest in the context of the subject of this article):*

.....  
Nessuno  
.....  
.....  
.....  
.....

Nome/Name **Marco Pradeva**  
.....  
Firma/Signature   
.....  
Data/Date **9-12-2015**  
.....

Per cortesia compili questo documento e lo carichi in Editorial Manager insieme al suo articolo.

Please fill in this document and upload it in Editorial Manager while submitting your manuscript.

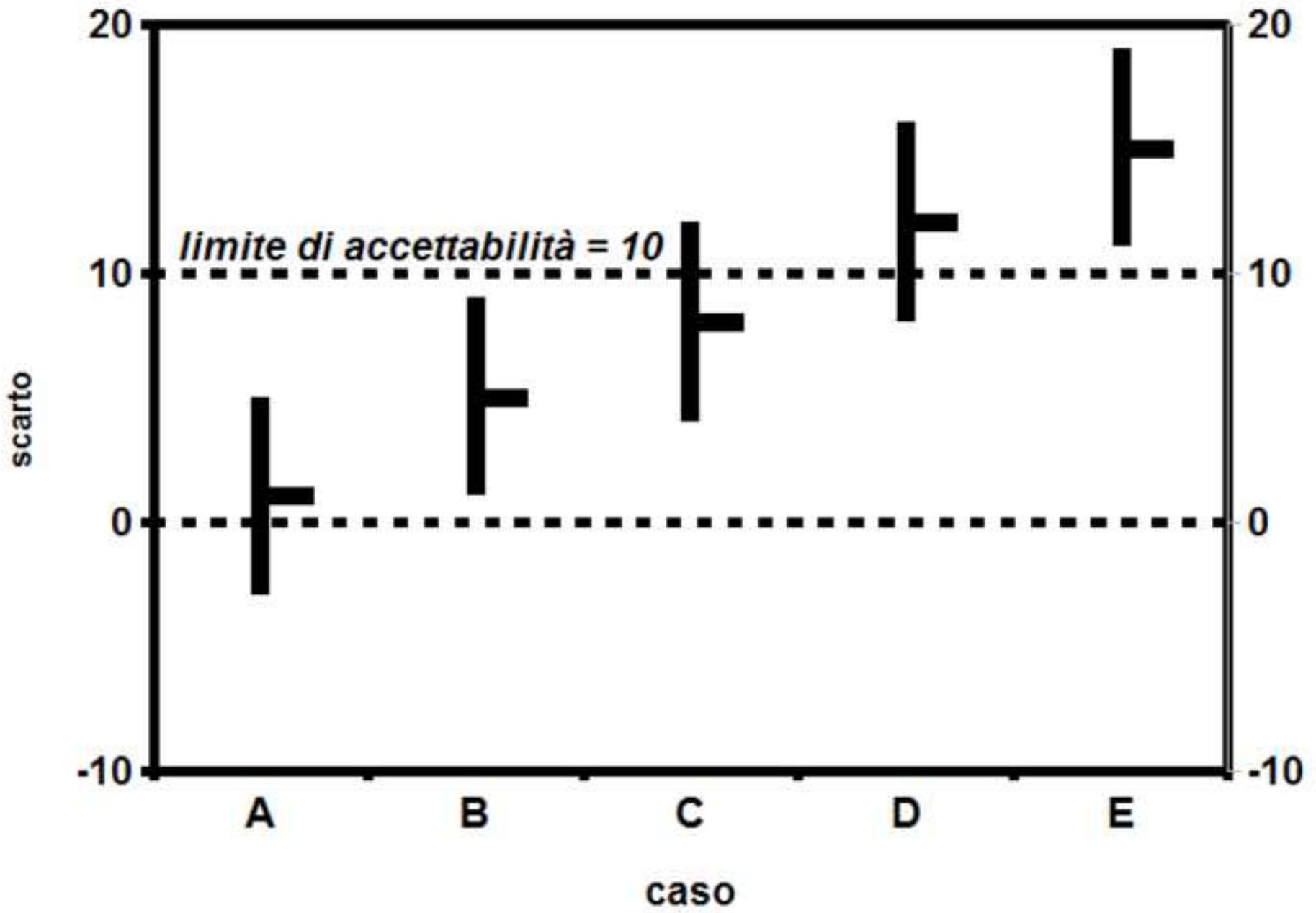
La Rivista Italiana della Medicina di Laboratorio – Italian  
Journal of Laboratory Medicine

Direttore Scientifico: Dorizzi, R.M.; Bizzaro, N.

ISSN: 1825-859X (print version)

ISSN: 2039-6821 (electronic version)

Journal no. 13631



[Click here to view linked References](#)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Rassegna

# Linee guida per il confronto di procedure di esami di laboratorio: utilizzo delle indicazioni di CLSI EP09-A3 ed EP31-A-IR

1  
2 **Guidelines for the comparison of laboratory examination**  
3 **procedures: use of indications of CLSI EP09-A3 and EP31-**  
4 **A-IR**  
5  
6  
7  
8  
9

10  
11 Marco Pradella • Bruno Mario Cesana  
12  
13

14  
15 M. Pradella<sup>1</sup>, B. M. Cesana<sup>2</sup>  
16  
17  
18

19  
20 <sup>1</sup> Laboratorio analisi degli Ospedali di Castelfranco Veneto e Montebelluna (TV), Italia; Società  
21 Italiana di Patologia Clinica e Medicina di Laboratorio – SIPMeL, Commissione Nazionale Qualità  
22 ed Accredimento (COMMQUA); Comitato Italiano per la Standardizzazione dei Metodi  
23 Ematologici e di Laboratorio (CISMEL).  
24

25  
26 <sup>2</sup> Unità di Statistica e Biomatematica, Dipartimento di Medicina Molecolare e Traslazionale,  
27 Università degli Studi di Brescia; Società Italiana di Statistica Medica ed Epidemiologia Clinica –  
28 SISMEC, Gruppo di lavoro “Misura nel laboratorio biomedico e nella clinica (MisMed)”  
29  
30  
31

32  
33  M. Pradella  
34

35 marco.pradella@yahoo.it  
36  
37  
38

39 *NOTA: dal documento Q7 presentato come Linea Guida della Commissione SIPMeL Qualità e*  
40 *Accreditamento al 1° Congresso Nazionale SIPMeL, Roma 2015.*  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 **Riassunto** *Premesse.* I requisiti degli standard di accreditamento (ISO 15189 e ISO 22870)  
2 obbligano sia il laboratorio medico sia il produttore di sistemi e reagenti a comparare i metodi che  
3 forniscono risultati per lo stesso esame. Le linee guida CLSI EP09 ed EP31 forniscono alcuni  
4 strumenti operativi e statistici. *Metodi.* Il testo delle linee guida CLSI EP09 ed EP31 è stato  
5 analizzato criticamente e collegato ai principali riferimenti di letteratura. È stata considerata anche  
6 la linea guida CLSI EP30 per il caso particolare di taratura con calibratori commutabili. *Risultati.*  
7 Le linee guida forniscono indicazioni molto diverse per il Laboratorio medico e per il produttore di  
8 sistemi diagnostici. Fondamentale è la rappresentazione grafica dei risultati. Le tabelle  
9 interpretative di CLSI EP31 si presentano abbastanza complesse. *Conclusioni.* L'attività di  
10 confronto di metodi è uno dei principali impegni gestionali del Laboratorio medico, a cui vanno  
11 dedicate notevoli risorse umane e soprattutto culturali. Le linee guida CLSI EP09 ed EP31  
12 forniscono strumenti, talvolta complessi, diversi per il Laboratorio e per il produttore di sistemi  
13 diagnostici.

14 **Parole chiave** Comparazione metodi • Statistica • Accredimento • ISO • CLSI  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

25 **Summary** *Background.* Accreditation standards (ISO 15189 and ISO 22870) require both the  
26 medical laboratory and the manufacturer of systems and reagents to compare methods that provide  
27 results for the same examination. CLSI guideline EP09 and EP31 provide some operational tools  
28 and statistics. *Methods.* The text of the CLSI guidelines EP09 and EP31 has been critically analyzed  
29 and linked to key literature references. Was also considered the guideline CLSI EP30 for the  
30 particular case of commutable calibrators. *Results.* The guidelines provide very different indications  
31 to the medical Laboratory and the manufacturer of diagnostic systems. Fundamental is a graphical  
32 representation of the results. The tables for interpretation of CLSI EP31 appear quite complex.  
33 *Conclusions.* The comparison of methods of examinations is one of the main managerial  
34 commitments of medical Laboratory, to which are devoted considerable human resources and  
35 mainly cultural. CLSI guidelines EP09 and EP31 provide different tools for the laboratory and for  
36 the manufacturer of diagnostic systems, sometimes complex.

37 **Keywords** Methods comparison • Statistics • Accreditation • ISO • CLSI  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## **T1 Premessa: requisiti di accreditamento sulla comparabilità dei risultati**

In questa linea guida vengono presentate e discusse criticamente le raccomandazioni dei documenti CLSI EP09, EP30 ed EP31 come strumenti per soddisfare ai requisiti di accreditamento ISO 15189 e ISO 22870 (Tab. 1).

## **T2 ISO 15189:2012 punto 5.6.4 comparabilità dei risultati degli esami**

Lo standard per l'accREDITAMENTO dei Laboratori medici prescrive che il Laboratorio debba avere una modalità definita per confrontare procedure di esame, strumenti e metodi, al fine di accertare la comparabilità dei risultati dei campioni dei pazienti negli intervalli clinicamente appropriati. Ciò vale in tutti i casi, ovvero per metodi dello stesso tipo o diversi, per strumenti, per sedi differenti o per l'insieme di tutte le possibili combinazioni.

Nel caso particolare di risultati riferibili alla stessa origine metrologica, si possono considerare anche metrologicamente comparabili se i calibratori sono commutabili.

Il Laboratorio deve informare gli utenti di qualsiasi differenza nella comparabilità dei risultati e deve discuterne le implicazioni per la pratica clinica. Per esempio, quando i sistemi analitici forniscono intervalli di misura diversi per lo stesso misurando o quando i metodi di misura vengono modificati.

Il Laboratorio deve documentare, registrare e, se opportuno, agire prontamente in base ai risultati delle comparazioni. I problemi o i difetti riscontrati devono essere trattati e adeguatamente sistemati e va conservata la registrazione degli interventi.

## **T2 ISO 22870:2006 punto 5.6.8, comparabilità di differenti metodi, strumenti o sedi**

Lo standard per il *point of care* ISO 22870 al punto 5.6.8 prescrive questo: “Si applica il punto 5.6.7 della ISO 15189:2003 e, tra l'altro, quanto segue. Il direttore del Laboratorio deve convalidare i processi seguenti per la fornitura dei servizi. a) L'esattezza e la precisione e, se appropriato, la linearità della risposta degli strumenti devono essere verificate con il programma di controllo della qualità. b) Devono essere utilizzati campioni dei pazienti suddivisi, o altri materiali accettabili per il controllo della qualità, per verificare le prestazioni dei sistemi per le analisi decentrate utilizzati in più siti...”.

*NOTA STATISTICA. In accordo alla terminologia statistica, è più comunemente utilizzato il termine “accuratezza”, anche come immediata traduzione del termine inglese “accuracy”, con cui si esprime la proprietà di una misurazione di tendere al “valore vero” (mancanza di errore sistematico o “bias” o distorsione della misurazione effettuata). Comunque, il termine “esattezza” più riferibile all'inglese “trueness” può essere considerato equivalente, senza coinvolgere il concetto dell'errore casuale che si associa alla misurazione e che è quantificato, secondo la terminologia statistica più usata, con il termine di “precisione”.*

## T1 Confronto di metodi e stima dello scarto sistematico (*bias*)

La situazione tipica in cui è necessario confrontare due metodi è quella del produttore o del Laboratorio clinico che dispone di un metodo normalizzato (standard) o di riferimento. Il produttore vuole in questo caso definire le caratteristiche prestazionali di un suo metodo commerciale per quanto riguarda l'aspetto dell'esattezza o differenza dal valore "vero" (validazione), mentre il Laboratorio ha bisogno di verificare le dichiarazioni del produttore sotto il medesimo aspetto (verifica).

Più spesso, invece, non si dispone di metodi di riferimento, ma solo di un metodo "in uso". Quindi non si confronta con un valore "vero" e il risultato dell'esperimento è da considerarsi semplicemente una "differenza" e non un "*bias*".

La linea guida CLSI EP09-A3 suggerisce due approcci distinti, uno per il caso del produttore e uno per il caso del Laboratorio medico. Nel primo caso il produttore per validare il metodo deve analizzare almeno 100 campioni e applicare l'analisi statistica della regressione. Nel secondo caso il Laboratorio per verificare le dichiarazioni del produttore può analizzare almeno 40 campioni e limitarsi a un grafico delle differenze (Tab. 2).

I metodi da confrontare vengono convenzionalmente denominati "procedura di misura candidata" e "procedura di misura comparativa".

Si può verificare anche il caso in cui la comparazione debba essere effettuata all'interno dello stesso metodo, per esempio per confrontare il comportamento di tipi differenti di campioni o di provette, lotti diversi dello stesso reagente o strumenti di backup (a questo scopo è stata predisposta la linea guida CLSI EP31).

CLSI EP31 è concepita per il confronto di procedure di esame all'interno di una struttura sanitaria, ovvero un insieme di Ambulatori medici, Cliniche, Ospedali, Laboratori di riferimento, governati dalla stessa amministrazione e in cui i risultati di laboratorio possono essere utilizzati indifferentemente da qualsiasi clinico inserito nel sistema. Con EP31 si possono confrontare al massimo 10 strumenti.

La situazione tipica in cui si applica CLSI EP31 è quella del *point of care* (POCT). La comparazione dei POCT con gli strumenti del Laboratorio a cui si riferiscono spesso non può essere condotta sugli stessi campioni, bensì su campioni diversi (come sangue capillare e plasma) raccolti simultaneamente.

I motivi per avviare una comparazione sono tipicamente: una scadenza programmata (frequente, ovvero giornaliera o settimanale; periodica, ovvero trimestrale o semestrale), l'esito negativo di un monitoraggio programmato, l'esito negativo di un esercizio interlaboratorio, una variazione osservata nel controllo di qualità statistico o in un altro tipo di monitoraggio (per esempio, medie mobili con i risultati dei pazienti), un cambio di lotto di reagenti o calibratori, infine una necessità *ad hoc* a seguito di un intervento di manutenzione importante, della sostituzione di un componente, di un aggiornamento del software o (meno raramente di quanto si pensi) di una segnalazione da parte dei clinici con il sospetto di risultati non accurati.

1 Per entrambe le procedure “candidata” e “comparativa” gli operatori devono innanzitutto  
2 familiarizzare con le operazioni sullo strumento e le fasi della misura, con la manutenzione dei  
3 sistemi analitici, con la preparazione dei campioni e infine con le funzioni di taratura e di controllo  
4 di qualità. Nella familiarizzazione si comprende l’addestramento fornito dal produttore. Si  
5 raccomanda un periodo di cinque giorni come durata media per la familiarizzazione, ma per sistemi  
6 semplici possono essere sufficienti meno giorni, mentre per sistemi complessi possono essere  
7 necessari più giorni.  
8  
9

10  
11 I campioni vanno raccolti secondo le regole della fase pre-esame (ISO 15189 clausola 5.4),  
12 registrando eventuali informazioni cliniche, anagrafiche e pre-analitiche (emolisi, ittero, lipemia  
13 ecc.). Per avere la possibilità di determinazioni replicate se i campioni individuali sono piccoli, si  
14 può ricorrere a “minipools”, ovvero mescolanze di pochi campioni (se possibile non più di due) con  
15 simili livelli di misura e storie cliniche. È possibile anche mescolare campioni con livelli diversi al  
16 fine di ottenere concentrazioni in intervalli altrimenti mancanti. Per mescolare sangue intero è  
17 necessario verificare la compatibilità sierologica. Comunque, questi campioni “manipolati” non  
18 dovrebbero essere più del 20% del totale.  
19  
20  
21  
22  
23

24 Diluizione, impoverimento o arricchimento (“spiking”) dei campioni dovrebbero essere evitati,  
25 poiché è stato dimostrato che i campioni così manipolati perdono la commutabilità tra metodi (v.  
26 sezione 6.1 *Study Samples*, Capitolo 5). Comunque i campioni modificati devono essere identificati  
27 e riconoscibili nelle rappresentazioni grafiche.  
28  
29  
30

## 31 **T1 Validazione della procedura di esame da parte del** 32 **produttore e verifica della procedura di esame da parte del** 33 **Laboratorio** 34 35 36

37 Quando il produttore di un sistema analitico dispone di un metodo di riferimento, dovrebbe ottenere  
38 uno scarto (*bias*) tra i due non clinicamente rilevante. Spesso, però, siamo in presenza di un nuovo  
39 metodo automatizzato (candidato) che dovrebbe sostituire il metodo precedente, manuale o  
40 parzialmente automatizzato. In questo caso l’obiettivo è stabilire l’entità della differenza.  
41  
42

43 Il Laboratorio che sostituisce un metodo in uso con uno nuovo ha la necessità di misurare lo scarto  
44 (*bias*) tra i due, per modificare eventualmente i limiti di riferimento o i valori decisionali. Dovrebbe  
45 quindi analizzare almeno 40 campioni distribuiti nell’intervallo analitico, in singolo. Se ne esiste la  
46 possibilità, ottenere risultati di misure replicate riduce l’incertezza della stima dello scarto (*bias*).  
47  
48

49 EP09-A3 raccomanda al produttore di utilizzare per la validazione almeno 100 campioni distribuiti  
50 lungo tutto l’intervallo analitico. Lo schema sperimentale dovrebbe comprendere fattori di  
51 variabilità come taratura, serie, giornata, lotto di reagente, lotto di calibrante e attrezzatura. Al  
52 contrario, la ripetibilità stretta dovrebbe essere minimizzata utilizzando la media di più replicati. Se  
53 i replicati sono almeno tre, è possibile utilizzare la mediana.  
54  
55  
56

57 Diversa è la situazione del produttore che intende verificare che una procedura di esame sia  
58 conforme alle caratteristiche prestazionali dichiarate. Anche in questo caso si raccomandano almeno  
59 100 campioni, misurati in singolo. Ottenere misure replicate, tuttavia, consente di visualizzare  
60  
61  
62  
63  
64  
65

1 meglio lo scarto (*bias*) e di applicare direttamente la regressione di Deming [1] (vedi EP09-A3E,  
2 Appendix E., pag. 56-59 e Appendix F., Constant CV (Weighted) Deming regression, pag. 60-61).  
3 L'esperimento dovrebbe essere condotto in diverse sedi esterne al sito di produzione, per tre-cinque  
4 giorni in ciascuna sede.  
5  
6

7 Se si sospetta che popolazioni differenti si comportino diversamente nella comparazione,  
8 l'esperimento deve essere condotto con almeno 100 campioni per ogni popolazione. Per esempio,  
9 donne in gravidanza rispetto a maschi adulti per diversi metodi immunometrici o pazienti in dialisi  
10 rispetto a soggetti non dializzati per l'ormone paratiroideo.  
11  
12

13 EP31 prescrive che nella scelta dei materiali per la comparazione si cerchi innanzitutto di garantire  
14 la commutabilità, ovvero l'equivalenza della relazione matematica tra i risultati di diverse  
15 procedure di misurazione di un materiale di riferimento e di campioni rappresentativi di ciò che si  
16 intende misurare (vedasi capitolo 5). Campioni freschi ottenuti da pazienti costituiscono il materiale  
17 ideale per effettuare il confronto. La commutabilità va verificata per tutti i materiali diversi dai  
18 campioni freschi dei pazienti. L'utilizzo di materiali per il controllo di qualità, "controlli" forniti dal  
19 produttore o materiali provenienti da programmi interlaboratorio (VEQ, PR, CQ "esteso") espone a  
20 notevoli rischi e va, quando possibile, evitato. I calibratori e i materiali per la verifica di linearità  
21 sono decisamente da escludersi.  
22  
23  
24  
25  
26

27 Nelle comparazioni rapide, secondo EP31, può apparire difficile e oneroso coprire l'intero  
28 intervallo analitico. Quindi in Laboratorio si può scegliere di studiare i valori vicini ai limiti  
29 decisionali o semplicemente vicini ai valori dei materiali di controllo.  
30  
31

32 I risultati vanno raccolti in una tabella ed esaminati per individuare i valori macroscopicamente  
33 anomali (*outlier*). Questi valori vanno indagati approfonditamente. I casi con errori documentati  
34 vanno esclusi dall'analisi finale.  
35  
36

37 Anche per EP31 la semplice ispezione della tabella dei risultati o di un grafico elementare può  
38 fornire informazioni sufficienti. Il Direttore del Laboratorio deciderà poi se procedere a elaborazioni  
39 più sofisticate.  
40  
41

42 Quindi i risultati vanno esaminati visivamente nelle rappresentazioni grafiche elementari, il  
43 diagramma di dispersione (*scatter plot*) e il diagramma delle differenze (*difference plot*).  
44  
45

46 Nel diagramma di dispersione viene riportato sull'asse delle ordinate il risultato del metodo  
47 candidato e sull'asse delle ascisse il risultato del metodo comparativo. La nuvola di punti può  
48 assumere una forma a cilindro se lo scarto tipo (deviazione standard, DS, o scarto quadratico  
49 medio) è costante nell'intervallo analitico, o a cono se, invece, è costante il coefficiente di  
50 variazione (CV). Più spesso assume una forma intermedia tra le due o addirittura irregolare.  
51  
52  
53

54 *NOTA: il termine "scarto tipo", anche se meno diffuso, corrisponde esattamente alla traduzione*  
55 *ufficiale di "standard deviation" secondo UNI ISO 3534-1 [2] e UNI CEI ENV 13005 [3].*  
56

57 Si conoscono quattro tipi di diagramma delle differenze. Sull'asse delle ordinate viene riportata la  
58 differenza tra metodo candidato e metodo comparativo e sull'asse delle ascisse può essere riportato  
59 il valore del metodo comparativo oppure la media tra il risultato candidato e il risultato comparativo  
60  
61  
62  
63  
64  
65

1 (c.d. grafico di Bland-Altman). Inoltre, sull'asse delle ordinate può essere riportata la differenza  
2 assoluta tra i due metodi oppure la differenza relativa o percentuale, a seconda che si ritenga che lo  
3 scarto sia costante o proporzionale al risultato. Ne risultano quindi quattro possibili combinazioni  
4 (Tab. 3).  
5  
6

7 Poiché spesso la variabilità costante si accompagna a una variabilità proporzionale in un altro  
8 ambito di risultati, è consigliabile utilizzare più tipi di diagramma.  
9

10  
11 Accade spesso che in alcune parti dell'intervallo analitico si ottengano pochi risultati. Tipicamente,  
12 si tratta delle regioni con i livelli più elevati. In questi casi si consiglia la rappresentazione in  
13 diagrammi utilizzando il rango dei risultati in luogo del valore effettivamente ottenuto con la  
14 procedura comparativa o della media dei valori delle due procedure candidata e comparativa. Il  
15 rango del singolo risultato è costituito dalla sua posizione nella seriazione non decrescente di tutti i  
16 risultati dal minore al maggiore.  
17  
18  
19  
20

## 21 **T1** Analisi quantitativa dei risultati sperimentali

### 22 **T2** Stima dello scarto sistematico (*bias*) dal diagramma delle differenze

23  
24 L'analisi può essere condotta sia sul diagramma delle differenze sia su quello di dispersione. Per il  
25 Laboratorio che intende fare una verifica con almeno 40 campioni, è sufficiente l'analisi del grafico  
26 delle differenze. Con uno scarto tipo costante nell'intervallo analitico, la media delle differenze  
27 costituisce la stima dello scarto sistematico (*bias*). La rappresentazione delle differenze come  
28 istogramma o come "mountain plot" [4] secondo CLSI EP21 [5] può aiutare l'interpretazione dei  
29 risultati. Nel caso di distribuzioni fortemente asimmetriche, è preferibile utilizzare la mediana  
30 invece della media aritmetica.  
31  
32  
33  
34

35  
36 *NOTA STATISTICA. Il "mountain plot" è ottenuto plottando sull'asse delle ascisse le differenze tra*  
37 *i due metodi e sull'asse delle ordinate il rispettivo percentile da 1 a 50 e poi il complemento a 100*  
38 *del percentile in modo che, per esempio, in corrispondenza della 99<sup>a</sup> differenza il percentile sia*  
39 *ancora 1, come quello della prima differenza. Chi ha proposto questo plot [6] riporta che: (1) ha il*  
40 *vantaggio di visualizzare più facilmente la parte centrale (95%) dei dati, (2) è più facile stimare il*  
41 *percentile per differenze maggiori del 95%, (3) la forma del plot, a differenza di un istogramma,*  
42 *non è funzione degli intervalli, (4) rende più facile il confronto tra differenti distribuzioni e (5) è più*  
43 *facilmente interpretabile rispetto ai grafici delle distribuzioni empiriche cumulative. Ovviamente,*  
44 *non si dovrà confondere il "mountain plot" con un tradizionale istogramma.*  
45  
46  
47

48  
49 Nel caso di un comportamento con CV costante, lo scarto sistematico (*bias*) viene stimato dalla  
50 media delle differenze relative percentuali o, se la distribuzione è molto asimmetrica, dalla loro  
51 mediana.  
52

53  
54 Nelle situazioni miste, con DS costante in un certo intervallo e CV costante in un altro intervallo, è  
55 necessario identificare il rango (k, diciamo) del risultato che divide le due zone, ovvero la sua  
56 posizione nella seriazione non decrescente, e, quindi, calcolare il bias con il metodo appropriato per  
57 ciascuna zona. Almeno 20 campioni per zona sono necessari per considerare la stima  
58 ragionevolmente affidabile. La stima deve essere accompagnata dall'intervallo dei risultati a cui si  
59 riferisce.  
60  
61  
62  
63  
64  
65

1 Se il *bias* si modifica in funzione del livello del risultato, è necessario effettuare l'analisi della  
2 regressione partendo dal diagramma delle differenze. Può accadere anche che la relazione tra  
3 differenze e risultati non sia lineare: in questo caso è necessario utilizzare l'analisi della regressione  
4 non lineare.  
5  
6

7 I risultati "solitari" (*outlier*) devono innanzitutto essere confermati con l'analisi statistica e, quindi,  
8 attribuiti a una causa per essere trattati come "aberranti". L'analisi si basa sul residuo  
9 "studentizzato" estremo generalizzato (ESD). Fissato un livello di significatività (0,05 o 0,01), si  
10 ottiene il numero dei possibili risultati *outliers* (h): per esempio, al 5% di 44 risultati, h è pari a 2.  
11 Quindi, si calcola la deviazione dalla media in unità di scarto tipo (scarto quadratico medio, DS) per  
12 ciascun potenziale *outlier*:  $EDS_1 = \max (|d_j - d|)/DS$ .  
13  
14  
15

16 Va poi individuato il valore del quantile della distribuzione t di Student per i gradi di libertà (v) pari  
17 a N-i-1, dove i è la sequenza del risultato outlier (i = 1, 2, ..., h), e per la probabilità  $p = 1 - \alpha / 2(N -$   
18  $i + 1)$ . Con questo valore di  $t_{v,p}$  si ottiene il valore critico  $\lambda$  pari a  $t_{v,p} (n-i)/\sqrt{(n-i+1)(v + t_{v,p}^2)}$   
19 [EP09-A3 Appendix B Detecting aberrant results (outliers)].  
20  
21  
22  
23  
24  
25

## 26 **T2 Stima dello scarto sistematico (*bias*) dal diagramma di dispersione**

27 L'analisi della regressione è obbligatoria per il produttore del sistema analitico, mentre per il  
28 Laboratorio è facoltativa, anche se raccomandata (Tab. 4).  
29  
30

31 La regressione lineare ordinaria (ORL) serve per ottenere l'equazione tipo  $y = a + bx$  che mette in  
32 relazione i risultati del metodo candidato con il metodo comparativo se si verificano tre precise  
33 condizioni: scarto tipo (scarto quadratico medio, DS) costante, distribuzione dei risultati omogenea  
34 nell'intervallo di misura e coefficiente di determinazione ( $r^2$ , il quadrato del coefficiente di  
35 correlazione r) pari o superiore a 0,95.  
36  
37  
38

39 È importante ricordare che  $r^2$  superiore a 0,95 non significa un'elevata concordanza tra i due  
40 metodi, ma solo l'accettabilità di procedere calcolando una regressione lineare ordinaria. Non deve  
41 essere nemmeno utilizzato come indicatore della linearità, un tema affrontato dalla linea guida CLSI  
42 EP06 [7].  
43  
44  
45

46 Un'importante limitazione dell'ORL è il fatto che le variabili Y (corrisponde al metodo candidato) e  
47 X (corrisponde al metodo comparativo) siano misurate "senza errore", cosa che, evidentemente, non  
48 è proponibile nel contesto del confronto tra metodi di laboratorio. Tra l'altro, la regressione  
49 "classica" (iniziale formulazione di questo modello statistico) prevede che la variabile indipendente  
50 X sia anche "fissa" e, quindi, una variabile non casuale, senza una distribuzione di probabilità di  
51 riferimento. Per tenere conto dell'errore di misura, si deve ricorrere a modelli di regressione per  
52 "errori di misura" ("*measurement error models*") di cui il metodo proposto da Deming in cui la  
53 variabilità dovuta all'errore di misura è posta uguale per i due metodi (rapporto pari a 1, regressione  
54 ortogonale). Questi metodi si basano sul minimizzare la distanza dei punti dalla retta (proiezione  
55 ortogonale) invece della distanza verticale dei punti dalla retta della regressione OLR. Possiamo  
56 facilmente renderci conto della cosa invertendo gli assi della comparazione: la regressione di  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 Deming fornisce lo stesso risultato, quella ORL invece no.  
2

3 Per calcolare la regressione di Deming (stime dei parametri) sono disponibili diversi pacchetti  
4 statistici anche online. Per esempio “Had2Know.com” [8], e “Peltier Tech Blog” [9], nonché diversi  
5 pacchetti statistici commerciali come Analyse-it [10], EP Evaluator [11], GraphPad Prism [12],  
6 MedCalc [13], NCSS (*statistical software*) [14], R [15], S-PLUS [16] e StatsDirect [17].  
7  
8

9 Se la relazione tra i due metodi presenta costante il coefficiente di variazione (ovvero DS  
10 proporzionale), non si possono utilizzare né la regressione ORL né quella di Deming; si deve usare  
11 la regressione ponderata dei minimi quadrati (WLS). Il peso è fornito dall’inverso del quadrato del  
12 risultato riportato sull’asse delle ascisse, in modo che i punti sulla destra, affetti da maggiore  
13 variabilità, pesino meno di quelli sulla sinistra. La WLS ha però gli stessi limiti descritti per la  
14 ORL, che possono essere superati dalla procedura ponderata di Deming. In questa procedura, a ogni  
15 valore viene assegnato un peso ( $w$ ) che tiene conto del risultato del metodo candidato e del risultato  
16 del metodo comparativo, unitamente al rapporto tra gli errori di misura.  
17

18 In questo modo, i valori sulla destra pesano meno di quelli sulla sinistra del diagramma. La  
19 procedura deve essere ripetuta con diversi valori dei pesi fino a quando le differenze tra le stime  
20 della pendenza e dell’intercetta ottenute dalla regressione attuale e quella precedente siano  
21 praticamente nulle [EP09-A3, *Appendix H constant CV (weighted) Deming regression*].  
22

23 Non è possibile inserire nella regressione valori di concentrazione pari a zero né per il metodo  
24 candidato né per il metodo comparativo.  
25

26 Gli approcci di Deming si basano sulla comparazione di risultati in singolo. Se la comparazione  
27 utilizza medie di replicati, si può ottenere una stima del rapporto tra le imprecisioni dei due metodi,  
28 identificata con  $\lambda$ . Infine, se le due imprecisioni non sono note,  $\lambda$  viene posto pari a 1, come  
29 proposto da Deming.  
30

31 La trasformazione dei risultati come logaritmi o funzioni di potenza può dare in alcuni casi la  
32 sensazione di una distribuzione omogenea nell’intervallo. Può servire per stimare lo scarto  
33 sistematico (*bias*) ma non per valutarne l’accettabilità (v. oltre).  
34

35 La regressione proposta da Passing-Bablok [18] [19] [20] richiederebbe un numero di risultati  
36 superiore a quello indicato di 100 per il produttore e 40 per il Laboratorio.  
37

38 *NOTA STATISTICA. In generale, una procedura non parametrica ha bisogno di un numero*  
39 *maggiore di soggetti, ma ciò dovrebbe essere verificato nel caso specifico.*  
40

41 L’applicazione di tecniche statistiche sofisticate non dovrebbe prescindere dall’esame visivo del  
42 diagramma delle differenze. Se da questo emerge la combinazione di variabilità costante (DS) e  
43 proporzionale (CV), allora la regressione secondo Deming diventa meno valida. In questo caso, la  
44 regressione non parametrica secondo Passing-Bablok dà risultati che si avvicinano maggiormente  
45 a quelli del diagramma delle differenze.  
46

47 Infine va sottolineato che le regressioni lineari, di qualunque tipo, non possono risolvere il problema  
48 di una relazione non lineare.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## T2 Intervallo di confidenza dello scarto sistematico (*bias*)

Con le ORL semplice e ponderata, il calcolo dell'intervallo di confidenza al 95% è immediato, dagli errore standard della pendenza, dell'intercetta e di un valore fittato che considerano la varianza residua della regressione (scarto tipo dei residui:  $s_{yx}$ ).

Con gli altri modelli di regressione sono state proposte stime degli errori standard mediante tecniche iterative come la tecnica *jackknife*. Il metodo *jackknife* è una procedura di ricampionamento utilizzata in statistica per stimare l'errore standard di una grandezza. L'idea fondamentale sottostante il metodo *jackknife* sta nello stimare più volte il parametro di interesse escludendo dal campione un'osservazione alla volta.

Per la regressione di Passing-Bablok, in cui non è appropriato ricorrere al metodo *jackknife*, è consigliata la procedura *bootstrap*. Il *bootstrap* (o *bootstrapping*) è un metodo statistico per stimare la distribuzione di campionamento di uno stimatore. Si tratta di ottenere un elevato numero di campioni campionando con sostituzione dall'originale insieme di dati e, quindi, di combinare opportunamente le stime calcolate da tutti i campioni ottenuti.

Per i Laboratori, a differenza dei produttori, è più interessante la stima degli intervalli di confidenza dallo scarto stimato con il diagramma delle differenze (d). Con scarto tipo costante (scarto quadratico medio, DS) o proporzionale (CV), ma simmetrico, è sufficiente calcolare l'errore standard della media delle differenze ( $se_d = SD/\sqrt{N}$ ), da moltiplicare per il quantile della distribuzione t di Student con i pertinenti gradi di libertà scelto in accordo al livello di confidenza voluto. Per esempio, con 20 risultati (19 gradi di libertà), il 97,5° percentile della distribuzione t di Student è pari a 2093.

## T2 Interpretazione dei risultati del confronto in funzione dei criteri prefissati

Avendo ottenuto lo scarto sistematico (*bias*) e il suo intervallo di confidenza, resta da stabilire se il *bias* è accettabile o meno. Il criterio, ovvero il limite da non superare, va fissato in anticipo, consultando la letteratura tecnica e i clinici che utilizzano i risultati.

La differenza tra una procedura di misurazione comparativa e una procedura candidata a sostituirla è di interesse sull'intero intervallo di misurazione (usualmente espressa come la pendenza della retta di regressione) o a una o più concentrazione/i clinicamente rilevante/i. Il giudizio sulla concordanza tra i due metodi è basato sul confronto dell'intervallo di confidenza (IC) dello scarto misurato (*bias* medio sull'intervallo di misurazione o *bias* a una specifica concentrazione, vedasi la sezione 9.3 della EP09-A3) con i limiti del *bias* (distorsione) accettabile, definiti prima dello studio. Ciascun produttore o Laboratorio clinico dovrebbero sviluppare i propri criteri (previa consultazione dei clinici esperti sull'argomento e/o la letteratura tecnica). Questi criteri dovrebbero essere definiti prima di effettuare lo studio che confronta le procedure di misurazione, specialmente se lo studio sarà valutato da un'autorità regolatoria. Questi criteri dovrebbero includere le regole decisionali rispetto ai possibili risultati del *bias* e del suo IC 95% come riportato nella Fig. 1, in cui

ogni punto solido è un esempio di stima del *bias* e la linea verticale rappresenta il suo pertinente IC 95%.

Il risultato A è un esempio in cui l'IC 95% del *bias* misurato include lo zero. La conclusione può quindi essere che è stato osservato un *bias* non statisticamente significativo tra la procedura candidata e quella comparativa.

Per il risultato B, i limiti predefiniti dell'IC 95% includono il *bias* accettabile. Quindi, il *bias* della procedura candidata rispetto alla procedura di misurazione comparativa incontra i criteri di accettazione del *bias* con una confidenza del 95% ed è accettabile per l'applicazione definita.

Se il *bias* stimato è entro i limiti predefiniti ma IC no (risultato C), allora non si può affermare che i criteri di accettabilità sono stati soddisfatti al 95%. Poiché il *bias* stimato è inferiore del limite del massimo accettabile *bias*, si potrebbe concludere che il *bias* è accettabile. Tuttavia, è anche possibile una conclusione alternativa che una non appropriata percentuale di risultati campionari abbia un *bias* al di fuori i limiti.

Se il *bias* stimato è all'esterno dei predefiniti limiti del *bias* accettabile, ma IC no (risultato D), allora i dati al 95% di confidenza non presentano un *bias* oltre l'accettabile. Poiché l'IC include il limite del più alto accettabile *bias*, si può ancora concludere che il *bias* è accettabile. Tuttavia, vi è una minore confidenza su questa conclusione rispetto a quella conseguente al risultato C.

Se il *bias* stimato e il suo rispettivo IC sono fuori i limiti del *bias* accettabile (risultato E), allora la performance della procedura di misurazione candidata non è accettabile per la specifica applicazione.

Il risultato del confronto potrebbe indirizzare un Laboratorio clinico ad aggiustare gli intervalli di riferimento utilizzando i risultati dello studio di confronto riportato nel documento CLSI EP28 [21], invece di concludere che una procedura candidata di misurazione non è accettabile. Questo non sarà possibile per i livelli decisionali. Tali soglie sono state infatti stabilite da ampi studi clinici o sulla base dell'esperienza clinica, quindi non possono essere cambiate.

Se un produttore mette a disposizione dati di confronto per il metodo candidato, il *bias* può essere un criterio di accettazione del Laboratorio. Se l'affermazione del produttore per il *bias* è inclusa nell'IC 95% (risultato C e D), allora il Laboratorio clinico può concludere che la procedura di misurazione candidata soddisfa le affermazioni sul *bias* del produttore.

I casi A e B sono quindi esiti accettabili di comparazione. Il caso E è un risultato non accettabile. I casi C e D potrebbero essere accettabili, ma con confidenza inferiore al 95% e, nel caso D, abbastanza bassa.

CLSI EP31 (Punti 7.1, 7.2 e 7.3) fornisce indicazioni per l'individuazione dei limiti di accettabilità dei risultati di una comparazione. Il metodo ideale, ovvero l'utilizzo degli esiti clinici (*outcome*), è disponibile in pochissimi casi. Quello che si basa su opinioni dei clinici raccolte mediante questionari è interessante dal punto di vista teorico, ma soggetto a molte importanti limitazioni. La variabilità biologica fornisce un criterio apparentemente oggettivo, ma anch'esso presenta importanti limiti. Il principale vantaggio è la semplicità, poiché si assume che la differenza

1 accettabile debba essere non superiore a  $< 0,33 \times CV_i$ , ovvero un terzo della variabilità biologica.

2  
3 In qualche caso è possibile utilizzare le raccomandazioni pubblicate di esperti della professione. Per  
4 esempio, il *National Cholesterol Education Program* indica come *bias* accettabile tra metodi per il  
5 colesterolo totale un valore  $\leq 3\%$ . Questi limiti però non hanno basi statistiche né biologiche e  
6 spesso sono concepiti per la differenza con il valore di riferimento di taratura piuttosto che tra due  
7 metodi diagnostici.  
8  
9

10 Gli enti di autorizzazione e di accreditamento talvolta definiscono obiettivi analitici sia per  
11 l'imprecisione sia per l'inesattezza (*bias*), che derivano da combinazioni di risultati interlaboratorio,  
12 consigli delle aziende produttrici e opinioni di esperti [22] [23] [24] [25]. Si tratta di fonti molto  
13 autorevoli, ma il risultato rischia di riflettere più lo stato dell'arte che il vero obiettivo clinico.  
14  
15  
16

17 Infine, talvolta è possibile ricavare indicazioni dai risultati delle prove valutative interlaboratorio  
18 (PT, VEQ), prendendo quelli dei sistemi analitici simili. Purtroppo, non raramente le differenze  
19 osservate nelle VEQ sono dovute a effetti matrice. In altri casi, non è possibile sapere se la  
20 differenza sia stata compensata con aggiustamenti dei valori di riferimento o dei limiti decisionali.  
21  
22  
23

24 EP31 ammette inoltre che gli approcci tradizionali al confronto di risultati (t di Student, analisi della  
25 regressione) non si prestano facilmente a confronti multipli e richiedono elevate numerosità.  
26

27 EP31 raccomanda invece tecniche adatte a campioni piccoli. Il test del range o test dell'intervallo di  
28 variazione è costruito molto semplicemente calcolando le medie di ciascun gruppo di risultati, per  
29 esempio di ciascuno strumento, ricavando la differenza tra la maggiore e la minore ( $w = X_{\max} -$   
30  $X_{\min}$ ). La valutazione della differenza richiede la conoscenza della differenza critica (CD), della  
31 precisione nella serie (SR) e totale (ST). Con questi valori e con il numero dei sistemi da valutare si  
32 possono consultare le tabelle rese disponibili da EP31 nell'appendice B. Gli aspetti statistici del test  
33 del range sono riportati nell'appendice C (pag. 55-59). Esempi dettagliati dell'effettuazione di  
34 questo test sono riportati nell'appendice A (pag. 28-33).  
35  
36  
37  
38  
39

40 SR e ST possono essere ottenute dai dati del controllo di qualità. Se è necessario verificare intervalli  
41 di risultati dove i controlli non arrivano, allora va stimata l'imprecisione con un protocollo standard  
42 come CLSI EP05 [26], con lo schema sperimentale su 20 giornate. Lo schema semplificato di CLSI  
43 EP15 [27], ovvero 3 replicati per 5 giorni, non è altrettanto adatto.  
44  
45

46 Nell'appendice B della EP31 sono riportate specifiche tabelle per il confronto tra due fino a dieci  
47 metodi. I limiti dell'intervallo di variazione sono calcolati per un errore di I tipo (probabilità di  
48 rifiutare erroneamente una "vera" ipotesi di concordanza dei metodi, indicata con  $\alpha$ , corrispondente  
49 al livello di significatività di 0,05 e un errore di II tipo (probabilità di non rifiutare una "falsa"  
50 ipotesi di concordanza, indicata con  $\beta$ , corrispondente al complemento a 1 della potenza) almeno  
51 inferiore a 0,2, ovvero una potenza di almeno 0,80 (80%).  
52  
53  
54

55 Le tabelle devono essere utilizzate per ottenere il numero delle serie (*runs*, prima colonna), il  
56 numero di replicati (quarta colonna) e i limiti di rifiuto dell'intervallo di variazione (quinta colonna,  
57 un valore che moltiplica la differenza critica: CD) da utilizzare nella valutazione del confronto tra i  
58 metodi.  
59  
60  
61  
62  
63  
64  
65

1 Inoltre, le tabelle riportano (ultime due colonne) i valori della potenza per il caso in cui ci sia una  
2 situazione alternativa (di differenza tra i metodi) caratterizzata da una configurazione meno  
3 favorevole con il maggior potenziale errore di II tipo (uguaglianza di tutte le medie, eccetto due:  
4 una maggiore e l'altra inferiore) e la configurazione (più favorevole) in cui tutte le medie sono  
5 uguali eccetto una. Ovviamente, questi valori di potenza coincidono nel caso del confronto tra solo  
6 due metodi in cui le due suddette situazioni appunto coincidono. Per "potenza" si ricorda che in  
7 statistica si intende la probabilità di rifiutare l'ipotesi nulla dimostrando la presenza di una  
8 differenza quando l'ipotesi alternativa è vera (la differenza esiste).  
9

10  
11  
12  
13 Questo è il procedimento per usare le tabelle. Partendo dalla tabella pertinente al numero di metodi  
14 che si vuole confrontare, si comincia a considerare nella prima colonna il numero di *runs* pari a 1.  
15 Poi si cerca (seconda colonna) il valore del rapporto tra la differenza critica (CD) e la variabilità  
16 totale (scarto tipo, SD), definito CD/ST, che meglio corrisponde all'analisi che deve essere  
17 confrontato. In corrispondenza di questo valore si cerca nella colonna adiacente il valore del  
18 rapporto tra la variabilità entro serie (SR) e la variabilità totale (ST), definito SR/ST, che meglio  
19 corrisponde all'analisi da confrontare. A questo punto, sulla stessa riga si legge il numero di  
20 replicati necessari, il coefficiente che moltiplica la CD per il calcolo del limite di rifiuto del range e  
21 la potenza per le due situazioni di ipotesi alternativa considerate. Se non si trovano le combinazioni  
22 dei rapporti CD/ST e SR/ST, oppure il numero dei replicati non è praticabile, si ripete la ricerca per  
23 le righe con numero delle serie pari a 2. Trovata la riga ottimale, la casella nella colonna "limite di  
24 rifiuto del range" riporta il moltiplicatore della CD da utilizzare.  
25  
26  
27  
28  
29  
30

31 Per esempio, dalla Tabella B1 di CLSI EP31 (Confronto di due metodi), con una sola serie, rapporti  
32 CD/ST e SR/ST pari a 1 e numero di replicati pari a 16, si ottiene un limite di rifiuto dell'intervallo  
33 di variazione pari a  $0,70 \cdot CD$ . La potenza per dimostrare l'esistenza di una differenza tra i metodi è  
34 pari a 0,802, ovviamente per entrambi gli scenari (Tab. 5).  
35  
36  
37

38 La Tabella B1 termina con la riga di due serie, rapporto CD/ST pari a 4,00, SR/ST non rilevante,  
39 replicati solo uno, limite dell'intervallo  $0,49 \cdot CD$  per una potenza di 0,979.  
40

41 La Tabella B9 di CLSI EP31 fa riferimento al confronto di 10 metodi. In questo caso, con una sola  
42 serie, rapporti CD/ST e SR/ST pari a 1 e numero di replicati pari a 16, si ottiene un limite di  
43 accettabilità dell'intervallo pari a  $1,12 \cdot CD$ . In questa situazione le potenze sono diverse, ovvero  
44 0,807 per la differenza di un solo metodo, 0,473 per lo scenario meno favorevole.  
45  
46  
47

48 La Tabella termina con la riga di due serie, rapporto CD/ST pari a 4,00, SR/ST non rilevante,  
49 replicati solo uno, limite dell'intervallo  $0,80 \cdot CD$  per una potenza di 0,989 nello scenario migliore,  
50 0,834 in quello meno favorevole.  
51  
52  
53

## 54 **T1** Commutabilità dei materiali

55  
56 ISO 15189 (punto 5.6.4 *Comparability of examination results*) prescrive che nel caso particolare di  
57 risultati riferibili alla stessa origine metrologica, si possono considerare anche metrologicamente  
58 comparabili se i calibratori sono commutabili. CLSI EP09 (punto 6.1 *Study Samples*) afferma che  
59 diluizione, impoverimento o arricchimento ("*spiking*") dei campioni dovrebbero essere evitati,  
60  
61  
62  
63  
64  
65

1 poiché è stato dimostrato che i campioni così manipolati perdono la commutabilità tra metodi. CLSI  
2 EP31 (paragrafo 6.1 *Commutability*) prescrive che nella scelta dei materiali per la comparazione si  
3 cerchi innanzitutto di garantire la commutabilità, ovvero l'equivalenza matematica della relazione  
4 tra campioni dei pazienti e un determinato materiale tra metodi diversi. La commutabilità va  
5 verificata per tutti i materiali diversi dai campioni freschi dei pazienti.  
6  
7

8  
9 CLSI EP30 descrive le modalità per verificare la commutabilità dei materiali. Le caratteristiche  
10 minime di un materiale da usare per taratura e verifica dell'esattezza (*trueness*) sono omogeneità,  
11 stabilità, riferibilità metrologica e incertezza per il valore di una proprietà, e infine commutabilità.  
12  
13

14 La commutabilità viene validata utilizzando la regressione lineare, come nel confronto di metodi o  
15 procedure di esame (EP30 punto 7.3.1.2, *Regression Analysis Using the 95% Prediction Interval*).  
16 La regressione può essere quella ordinaria (ORL), quella di Deming o quella di Passing-Bablok.  
17 Alla regressione vanno applicati gli intervalli predittivi al 95%, calcolati dallo scarto tipo dei valori  
18 Y stimati con la regressione. Per questo calcolo può essere utilizzato anche l'approccio semplificato  
19 presentato da CLSI EP14 [28], che assume la mancanza di variabilità nelle misure del metodo di  
20 comparazione, sull'asse delle X.  
21  
22  
23

24 I materiali che danno risultati compresi nell'intervallo predittivo possono essere definiti  
25 commutabili.  
26  
27

## 28 **T1 Conclusioni**

29 Per soddisfare i requisiti di accreditamento il Laboratorio deve attrezzarsi con strumenti per valutare  
30 la comparabilità tra le procedure di esame o metodi, anche nel caso dei POCT. La comparazione nel  
31 caso del produttore è una vera e propria valutazione, con molti campioni e metodi statistici  
32 complessi. Nel caso del Laboratorio è invece una verifica, con meno campioni e metodi di analisi  
33 dei risultati più semplificati. La rappresentazione grafica dei dati ottenuti è di fondamentale  
34 importanza. L'interpretazione dei risultati deve essere basata su criteri di prestazioni (performance)  
35 conosciuti in anticipo. CLSI EP31 mette a disposizione tabelle per l'interpretazione che, nonostante  
36 l'apparente complessità, sono in definitiva utilizzabili abbastanza facilmente.  
37  
38  
39

40 In conclusione, l'attività di confronto di metodi è uno dei principali impegni gestionali del  
41 Laboratorio medico, a cui vanno dedicate notevoli risorse umane e soprattutto culturali.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**Conflitti di interessi** Nessuno

**Figura 1** Interpretazione dei risultati del confronto: scarto (bias) e intervalli di confidenza con limiti di accettabilità (modificata dalla Fig. 20 di EP9-A3).

## Bibliografia

- 
- 1 Martin RF (2000) General Deming regression for estimating systematic bias and its confidence interval in method-comparison studies. *Clin Chem* 46:100-104
  - 2 UNI ISO 3534-1:2000, Statistica - Vocabolario e simboli - Probabilità e termini statistici generali. Milano: Ente Nazionale Italiano di Unificazione
  - 3 UNI CEI ENV 13005:2000. Guida all'espressione dell'incertezza di misura. Milano: Ente Nazionale Italiano di Unificazione
  - 4 Krouwer JS, Monti KL (1995) A simple, graphical method to evaluate laboratory assays. *Eur J Clin Chem Clin Biochem* 33:525-527
  - 5 Clinical and Laboratory Standards Institute (CLSI) (2003) Estimation of Total Analytical Error for Clinical Laboratory Methods; Approved Guideline. CLSI document EP21-A (ISBN 1-56238-502-X). Clinical and Laboratory Standards Institute, 950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087 USA
  - 6 Krouwer JS, Monti KL (1995) A simple, graphical method to evaluate laboratory assays. *Eur J Clin Chem Clin Biochem* 33:525-527
  - 7 Clinical and Laboratory Standards Institute (CLSI) (2003) Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline. CLSI document EP06-A. Wayne, PA: Clinical and Laboratory Standards Institute
  - 8 Had2Know.com. How to Minimize the Sum of Perpendicular Offsets. Deming Linear Regression <http://www.had2know.com/academics/deming-regression-perpendicular-distances.html> (Accesso 12 dicembre 2015)
  - 9 Peltier Tech Blog. Deming Regression Utility. <http://peltiertech.com/deming-regression-utility/> (Accesso 12 dicembre 2015)
  - 10 Analyse-it Software, Ltd. The Tannery, 91 Kirkstall Road, Leeds, LS3 1HS, United Kingdom. <http://analyse-it.com/> (Accesso 12 dicembre 2015)
  - 11 EP Evaluator. Data Innovations Europe. 34 av. Jacques Brel, 1200 Brussels, Belgium. <https://www.datainnovations.com/products/ep-evaluator> (Accesso 12 dicembre 2015)
  - 12 GraphPad Prism. GraphPad Software, Inc. 7825 Fay Avenue, Suite 230, La Jolla, CA 92037 USA. <http://www.graphpad.com/scientific-software/prism/> (Accesso 12 dicembre 2015)
  - 13 MedCalc MedCalc Software bvba, Acaciaaan 22, 8400 Ostend, Belgium <https://www.medcalc.org/> (Accesso 12 dicembre 2015)
  - 14 NCSS (statistical software). NCSS, LLC, 329 North 1000 East, Kaysville, Utah 84037, USA <http://www.ncss.com/>. (Accesso 12 dicembre 2015)
  - 15 R. The R Foundation for Statistical Computing c/o Institute for Statistics and Mathematics, Wirtschaftsuniversität Wien, Welthandelsplatz 1, 1020 Vienna, Austria <https://www.r-project.org/> (Accesso 12 dicembre 2015)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

---

16 S-PLUS. TIBCO Software Inc. 3303 Hillview Avenue, Palo Alto, CA 94304 USA. <http://www.tibco.com/> (Accesso 12 dicembre 2015)

17 StatsDirect. StatsDirect Ltd, 9 Bonville Chase, Altrincham, CHESHIRE WA14 4QA, UK. <http://www.statsdirect.com/> (Accesso 12 dicembre 2015)

18 Passing H, Bablok W (1983) A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem* 21:709-720

19 Passing H, Bablok W (1984) Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part II. *J Clin Chem Clin Biochem* 22;431-445

20 Bablok W, Passing H, Bender R, Schneider B (1988) A general regression procedure for method transformation. Application of linear regression procedure for method comparison studies in clinical chemistry. Part III. *J Clin Chem Clin Biochem* 26;783-790

21 Clinical and Laboratory Standards Institute (CLSI) (2010) Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. CLSI document EP28-A3c. Wayne, PA: Clinical and Laboratory Standards Institute

22 Petersen PH, Hørder M (1988) Ways of assessing quality goals for diagnostic tests in clinical situations. *Arch Pathol Lab Med* 112:435-443

23 Skendzel LP, Barnett RN, Platt R (1985) Medically useful criteria for analytic performance of laboratory tests. *Am J Clin Pathol* 83:200-205

24 Petersen PH, Fraser CG, Westgard JO, Larsen ML (1992) Analytical goal-setting for monitoring patients when two analytical methods are used. *Clin Chem* 38:2256-2260

25 Ricós C, Iglesias N, García-Lario JV et al (2007) Within-subject biological variation in disease: collated data and clinical consequences. *Ann Clin Biochem* 44(Pt 4):343-352

26 Clinical and Laboratory Standards Institute (CLSI) (2014) Evaluation of Precision of Quantitative Measurement Procedures; Approved Guideline—Third Edition. CLSI document EP05-A3. Wayne, PA: Clinical and Laboratory Standards Institute

27 Clinical and Laboratory Standards Institute (CLSI) (2014) User Verification of Precision and Estimation of Bias; Approved Guideline—Third Edition. CLSI document EP15-A3. Wayne, PA: Clinical and Laboratory Standards Institute

28 Clinical and Laboratory Standards Institute (CLSI) (2014) Evaluation of Commutability of Processed Samples; Approved Guideline—Third Edition. CLSI document EP14-A3. Wayne, PA: Clinical and Laboratory Standards Institute

## TABELLE

### **Tabella 1. I riferimenti normativi della linea guida**

---

CLSI. *Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition*. CLSI document EP09-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2013

CLSI. *Verification of Comparability of Patient Results Within One Health Care System; Approved Guideline (Interim Revision)*. CLSI document EP31-A-IR. Wayne, PA: Clinical and Laboratory Standards Institute; 2012

CLSI. *Characterization and Qualification of Commutable Reference Materials for Laboratory Medicine; Approved Guideline*. CLSI document EP30-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2010

UNI EN ISO 15189:2013 Laboratori medici - Requisiti riguardanti la qualità e la competenza. Milano: UNI 2013

UNI EN ISO 22870 Analisi decentrate (*point-of-care testing*, POCT) Requisiti per la qualità e la competenza Milano: UNI 2006

---

### **Tabella 2. Numerosità dei campioni per il confronto tra metodi (secondo CLSI EP09-A3)**

Campioni per il produttore	$\geq 100$
Campioni per il laboratorio	$\geq 40$
Campioni manipolati (" <i>minipools</i> ")	$\leq 20\%$

---

**Tabella 3. Equazioni per i diagrammi delle differenze**

Asse orizzontale (z)	Asse verticale (d)	
	Differenza costante nell'intervallo (DS costante)	Differenza proporzionale nell'intervallo (CV costante)
Procedura di esame comparativa	$z_i = d_i = y_i - x_i$	$z_i = x_i$ $d_i = (y_i - x_i) / x_i$
Media tra comparativa e candidata	$z_i = (y_i + x_i) / 2$ $d_i = y_i - x_i$ (Bland-Altman)	$z_i = (y_i + x_i) / 2$ $d_i = (y_i - x_i) / [(y_i + x_i) / 2]$

$x_i$  è il risultato della procedura comparativa per l'i-esimo paziente;  $y_i$  è il risultato della procedura candidata per l'i-esimo paziente; ( $z_i$ ,  $d_i$ ) sono il risultato sulle asse delle coordinate per l'i-esimo paziente.

**Tabella 4.** Stima dello scarto sistematico (bias) [secondo CLSI EP09-A3]

<b>Attore</b>	<b>Obiettivo</b>	<b>Procedura</b>
Produttore	Validazione della procedura di esame	Analisi statistica della regressione
Laboratorio	Verifica della procedura di esame	Analisi del diagramma delle differenze

**Tabella 5.** Estratto dalla Tabella B1 di CLSI EP31. Confronto di due metodi

<b>Serie</b>	<b>CD/ST</b>	<b>SR/ST</b>	<b>Numero dei replicati</b>	<b>Limite di rifiuto del range</b>	<b>Potenza scarto di metodo</b>	<b>di Potenza un configurazione meno favorevole</b>	<b>della</b>
1	1,00	1,00	16	$0,70 \cdot CD$	0,802	0,802	
1	1,25	1,00	11	$0,67 \cdot CD$	0,833	0,833	
...							
2	2,50	0,50	6	$0,70 \cdot CD$	0,800	0,800	

CD, differenza critica; SR, scarto tipo (scarto quadratico medio) nella serie; ST, scarto tipo (scarto quadratico medio) totale.