



**Raccomandazioni dell'Associazione Tecnico Scientifica – S.I.T.La.B**  
***Gruppo di lavoro Biochimica clinica e Patologia clinica***

N.055/23

**L'incertezza di misura nel laboratorio clinico:  
casi particolari di stima della precisione analitica nei test quantitativi**

E. Cerelli (Reggio Emilia), A. D'Addiego (Teramo), E. Iannello (Messina), V. Nocito (Cosenza),  
A. Ricotta (Agrigento), I. Rossini (Varese), D. F. Santarcangeli (Milano)

**Rev. 1.0**

**SITLaB news**

**Publicato: 26 Ottobre 2023**

**Copyright: © SITLaB**

## Avvertenze

- I. Per evitare le ambiguità semantiche presenti nell'uso dei vocaboli metrologici, il lessico utilizzato in questo documento fa riferimento alla: “*Guida Eurachem - Terminologia per le misurazioni analitiche – Introduzione al VIM 3 (International Vocabulary of Metrology)*” e al glossario comparato inglese/francese/italiano proposto da SIPMeL: “*Recommendations for the glossary in the new ISO standards for medical laboratories on calibration, labels and quality control (ISO 17511, ISO 18113, and 15198)*”.
- II. Nel documento, a parte alcuni richiami indispensabili per la comprensione del testo, non sono trattate le basi matematiche della stima delle incertezze in metrologia, né i fondamenti teorici, informatici e grafici del controllo statistico di qualità nel laboratorio clinico, reperibili in ottimi manuali e linee guida (vedi bibliografia essenziale).
- III. In allegato, un “tutorial” in formato EXCEL ([SITLAB-precisione.xls](#)) protetto da modifiche accidentali, ma senza alcuna password e quindi liberamente modificabile, consente di verificare le funzioni utilizzate e di sperimentare i casi pratici qui esaminati, adattandoli alla propria pratica quotidiana. L'allegato è possibile trovarlo al seguente link:

<https://docs.google.com/spreadsheets/d/1x0MIdHchX1Ba6INMgHEEmB4kELINNmKJ/edit?usp=sharing&ouid=116629742702071009141&rtpof=true&sd=true>

## Sommario

La “garanzia” di Qualità nel laboratorio clinico .....	4
Problemi di carattere generale relativi alla stima dell'incertezza .....	6
Il controllo di qualità analitico (QC) nel laboratorio clinico .....	7
Breve nota sull'importanza del materiale di controllo.....	8
Il Controllo Interno della Qualità analitica (IQC).....	9
Quale precisione misurare e quale precisione è utile per la clinica? .....	9
Conclusioni .....	12
Pratica di laboratorio.....	13
Caso 1– Precisione intermedia: un analizzatore, due o più lotti di materiale di controllo con breve scadenza .....	13
Caso 2 –Riproducibilità intralaboratorio e interlaboratorio: due o più analizzatori che utilizzano lo stesso lotto di materiale di controllo .....	15
Analisi della varianza (ANOVA) univariata.....	15
Caso 3 – Precisione in assenza di un materiale di controllo stabile e/o commutabile: un analizzatore, utilizzo di replicati di campioni umani .....	18
Bibliografia .....	19

## Abstract

La capacità di produzione continua nei moderni analizzatori, se non adeguatamente governata, può comportare che i risultati siano forniti ai clinici con un TAT (*Turn Around Time*) molto ridotto, ma anche con una accuratezza di misura più ipotizzata che effettivamente verificata, con l'imbarazzante eventualità del richiamo di molti risultati ritenuti, a posteriori, non corretti (vedi lo scandalo in USA della start-up "*Theranos*", nel 2015).

Infatti è la conoscenza dell'incertezza effettiva (esattezza e precisione) associata al risultato di un test che rende la misura clinicamente affidabile: incertezza che andrebbe dichiarata nei documenti del TQM, per chiunque ne faccia richiesta.

Ma mentre l'esattezza si scontra ancora con la carenza di produttori e di materiali di riferimento certificati e commutabili, la stima della precisione ha storicamente rappresentato la "*core competence*" del Controllo Interno della Qualità Analitica (IQC), tramite l'analisi statistica di osservazioni ripetute di campioni di controllo commutabili in condizioni definite.

Tuttavia, nonostante l'industria fornisca comode soluzioni preconfezionate per il monitoraggio di questo fondamentale parametro analitico, la scarsa flessibilità dei software commerciali generalisti a volte impedisce di gestire correttamente le situazioni organizzative e di "flusso" dei campioni meno comuni fornendo, di conseguenza, valutazioni dell'incertezza non aderenti alla realtà e poco utili ai fini del giudizio clinico.

A tal riguardo, utilizzando il lessico internazionale della metrologia, in questo documento si esamineranno alcuni contesti di lavoro particolari ove la stima puntuale e aggiornata della precisione analitica evidenzia ancora criticità e lacune, proponendo qualche soluzione di facile implementazione nella pratica quotidiana.

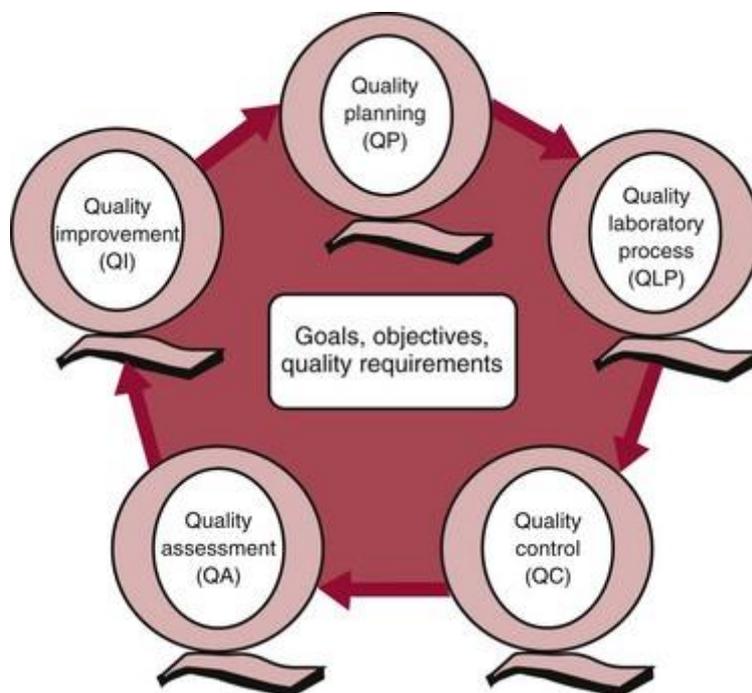
## La “garanzia” di Qualità nel laboratorio clinico

Il “Controllo totale della qualità” (*Total quality control - TQC*) in un laboratorio clinico, detto anche “Governo totale della qualità” (*Total quality management – TQM*), può essere definito come un processo ciclico (idealmente senza fine) avente come scopo l’azzeramento, o più verosimilmente, la riduzione al minimo delle “non conformità” agli standard organizzativi, analitici e clinici pianificati dal laboratorio, sulla base dello stato dell’arte più aggiornato.

Nella pratica il TQC si concretizza in un insieme di tecniche ad ampio raggio, per la maggior parte derivate da metodologie sviluppate per la pianificazione, il controllo e il miglioramento dei processi industriali. Al riguardo occorre sottolineare come nella comunità scientifica clinico-laboratoristica manchi ancora un consenso unanime su come stabilire le APS (*Analytical performance specifications*), cioè gli obiettivi clinico-analitici necessari per fornire una risposta adeguata alle esigenze dei pazienti, e su come applicare alla pratica medica alcune di queste tecniche (ad esempio, se sia appropriato utilizzare la metrica Sigma per il calcolo degli “Indici di capacità di processo Cp e Cpk” (*Process Capability*), utili per valutare le opportunità intrinseche ad un processo analitico stabile di non commettere errori clinicamente significativi (i cosiddetti *DPMO - difetti per milioni d’opportunità di commettere errori*).

La metodologia del TQC dovrebbe rappresentare lo strumento essenziale per garantire il **valore clinico** del dato di laboratorio, ovvero la sua reale utilità per la prevenzione, la diagnosi, il monitoraggio e la terapia delle malattie.

I suoi principi sono stati sintetizzati da James O. Westgard e altri in uno schema di processo detto delle **cinque Q** (figura 1).



**Figura 1.** Schema delle “cinque Q” di gestione della qualità totale (TQM) in un laboratorio sanitario. (Westgard JO, Burnett RW, Bowers GN. *Quality management science in clinical chemistry: a dynamic framework for continuous improvement of quality.* Clin Chem 1990;36:1712-6)

**QP** (*Quality planning*) rappresenta la fase di pianificazione; **QLP** (*Quality laboratory process*) stabilisce politiche, pratiche e procedure per il corretto svolgimento delle attività laboratoristiche; **QC** (*Quality control*) e **QA** (*Quality assessment*), tramite indicatori quali-quantitativi, misurano quanto bene queste attività sono state eseguite (**QC utilizza la matematica e la statistica per tenere sotto controllo l'accuratezza dei risultati prodotti**, mentre QA monitora altre prestazioni del laboratorio, come il Turn-Around-Time (TAT), l'identificazione corretta dei campioni e dei pazienti, l'adeguatezza dei fattori ambientali (temperature, umidità, gap temporale dal prelievo alla processazione, ecc); **QI** (*Quality improvement*) fornisce gli strumenti per agire sia in senso correttivo, se tali misure evidenziano scostamenti quantitativi o non-conformità qualitative dagli standard prestabiliti, sia in senso migliorativo.

Lo schema delle "cinque Q" può anche essere ricondotto al più noto *ciclo PDCA* (**Plan, Do, Check, Act**), o ciclo del "miglioramento continuo" ("Kaizen" in giapponese), ideato e implementato da E. Deming nel Giappone degli anni 60-70 e utilizzato normalmente in molti ambiti di lavoro come base per decisioni obiettive.

E' bene sottolineare come le fasi del ciclo denominate QLP, QC e QA, storicamente enfatizzate nei laboratori clinici, siano deputate alla "valutazione" della qualità e non alla "garanzia" della medesima. Misurare le prestazioni, infatti, necessariamente non le migliora e spesso non rileva i problemi in tempo utile per prevenire esiti negativi. La garanzia della qualità richiede, quindi, che le cause dei problemi siano identificate sistematicamente, ed eliminate attraverso la ripianificazione continua della qualità stessa aprendo un nuovo ciclo. Inoltre le tecniche utilizzate per il QC e il QA devono essere in grado di rilevare i problemi in tempo reale o, comunque, prevenirne le conseguenze abbastanza precocemente da permettere di attivare azioni atte a riportare la situazione sotto controllo, prima dell'emissione dei referti.

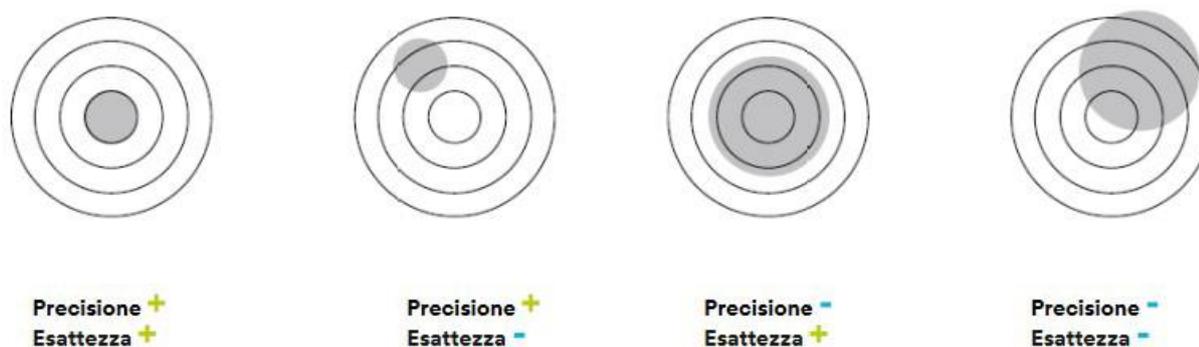
A tal proposito occorre notare come la capacità di produzione continua nei moderni analizzatori, se non adeguatamente governata, possa comportare che i risultati siano forniti ai clinici certamente con un TAT (*Turn Around Time*) molto ridotto, ma anche con una accuratezza più ipotizzata che effettivamente verificata. Occorre prestare molta attenzione a quest'aspetto, per evitare in primo luogo un rischio ai pazienti, ma anche l'imbarazzante eventualità del richiamo di molti risultati ritenuti, a posteriori, non corretti: lo scandalo della start-up USA "*Theranos*" nel 2015 è emblematico (vedi in bibliografia essenziale).

Si raccomanda che sia posta particolare cura agli algoritmi che regolano la trasmissione automatica dei risultati, affinché questa avvenga solo dopo le verifiche di affidabilità previste dall'IQC e dopo l'autorizzazione tecnica necessaria [cfr. <i>Raccomandazione SITLAB 9/20: L'Autorizzazione Tecnica (Technical Authorization) nel Laboratorio Biomedico: razionale e concetti-chiave</i> ].
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Problemi di carattere generale relativi alla stima dell'incertezza

Richiamiamo brevemente alcuni concetti fondamentali di metrologia.

L'accuratezza (*accuracy*) di un risultato di laboratorio (requisito che indica il suo avvicinarsi al valore "vero") è sempre "incerta", e dipende dalla capacità dei "sistemi di misura" (insieme di abilità umane, strumenti tecnologici e metodi chimico-fisici) di essere il più possibile esatti/giusti (*trueness*), cioè misurare effettivamente il valore "vero" - o considerato di "riferimento" - dell'analita, e precisi (*precision*), fornire risultati molto vicini in misure ripetute (*figura 2*).



**Figura 2.** Descrizione grafica dei due elementi dell'accuratezza: esattezza e precisione (Dr. Jan-Frederik Güth, Università di Monaco)

Ciò comporta che il risultato di una misura (singola o replicata) dovrebbe essere espresso abitualmente tramite due parametri: il **valore** strumentale osservato (*measurement result*) e la sua **incertezza di misura** (*measurement uncertainty*).

Quest'ultima, nelle sue componenti di inesattezza e imprecisione, caratterizza "l'errore" (senza alcuna accezione negativa!) insito nel dato di laboratorio, ed è definita sinteticamente da un numero espresso nella stessa unità di misura del risultato ovvero, se esiste proporzionalità col risultato, da una percentuale del valore stesso.

L'incertezza "estesa" (*expanded uncertainty*) indica l'intervallo dei valori entro i quali si colloca, con una probabilità definita (in genere il 95%), il risultato "vero" del misurando: più è piccola l'incertezza estesa, migliore è la misura.

<b>Risultato = Valore ± incertezza estesa</b> oppure <b>Risultato = Valore ± %</b>
------------------------------------------------------------------------------------

La conoscenza dell'incertezza estesa consente all'utilizzatore finale di avere informazioni sufficienti sull'affidabilità dell'analisi, soprattutto quando debba confrontarla con un limite prestabilito (ad esempio, con i valori di "riferimento normali" - *reference interval* o *reference range* - oppure con un limite stabilito per legge) ovvero con un risultato precedente (*delta-check*).

Nei laboratori clinici da anni si dibatte sull'utilità (e l'opportunità!) di riportare nel referto finale, assieme ad ogni risultato, la stima della sua incertezza.

Allo stato attuale non esiste alcun obbligo normativo in tal senso, e neppure "certificativo" su base volontaria. Anche le norme ISO 17025: "*General requirements for the competence of testing and calibration laboratories*" e le più specifiche ISO 15189: "*Medical laboratories - Requirements for quality and competence*" non impegnano in tal senso. Entrambe prescrivono che il laboratorio debba stabilire l'incertezza delle sue misure in fase di validazione dei metodi (secondo le specifiche tecniche della norma ISO/TS 20914: "*Medical laboratories — Practical guidance for the estimation of measurement uncertainty*") per essere consapevole dell'affidabilità dei propri risultati rispetto all'uso che se ne deve fare, ma che questa sia riportata nei rapporti di analisi solo quando sia richiesta

espressamente dall'utilizzatore finale oppure quando l'incertezza influisca sulla conformità a un limite normativo o di specifica (tipicamente indagini tossicologiche e forensi, di qualità ambientale, ecc).

Le due componenti dell'incertezza, esattezza e precisione, non sono stimabili con la stessa facilità. Mentre l'esattezza delle misure in campo clinico si scontra (anche per moltissimi misurandi che rivestono un valore fondamentale nella pratica clinica come, ad esempio, i marcatori tumorali), con la carenza di produttori (*RMP: Reference Material Producer*) e di materiali di riferimento certificati e commutabili (*CRM: Certified Reference Material*), la stima della precisione analitica ha storicamente rappresentato la "core competence" dell'IQC, tramite l'analisi statistica di osservazioni ripetute di campioni di controllo in condizioni definite.

E' comunque importante evidenziare come, nella pratica clinica quotidiana, un errore costante nell'esattezza possa essere "compensato" più facilmente dall'esperienza professionale rispetto a un aumento improvviso e significativo dell'imprecisione. Infatti, mentre non sempre è necessario e utile conoscere il valore "vero" di un analita per confrontarlo con un "riferimento" assoluto, l'ampiezza dell'oscillazione casuale di un risultato incide direttamente sia sul calcolo della "differenza critica" (*critical difference*), minima discrepanza considerata significativa tra due risultati ottenuti sullo stesso paziente in tempi diversi, sia soprattutto sul processo psicologico d'interpretazione ed elaborazione di questa differenza da parte dei curanti.

D'altro canto qualunque modificazione "accidentale" di un valore che sia generata semplicemente da cause tecniche, come la messa in uso di un nuovo lotto di reagenti, in particolare nei test immunometrici ed enzimatici, può essere confusa con una vera variazione nell'omeostasi del paziente ed esitare in trattamenti non appropriati.

Tale eventualità è più remota per i misurandi utilizzati per la diagnosi nel breve periodo piuttosto che per il monitoraggio a lungo termine. Sicuramente la troponina cardiaca nella diagnosi d'infarto miocardico soffre difficilmente di variazione tra i lotti, perché le misurazioni seriali sono eseguite in un lasso di tempo così breve da rendere improbabile (e inopportuno!) l'utilizzo di lotti diversi di reagente. Non così nel monitoraggio a lungo termine di misurandi come l'antigene prostatico specifico (PSA) o l'emoglobina glicata (HbA1c), fondamentali nella gestione del cancro alla prostata e del diabete mellito, dove l'estensione temporale rende quasi certo il cambio di lotto da una misura all'altra.

Questo esempio rende conto della diversa importanza che può assumere la conoscenza della variabilità del risultato nel breve, medio o lungo periodo, e comunque della necessità di mantenerla la più bassa e stabile possibile anche per molti anni.

## **Il controllo di qualità analitico (QC) nel laboratorio clinico**

Normalmente la gestione della qualità ANALITICA di un laboratorio clinico è effettuata tramite tre tipologie di programmi, complementari tra loro.

- Il classico controllo interno, (***Internal Quality Control – IQC***) comprendente le note regole di Westgard, le carte di Shewhart e altro, applicato anche più volte per ogni ciclo di analisi. L'IQC è finalizzato al monitoraggio temporale della "stabilità" dei risultati all'interno di un'incertezza di misura predefinita e certa: presupposto necessario per il rilascio dei risultati alla refertazione tramite l'autorizzazione tecnica (*technical authorization*) e, quindi, per il corretto uso clinico dei valori di riferimento (*biological reference interval*) e del *delta-check* nei controlli seriali analitici e di *follow-up*.
- La valutazione esterna di qualità (***External Quality Assurance - EQA***), promossa da un ente esterno ed eseguita su campioni occasionali per scopi di "*proficiency test*" (conosciuti anche come *Ring Test*) oppure per la valutazione dell'esattezza (considerando come riferimento il valore di "consenso" ottenuto da tutti i laboratori partecipanti). Inoltre, quando il materiale di

controllo ha concentrazione e incertezza certificate tramite una catena di riferibilità metrologica (traceability chain) a uno standard di alto livello, può essere utilizzato per la stima dello scostamento sistematico di misura (*measurement bias*) dal valore vero e quindi dell'accuratezza.

- Una terza variante che aggrega IQC ed EQA: il controllo di qualità interno multicentrico, o allargato (***Extended Quality Control - EQC***). Utilizzando giornalmente una sola matrice commerciale e tramite l'ausilio di software esperti remoti, interfacciati con il sistema informatico del laboratorio (*Laboratory Information System – LIS*), il gestore esterno EQC raccoglie ed elabora in tempo reale tutte le misure effettuate, fornendo report di IQC ed EQA molto utili per i laboratori affiliati.

IQC, EQA ed EQC, insieme alle procedure di validazione e riconvalida di metodi e test, costituiscono il pilastro della qualità analitica nel laboratorio clinico. Il loro fine è monitorare e mantenere nel lungo periodo la stabilità dei sistemi di analisi, facendo in modo che i fattori chiave determinanti una buona qualità di misura - la capacità di determinare effettivamente la sostanza d'interesse e una bassa oscillazione casuale tra misure ripetute - non cambino durante l'uso routinario.

## **Breve nota sull'importanza del materiale di controllo**

Per l'IQC, l'EQA e l'EQC un aspetto sovente sottovalutato è la scelta del materiale di controllo, comunemente di origine commerciale.

Esso deve possedere caratteristiche chimico-fisiche ben precise:

- una perfetta commutabilità (*commutability*) per fornire una risposta strumentale equivalente a quella dei campioni umani nello stesso ambiente di analisi, compresi gli "errori" di misura sistematici e casuali (ad esempio, densità e viscosità devono essere simili, per evitare aliquotazioni imperfette; l'effetto "matrice" deve essere comparabile; dovrebbero essere assenti substrati proteici o enzimatici di origine non-umana; ecc). La commutabilità è fondamentale per poter affermare che l'incertezza di misura provata sui campioni di controllo è la stessa di quella dei campioni umani.
- una durata dei singoli lotti, per omogeneità, quantità e stabilità, la più lunga possibile (almeno un anno di lavoro);
- l'assenza di stabilizzanti/conservanti potenzialmente interferenti con i metodi di analisi in uso nel laboratorio. Se presenti, devono essere sempre dichiarati dal produttore.

Si raccomanda di verificare sperimentalmente, in fase di validazione di ogni metodo e test, la commutabilità dei materiali di controllo, confrontandoli con replicati di campioni umani. Nell'IQC particolare cura deve essere dedicata alla verifica dell'omogeneità della varianza rispetto ai campioni umani nei pressi dei livelli decisionali critici (test *F di Fisher* o similari).

Si raccomanda che i controlli commerciali siano forniti da produttori terzi qualificati, indipendenti dai fornitori dei sistemi analitici utilizzati (strumentazione e reagenti).

## Il Controllo Interno della Qualità analitica (IQC)

Nonostante l'industria fornisca comode soluzioni preconfezionate, non è scontato ricordare quanto sia importante la "personalizzazione" dell'IQC in ogni laboratorio: soprattutto in funzione dell'**organizzazione, della quantità dei test eseguiti e dell'uso, a volte specialistico, previsto per i medesimi** (ad esempio, per studi epidemiologici; per screening su popolazione per definizione sana, come i donatori di sangue, o sicuramente malata; per gli utenti ospedalizzati o ambulatoriali; per fare diagnosi di conferma o esclusione di una malattia; per il monitoraggio di una terapia farmacologica; per la medicina d'urgenza; per il follow-up clinico; ecc).

Ogni laboratorio deve documentare per iscritto l'intero sistema del TQM, compreso l'IQC. Particolarmente dettagliata deve essere la parte che riguarda i procedimenti operativi, la loro validazione e rivalidazione, la strumentazione analitica e i materiali necessari, il modo di utilizzo e registrazione dei controlli di qualità, i criteri (regole) di accettabilità, le tecniche statistiche di valutazione, gli interventi previsti nelle situazioni di "fuori controllo" (gestione delle non conformità analitiche), modalità e durata di archiviazione di risultati e controlli.

L'obiettivo dovrebbe essere quello di conciliare le caratteristiche di un ottimo Sistema Qualità con le risorse umane, materiali ed economiche disponibili e con le esigenze cliniche dell'utenza di riferimento: insomma, mediare l'ottimo col buono tenendo conto che, a volte, il primo è nemico del secondo e che ogni teoria non è che un modello imperfetto della realtà.

Il TSLB, analista, controllore e conduttore dei sistemi di analisi, deve studiare a fondo l'IQC in uso nel laboratorio e le sue criticità al fine di fornire, con cognizione di causa, il benessere tecnico (*Technical Authorization*) necessario al rilascio dei risultati prodotti.

## Quale precisione misurare e quale precisione è utile per la clinica?

Ovviamente è il modello sperimentale adottato nell'IQC che rende conto del **tipo** di precisione misurata. E' noto, infatti, che esistono molteplici fonti di variabilità "aggiuntive" rispetto alla variabilità minima intrinseca riscontrabile in una sola serie analitica, detta **ripetibilità (repeatability)**. Alcune possono manifestare una **frequenza ricorrente basata sul tempo**, come il turnover del personale o la variazione oraria, settimanale o stagionale (ad esempio, per fluttuazioni non compensate nella temperatura e umidità ambientali), altre possono seguire uno **schema basato su eventi**, come le ricalibrazioni, il cambio dei lotti di reagenti e dei calibratori, l'usura e i guasti delle macchine. La conseguenza è comunque sempre una "perturbazione" osservabile nelle carte di controllo dell'IQC, indicante un cambiamento nella precisione o nell'esattezza di misura.

Sia le fonti del primo tipo (temporali ricorrenti) sia le seconde (eventi) possono ampliare gli errori casuali (aumento del *CV*, *coefficiente di variazione*) o introdurre scostamenti sistematici nel valore medio (anche di segno opposto, la cui risultante nel lungo periodo può essere un annullamento reciproco).

Tutte assieme caratterizzano la **precisione intermedia (intermediate precision)**, così chiamata perché si situa a metà strada tra la **ripetibilità intraserie** e la **riproducibilità intra- o inter-laboratori (reproducibility)**: in quest'ultima si valuta anche il contributo fornito dall'uso di strumentazione/metodi differenti all'interno dello stesso laboratorio o in laboratori diversi.

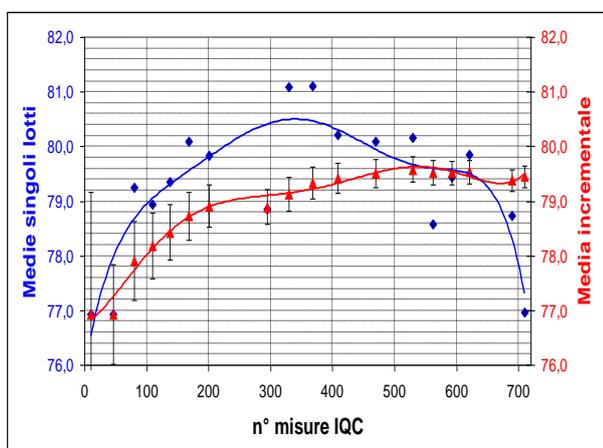
A titolo di esempio, i grafici di figure 3 e 4, relativi al controllo di qualità interno dell'enzima ALT nell'arco di 2 anni, mostrano le fluttuazioni nelle medie e nei CV osservate utilizzando un unico lotto stabile e commutabile di controllo commerciale. Le misure sono state fatte sempre sullo stesso analizzatore automatico multiparametrico, cambiando nei 2 anni di osservazione 16 lotti di reagente

ed effettuando alcune ricalibrizioni strumentali: ci troviamo dunque nel campo della “precisione intermedia” a lungo termine.

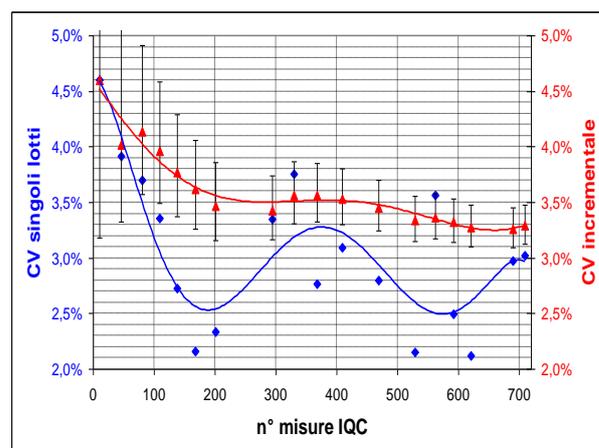
In figura 3 i rombi blu mostrano le medie calcolate per ogni lotto di reagente utilizzato, di numerosità variabile (interpolate dalla curva a “campana” di colore blu), mentre i triangoli di colore rosso, interpolati dalla linea di uguale colore, indicano la media incrementale da zero ai 710 dati raccolti, con il relativo intervallo di confidenza (*Confidence Interval - CI*) al 95% ( $p=0.05$ ; barre verticali). Lo scarto tra le medie riflette una dipendenza delle misure dal lotto di reagenti utilizzati, ma risente anche di altre variabili: sia di tipo “temporali” ricorrenti, come il cambio di temperatura stagionale non compensato sufficientemente dal sistema di termostattizzazione dell’analizzatore, sia basate su “eventi” più casuali, quali le ricalibrizioni periodiche, le manutenzioni, i guasti, ecc.

Parimenti, in figura 4 i rombi blu mostrano i CV calcolati per ogni lotto di reagente (interpolati dalla simil-sinusoidale di colore blu), mentre i triangoli rossi, interpolati dalla linea di uguale colore, indicano il CV incrementale dal primo all’ultimo lotto utilizzato, con il relativo CI al 95% ( $p=0.05$ ; barre verticali).

Dai due grafici si evince chiaramente sia lo scarto sistematico della media di ogni lotto verso la media cumulata sia la variabilità nella precisione tra i lotti (oscillazione del CV a ogni cambio di reagente). Per questo test occorrono almeno 300 dati cumulati per “normalizzare” l’andamento dei due parametri, ottenendo un intervallo di confidenza sufficientemente ristretto e stabile, comparabile con l’errore di arrotondamento all’ultima cifra significativa, l’unità, con la quale sono espressi i risultati. Ciò rende conto del perché sia richiesto generalmente un anno di misure per definire, in condizioni di stabilità, la precisione intermedia di un metodo.

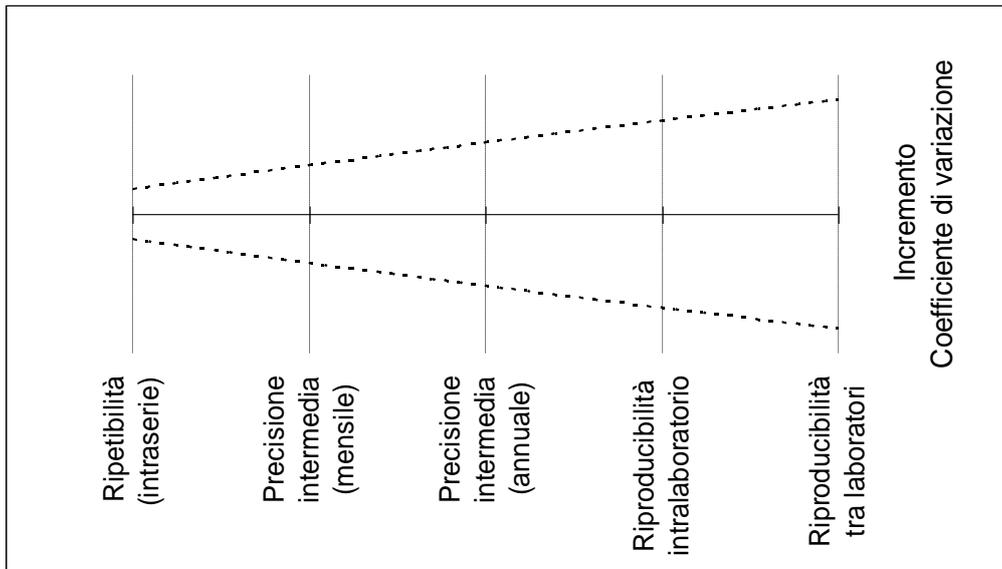


**Figura 3.** Andamento nel tempo, in U/L, delle medie relative a ogni lotto di reagente (rombi blu) e alla media incrementale (triangoli rossi). Le barre indicano l’intervallo fiduciale (CI) al 95%.



**Figura 4.** Andamento nel tempo dei CV relativi a ogni lotto di reagente (rombi blu) e al CV incrementale (triangoli rossi). Le barre indicano l’intervallo fiduciale (CI) al 95%.

Riassumendo, come schematizzato in figura 5, la precisione analitica misurata su un singolo analizzatore, sia intraserie sia nel breve periodo (giorni o settimane), può solo approssimare, per via di un minor numero di dati utilizzati e un minor numero di variabili ambientali e analitiche sperimentate, quella misurata per lungo tempo (tipicamente annuale) in condizioni di stabilità del metodo (e possibilmente per valori vicini ai livelli decisionali clinici). Parimenti anche la precisione intermedia sarà diversa, e normalmente migliore, di quella derivante dall’utilizzo di più analizzatori indipendenti (*riproducibilità nel laboratorio o tra laboratori*).



**Figura 5:** Quando aumentano le occasioni di variabilità nell'ambiente di misura, passando da condizioni di ripetibilità a condizioni di riproducibilità (secondo la terminologia VIM), il coefficiente di variazione (CV) generalmente aumenta.

## Conclusioni

Da quanto detto, appare chiaro come abbia poco significato il dichiarare per un test un certo valore di precisione senza indicare in quali condizioni specifiche (*specified precision conditions*) si sia ottenuta quella performance: se in condizioni di ripetibilità, di precisione intermedia o di riproducibilità (*specified condition of repeatability, intermediate precision or reproducibility*).

Infatti, anche se la sola componente della precisione in condizioni di ripetibilità può rappresentare un'evidenza di conformità nelle verifiche di parte terza, è la riproducibilità delle misure nel suo insieme (cioè quella dell'intero settore/laboratorio) che può essere la più rilevante ai fini medici, dipendendo fortemente dagli scenari organizzativi di ogni laboratorio.

E' quindi una valutazione che deve essere fatta caso per caso e dichiarata nei documenti del TQM, incidendo direttamente sulla "capacità" del processo analitico di rispettare le specifiche di prestazione predefinite (APS).

Per tutti questi motivi anche l'IQC, a volte, deve mettere da parte le tentazioni della semplicità eccessiva e degli automatismi "commerciali" e, criticamente, tenere conto delle ragioni della medicina. A tal fine si esamineranno alcuni contesti di lavoro particolari ove la stima puntuale e aggiornata della precisione analitica evidenzia ancora, nella pratica quotidiana, criticità e lacune.

Tipicamente la problematica si riscontra per la scarsa flessibilità dei software commerciali generalisti, incapaci di gestire situazioni al di fuori delle situazioni lavorative più comuni. Cercheremo, quindi, di fornire alcune indicazioni utili per dare una risposta alle seguenti domande:

1. Come si stima la precisione intermedia a lungo termine, quando la scarsa stabilità dell'analita (max poche settimane) costringe a cambiare frequentemente i lotti del materiale di controllo (dosaggi di alcuni ormoni, farmaci, esame emocromocitometrico, ecc), senza quindi avere la possibilità di raccogliere dati per un tempo significativamente duraturo?
2. Come si calcola la riproducibilità intralaboratorio e/o interlaboratori, quando si utilizzano più analizzatori intercambiabili per la stessa tipologia di analisi (ad esempio, 2 o più contaglobuli dei quali ognuno può essere strumento di backup per l'altro)?
3. Come si calcola la ripetibilità intraserie e/o la precisione intermedia, a breve o a lungo termine, in assenza di un materiale di controllo stabile o commutabile?

## Pratica di laboratorio

### Caso 1– Precisione intermedia: un analizzatore, due o più lotti di materiale di controllo con breve scadenza

Abitualmente, a ogni cambio lotto del materiale di controllo, i software commerciali di gestione dell'IQC prevedono l'azzeramento degli indici di posizione centrale e di dispersione riguardanti il lotto precedente e il loro ricalcolo sulla base dei primi dati del lotto nuovo, indipendentemente dalla durata dello stesso.

Questo fatto comporta che, in caso di scadenze molto brevi (massimo alcune settimane), si perda l'importante informazione derivante dall'analisi statistica cumulativa di misure effettuate per un tempo sufficientemente lungo, come invece avviene quando si hanno a disposizione lotti stabili per molto tempo (almeno un anno).

D'altronde è impossibile per l'industria, nonostante l'utilizzo di conservanti e di altri artefatti in fase di produzione, fornire materiali che siano contemporaneamente commutabili e dotati di una stabilità molto superiore a quella naturalmente riscontrabile nel sangue umano per alcuni analiti.

Per ovviare a questo limite occorre, quindi, operare un cambiamento di paradigma nell'analisi statistica usuale dell'IQC.

Ipotizziamo che per ogni lotto di controllo, alla fine del periodo di utilizzo, siano disponibili la stima di posizione centrale (media aritmetica), di dispersione (deviazione standard) ed il numero totale delle misure (ad esempio, tramite un software residente sullo strumento). Per questo caso specifico, molto comune, non servono i singoli dati giornalieri.

Anche se ogni partita di controllo commerciale idealmente rappresenta una popolazione a se stante, con una distribuzione dei valori caratterizzata da una propria media e deviazione standard, normalmente il produttore fa in modo che le differenze da lotto a lotto siano il più possibile contenute, così come le caratteristiche di commutabilità.

Quindi se le concentrazioni medie sono molto vicine (normalmente entro un 10%) possiamo senz'altro ritenere soddisfatto il requisito statistico dell'**omoschedasticità** della varianza (ovvero la dispersione delle misure non è significativamente diversa da lotto a lotto), condizione necessaria per applicare le statistiche parametriche ai dati. In caso di dubbio è possibile consultare il "*profilo d'imprecisione*" tracciato in fase di validazione locale del metodo oppure ricorrere a uno dei molti test statistici disponibili per valutare l'omogeneità delle varianze.

Quindi, per ogni lotto, avremo varianza omogenea ma medie lievemente discordanti, originate:

- sia dall'imprecisione metodologica (che è quella che realmente ci interessa);
- sia dal *bias* introdotto dalle variabili preparative industriali, diverse per ogni stock. Queste ultime differenze, seppur piccole, se non fossero eliminate dal calcolo si sommerebbero inevitabilmente all'imprecisione analitica a lungo termine falsandone, in eccesso, il risultato.

Il problema può essere risolto considerando ogni singola partita di materiale di controllo come un campione casuale estratto da una popolazione generale rappresentata dall'insieme di tutti i lotti, quest'ultima con:

- indice di posizione centrale la media aritmetica ponderata delle medie di tutti i lotti (rispetto alla media aritmetica semplice, la media ponderata è una stima migliore del valore centrale dell'intera popolazione, perché tiene conto del numero di misure effettuate su ogni singolo lotto);
- indice di dispersione globale la sommatoria delle devianze di ogni lotto (somma degli scarti al quadrato tra ogni valore e la media del proprio gruppo), facilmente ricavabile elevando al quadrato la deviazione standard di ogni gruppo (varianza) e moltiplicando per il relativo numero dei dati. [NB: è necessario sommare le devianze, e non le varianze, perché solo le prime

*hanno proprietà additive*]. Con procedimento inverso, dividendo la devianza globale per il numero totale delle misure meno uno (gradi di libertà) ed estraendo la radice quadrata si ottiene la deviazione standard campionaria relativa all'intera popolazione dei controlli: la miglior stima disponibile della precisione intermedia per il periodo compreso tra il primo e l'ultimo lotto considerato.

In questo modo si apprezza correttamente la componente della precisione ascrivibile unicamente agli errori originati dal sistema analitico (componente interna della varianza - *varianza entro i gruppi*), mentre la parte generata dalle variabili industriali insite nella produzione di ogni lotto (*varianza tra i gruppi* - dovuta ai valori medi differenti per ogni lotto) è esclusa dal calcolo (vedi caso 2 per maggiori delucidazioni sull'analisi della varianza).

Oltre che per verificare il mantenimento degli obiettivi analitici nel lungo periodo (si rammenta che gli indici di capacità di un processo analitico,  $C_p$  e  $C_{pk}$ , peggiorano con l'aumento dell'imprecisione!), conoscere la precisione intermedia tipica del metodo in uso può essere utile per calibrare meglio l'utilizzo delle regole di Westgard e i limiti di controllo superiore (*UCL: Upper Control Limit*) e inferiore (*LCL: Lower control limit*) delle carte di Shewhart, qualunque modello teorico si desideri utilizzare per il loro calcolo e utilizzo.

Interessante anche poter verificare l'andamento della precisione lotto per lotto (coi relativi intervalli fiduciali), costruendo apposite carte di controllo e identificando eventuali periodi temporali con imprecisioni significativamente anomale rispetto alle prestazioni tipiche a lunga scadenza.

La registrazione giornaliera degli eventi in un ***Quaderno di Laboratorio (Lab-Book)*** cartaceo o elettronico, strutturato secondo le esigenze distinte di ogni settore, strumento o processo analitico, aiuterà a comprendere se vi sono cause specifiche (manutenzioni, ricalibrizioni, guasti, ecc) di deterioramento della precisione [cfr. *Raccomandazione SITLAB 9/20: L'Autorizzazione Tecnica (Technical Authorization) nel Laboratorio Biomedico: razionale e concetti-chiave*].

## Caso 2 –Riproducibilità intralaboratorio e interlaboratorio: due o più analizzatori che utilizzano lo stesso lotto di materiale di controllo

Questa criticità sorge quando l'utilizzo contemporaneo (uso sincrono) di due o più analizzatori indipendenti, di modello uguale o diverso, comporta l'assegnazione dei campioni in modo casuale su uno di essi, oppure quando gli strumenti sono utilizzati a turno (uso asincrono) in giorni diversi della settimana (ad esempio, per averne sempre almeno uno efficiente di scorta in caso di guasti) oppure, ancora, quando strumenti decentrati (presso POCT o laboratori "Spoke") forniscono risultati che possono riguardare gli stessi pazienti in alternanza col laboratorio principale "Hub".

In questo caso è la riproducibilità delle misure nel suo insieme (cioè quella dell'intero settore/laboratorio o laboratori collegati) che può essere la più interessante ai fini clinici: in particolare per evitare valutazioni errate sulla differenza critica tra i risultati degli esami seriali di uno stesso paziente effettuati, in modo non tracciabile, su sistemi analitici diversi.

Utilizzando lo stesso lotto di materiale di controllo su strumenti diversi si può stimare, oltre alla precisione intermedia dei singoli analizzatori, anche la riproducibilità di tutti gli strumenti attivi all'interno del settore/laboratorio (o di più laboratori interconnessi), verificando inoltre che l'allineamento dei medesimi resti entro limiti prestabiliti.

Tabulando in modo ordinato le misure dei controlli di qualità (eseguite in periodi di tempo comparabili e con lo stesso materiale), si potrà osservare come le misure di posizione centrale (medie aritmetiche) e di dispersione (deviazioni standard) calcolate per ogni strumento non siano mai uguali.

Queste differenze, escludendo eventuali errori grossolani, sono dovute:

1. al solo errore casuale di misura;
2. all'errore casuale più l'eventuale errore sistematico del disallineamento di un analizzatore rispetto a un altro (per incertezze di calibrazione, variabilità nei sistemi di pipettaggio dei campioni, uso di lotti diversi di reagente, instabilità dei reattivi, manutenzioni effettuate in tempi differenti, guasti, ecc).

L'analisi della varianza a una via, o ANOVA (*ANalysis Of VAriance*), che consente di confrontare da un punto di vista della statistica inferenziale le medie di più gruppi, è lo strumento più adatto per risolvere il problema.

### Analisi della varianza (ANOVA) univariata

In generale l'ANOVA a una via permette di verificare se c'è relazione tra la varianza di una variabile dipendente QUANTITATIVA e una variabile indipendente QUALITATIVA che diversifica i suddetti gruppi, ad esempio per il sesso (maschi/femmine), la classe di età (bambini/adulti/anziani), la salute (sani/malati), analizzatori diversi (strumento 1/ 2/...n); ecc ecc.

Nel caso di ANOVA a una via con una variabile qualitativa dicotomica (che assume cioè solo due valori) si ottengono gli stessi risultati del *test t di Student* per il confronto di due medie.

I requisiti canonici che devono essere rispettati per condurre un'ANOVA univariata sono l'indipendenza delle osservazioni, la normalità della variabile dipendente (distribuzione dei valori di tipo gaussiano, soprattutto per campioni di piccole dimensioni), l'omoschedasticità (omogeneità) delle varianze dei diversi gruppi.

Poiché ai fini dell'IQC la variabile dipendente quantitativa è rappresentata dai risultati di un medesimo controllo di qualità testato su più dispositivi, mentre la variabile indipendente qualitativa designa gli analizzatori stessi (fino a 5 nel tutorial EXCEL di esempio), si ha che: il primo requisito è sempre soddisfatto; il secondo richiede verifiche (test di asimmetria e curtosi) solo in casi particolari (in quanto le misure analitiche ripetute sullo stesso campione tendono quasi sempre ad approssimare

una distribuzione gaussiana); l'omoschedasticità, invece, andrebbe verificata per poter utilizzare con tranquillità il test F per l'ANOVA.

Perché si parla di analisi della varianza quando, di fatto, questa tecnica è utilizzata abitualmente per verificare se le differenze tra le medie dei gruppi possono essere considerate non casuali, ma statisticamente significative con un certo livello di probabilità?

Quando si mettono a confronto due o più gruppi di dati, normalmente si osserva la loro media (indice di posizione centrale) e deviazione standard (indice di dispersione dei dati).

La media aritmetica di campioni numerosi estratti da una popolazione con distribuzione gaussiana è un valore ben definito, con poco margine di variazione, eccetto che in presenza di outlier. La varianza totale, invece, è un valore più complesso e può essere suddivisa matematicamente in **varianza entro i gruppi e varianza tra i gruppi**.

Nel nostro caso, la prima è dovuta alle differenze tra i singoli risultati con la media del proprio strumento (differenza "interna" ai gruppi), mentre la seconda si riferisce alle differenze tra le medie di ogni strumento e la media aritmetica ponderata generale (differenza "tra" i gruppi).

Questa scissione della varianza permette di capire se la differenza tra le medie sia causata da un vero disallineamento sistematico (*bias strumentale*) oppure sia l'effetto, per ogni strumento, di grandi differenze tra le singole osservazioni. In altre parole, se la variabilità generata da ogni strumento è elevata (precisione strumentale scarsa!) rispetto alla variabilità che si riscontra tra gli strumenti, allora probabilmente la differenza tra i gruppi è casuale e dipende solamente dalla variabilità interna. Viceversa, più è piccola la varianza di ogni strumento (precisione molto buona!) e più è grande la varianza tra gli strumenti, ci si troverà di fronte molto probabilmente a un disallineamento sistematico tra due o più analizzatori (medie statisticamente diverse).

Che ciò sia significativo ai fini clinici dipende esclusivamente dagli obiettivi analitici che il laboratorio si pone e da come il laboratorio sia organizzato riguardo al proprio bacino di utenza: in particolare, qual è il flusso dei campioni rispetto agli analizzatori disponibili.

Sfruttando le proprietà additive della devianza (unico indice di dispersione statistico che si può sommare o sottrarre) la devianza totale (***SS - Sum of Squares***), che rappresenta la riproducibilità tra tutti gli analizzatori (compreso, quindi, il contributo dei disallineamenti tra gli analizzatori), è scomposta in due parti (*tabella 1*):

- una componente attribuibile alla differenza tra le medie dei gruppi (nel nostro caso gli analizzatori) e la media aritmetica pesata di tutti gli strumenti (***SSB - Sum of Squares Between***), componente della variabilità "esterna" riconciliabile ai bias tra gli analizzatori;
- una seconda componente dovuta alle differenze riscontrate, all'interno di ogni gruppo, tra ogni valore e la media del proprio gruppo (***SSW - Sum of Squares Within***), componente della variabilità "interna" ascrivibile alla precisione intermedia di tutti i sistemi analitici.

Dividendo ciascuna somma dei quadrati (le devianze) per i rispettivi gradi di libertà, si ottengono tre varianze, o medie dei quadrati: la media dei quadrati tra gruppi (***MSB - Mean of Squares Between***), la media dei quadrati all'interno dei gruppi (***MSW - Mean of Squares Within***) e la media dei quadrati totale (***MST - Mean of Squares of the Total***), somma delle prime due.

Se prevale la componente interna significa che la varianza NEI gruppi (MSW) sarà la più grande, viceversa se quella esterna è la principale, allora la varianza TRA gruppi (MSB) sarà la più elevata.

<p><b>Devianza totale</b>  <b>SST</b> = mettendo in un'unica lista <u>tutti i dati</u> e calcolando la somma dei quadrati delle differenze <u>rispetto alla media totale</u></p> <p><b>SST = SSB + SSW</b></p>	<p><b>Devianza tra i gruppi</b>  <b>SSB</b> = calcolando la somma dei quadrati delle differenze tra le <u>medie di ciascun gruppo</u> rispetto alla <u>media totale</u> (si annulla la variazione entro i gruppi)</p>	<p><b>Devianza entro i gruppi</b>  <b>SSW</b> = calcolando le somme dei quadrati delle differenze tra <u>ogni dato</u> e la rispettiva <u>media di gruppo</u> (entro ciascun gruppo), facendone poi la sommatoria (si annulla la variazione tra gruppi)</p>
<p><b>GDL<sub>SST</sub></b> = n° totale dati - 1</p>	<p><b>GDL<sub>SSB</sub></b> = n° gruppi - 1</p>	<p><b>GDL<sub>SSW</sub></b> = n° totale dati - numero gruppi</p>
<p>Varianza<sub>SST</sub>  <b>MST</b> = SST/ GDL<sub>SST</sub></p>	<p>Varianza<sub>SSB</sub>  <b>MSB</b> = SSB/ GDL<sub>SSB</sub></p>	<p>Varianza<sub>SSW</sub>  <b>MSW</b> = SSW/ GDL<sub>SSW</sub></p>

**Tabella 1:** modalità di scomposizione della varianza totale nelle sue componenti principali e relativi gradi di libertà.

Per verificare se esiste un bias sistematico tra gli analizzatori (cioè se le medie siano statisticamente diverse tra loro) si utilizza il **test F per l'ANOVA a una via**, ottenuto come rapporto tra MSB (con GDL<sub>SSB</sub>) e MSW (con GDL<sub>SSW</sub>). Più questo rapporto è alto, maggiore è la probabilità che una o più medie differiscano tra loro in modo non casuale.

Oltre a calcolare gli intervalli di confidenza di medie e CV, il tutorial EXCEL fornisce automaticamente il valore del test F e della probabilità statistica che la differenza sia casuale: normalmente una probabilità del 5% ( $p=0.05$ ) o superiore è considerata accettabile per rifiutare l'ipotesi di bias sistematico tra le medie.

Per utilizzare correttamente il test F, occorre prima verificare che tutti gli analizzatori abbiano una varianza omogenea. Il tutorial EXCEL a questo fine utilizza il **test di Levene**, il quale restituisce la probabilità statistica che le differenze riscontrate tra le varianze siano casuali: come per il test F, una probabilità superiore al 5% ( $p=0.05$ ) è considerata accettabile per rifiutare l'ipotesi di differenze statisticamente significative tra le varianze.

Occorre comunque rammentare che all'aumentare del numero dei dati raccolti la potenza di questi test aumenta, evidenziando anche minime (insignificanti ai fini pratici!) variazioni nelle medie e nelle varianze: per cui, nel laboratorio clinico, questi test trovano i loro limiti operativi in traguardi analitici realistici collegati alle esigenze mediche dei pazienti.

In definitiva, anche se un test di Levene o un test F ANOVA significativo può essere preoccupante da un punto di vista rigorosamente matematico, non necessariamente lo deve essere da un punto di vista pratico.

A questo riguardo, il tutorial EXCEL permette di confrontare un **disallineamento "di allarme"** inseribile manualmente (ricavabile, ad esempio dalla letteratura, dall'esperienza, dalla variabilità biologica o altro) con il massimo disallineamento rilevato, evidenziando l'opportunità di interventi correttivi solo quando effettivamente necessario.

Per ultimo, occorre rimarcare come ogni laboratorio, qualora utilizzi più sistemi analitici per gli stessi test, debba prestare molta cura alla validazione iniziale e alla verifica regolare dell'allineamento strumentale, al fine di rendere perennemente indipendente il risultato di una misura dall'analizzatore utilizzato.

Le modalità di validazione di metodi e strumenti esulano dai fini del presente documento e, pertanto, si rimanda alla letteratura specifica per ogni approfondimento.

### Caso 3 – Precisione in assenza di un materiale di controllo stabile e/o commutabile: un analizzatore, utilizzo di replicati di campioni umani

Quando, per insuperabili problemi di commutabilità, l'utilizzo dei risultati ottenuti da controlli commerciali possa essere fuorviante oppure quando l'industria non li fornisca (per scarsa stabilità dell'analita o per test talmente rari da rendere non sostenibile economicamente la loro produzione), è necessario ricorrere a strategie differenti per monitorare le prestazioni degli analizzatori.

In questo caso può essere utile utilizzare replicati (misure in doppio) di campioni umani.

Si tratta di una via alternativa all'utilizzo di materiale prodotto dall'industria: semplice e poco costosa.

E' stata molto utilizzata agli albori della chimica-clinica moderna (fino agli anni '70-'80 del secolo scorso), quando i laboratori erano più piccoli, gli esami prevalentemente manuali, i kit analitici commerciali poco diffusi e gli analizzatori di grandi dimensioni e produttività ancora in divenire. In alcuni casi può essere utilizzata in aggiunta ai controlli commerciali (ad esempio quando controlli multicomponenti alla prova dei fatti si rivelino commutabili/idonei solo per alcuni test e non per altri).

Anche in questo caso **il significato della precisione misurata è condizionato in modo determinante dal disegno sperimentale adottato**: così se le misure in doppio sono ottenute nella stessa serie analitica si otterrà una stima della precisione entro la serie (ripetibilità) mentre se le misure in doppio sono ottenute in giorni, e quindi in serie analitiche diverse, misurando ogni giorno un campione già analizzato per lo stesso componente il giorno precedente (metodo ieri-oggi), si otterrà una stima della precisione tra le serie (precisione intermedia).

Come sempre il campione sottoposto a doppia analisi non deve godere, al pari dei controlli commerciali, di nessun trattamento privilegiato.

Poiché le differenze tra le coppie di valori possono essere influenzate, in modo variabile, dalla concentrazione (da valutare tramite il profilo d'imprecisione del metodo), è importante scegliere per la ripetizione prelievi con concentrazione non molto diverse tra di loro, in genere ricomprese in un ventaglio di valori predeterminato (non necessariamente l'intervallo di "normalità" ma, più utilmente, un range attorno ai valori decisionali critici).

Sono pure da evitare, specialmente nel caso di analisi ripetute in giorni diversi, prelievi suscettibili d'instabilità e interferenze note, come campioni da pazienti leucemici, con agglutinine a freddo, emolizzati, lipemici, itterici, eccetera. Normalmente, nel caso di analisi replicata nella serie, la provetta rimane sul banco di lavoro, chiusa col tappo, mentre nel caso di analisi ripetuta nel giorno successivo va conservata in frigorifero (2-8° C), sempre ben sigillata e con la parte corpuscolata separata dal siero/plasma (utilizzo del gel-separatore oppure separazione manuale).

A partire dalle coppie di valori (nella serie o tra giorni) è possibile calcolare la deviazione standard descrittiva dell'imprecisione del metodo e generare i valori del CV%.

Eventuali fluttuazioni dell'imprecisione possono essere verificate, esaminando con il test F di Fisher per la varianza, la significatività statistica di differenze tra valori ottenuti in tempi successivi.

I metodi basati sull'analisi di replicati non sono sensibili nel rilevare errori sistematici, ma con le *Cumulative Sum Chart (CUSUM)* applicate al metodo ieri-oggi è possibile osservare l'instaurarsi di tendenze.

L'esempio riportato nel tutorial EXCEL è tratto dalle linee guida della Regione Lombardia sul Controllo di Qualità Interno nel Servizio di Medicina di Laboratorio", a cura della "Rete di riferimento per la qualità dei servizi di medicina di laboratorio".

Esso permette di confrontare, tramite il test F, varianze osservate in periodi diversi. Inoltre, adoperando la deviazione standard tipica del metodo/strumento fornita dal produttore oppure la deviazione standard calcolata su un numero sufficiente di duplicati in condizioni di stabilità del metodo (da utilizzare come riferimento), consente di calcolare il "Limite dell'intervallo dei duplicati", cioè la massima differenza normalmente riscontrabile, con una certa probabilità, entro un duplicato. Se tra due ripetizioni si ottiene una differenza maggiore del "limite", in particolare se questo comportamento si mantiene nel tempo, è probabile che la precisione si stia deteriorando in modo significativo.



## Bibliografia

1. Per una panoramica sullo “scandalo Theranos”  
[https://www.repubblica.it/tecnologia/2022/11/23/news/3\\_cose\\_da\\_imparare\\_dal\\_caso\\_theranos\\_e\\_dalla\\_condanna\\_di\\_elizabeth\\_holmes-375631792/](https://www.repubblica.it/tecnologia/2022/11/23/news/3_cose_da_imparare_dal_caso_theranos_e_dalla_condanna_di_elizabeth_holmes-375631792/)  
[https://www.italian.tech/2022/01/04/news/la\\_startup\\_theranos\\_era\\_una\\_truffa\\_condannata\\_la\\_donna\\_che\\_voleva\\_essere\\_steve\\_jobs-332625726/](https://www.italian.tech/2022/01/04/news/la_startup_theranos_era_una_truffa_condannata_la_donna_che_voleva_essere_steve_jobs-332625726/)
2. *Rapporti ISTISAN 13/41 - Guida Eurachem. Terminologia per le misurazioni analitiche – Introduzione al VIM 3 (International Vocabulary of Metrology)*  
[https://www.eurachem.org/images/stories/Guides/pdf/TAM\\_2011\\_IT.pdf](https://www.eurachem.org/images/stories/Guides/pdf/TAM_2011_IT.pdf)
3. SIPMeL - L5Q15 Raccomandazioni per il glossario nelle nuove norme ISO per i laboratori medici su taratura, “etichette” e controllo di qualità (ISO 17511, ISO 18113 e ISO 15198)  
[https://www.sipmel.it/download/118108-L5Q15glossarioISO1751118113\\_pw.pdf](https://www.sipmel.it/download/118108-L5Q15glossarioISO1751118113_pw.pdf)
4. SIPMeL - Q16 - Raccomandazioni per la stima dell'incertezza di misura nei laboratori medici (ISO 15189 e ISO 20914)  
[https://www.sipmel.it/download/118104-Q16incertezzaISO20914\\_agg7sept21\\_pw.pdf](https://www.sipmel.it/download/118104-Q16incertezzaISO20914_agg7sept21_pw.pdf)
5. SIBioC - Linee guida per gestione dei programmi di Controllo di Qualità Interno  
<https://sibioc.it/wp-content/uploads/2021/05/1042-ottomano-1.pdf>
6. SIBioC - Linee guida per la gestione dei Programmi di Valutazione Esterna di Qualità  
[https://sibioc.it/wp-content/uploads/2021/04/305-Valutazione\\_esterna\\_qualita\\_2009.pdf](https://sibioc.it/wp-content/uploads/2021/04/305-Valutazione_esterna_qualita_2009.pdf)
7. SIBioC - Protocollo per la comparazione di due metodi analitici di laboratorio  
<https://sibioc.it/wp-content/uploads/2021/06/724-1181-129-142-Vidali-doc-sibioc-5-2.pdf>
8. SIBioC - Protocollo operativo per la verifica della comparabilità dei risultati di laboratorio ottenuti su più procedure analitiche  
<https://sibioc.it/wp-content/uploads/2021/05/4410-vidali.pdf>
9. Rapporti ISTISAN 16/39 – Guida Eurachem – Idoneità per lo scopo dei metodi analitici  
[https://www.iss.it/documents/20126/45616/16\\_39\\_web.pdf/03c10bed-c18a-7c7a-4e01-8995253f50f2?t=1581099238778](https://www.iss.it/documents/20126/45616/16_39_web.pdf/03c10bed-c18a-7c7a-4e01-8995253f50f2?t=1581099238778)
10. ERM – Application note 8 – Valutazione della commutabilità dei materiali di riferimento  
[https://crm.jrc.ec.europa.eu/graphics/cms\\_docs/erm8\\_italian.pdf](https://crm.jrc.ec.europa.eu/graphics/cms_docs/erm8_italian.pdf)
11. DECRETO DIREZIONE GENERALE SANITA' N. 32856 DEL 19.12.2000 (Lombardia)  
Linee guida su “Controllo di Qualità Interno nel Servizio di Medicina di Laboratorio”  
[http://www.qualitalaboratorilombardia.it:8080/front/public/1588172930106\\_DD2000\\_32856\\_\(1\).pdf](http://www.qualitalaboratorilombardia.it:8080/front/public/1588172930106_DD2000_32856_(1).pdf)
12. Soliani L. – Manuale di statistica per la ricerca e la professione  
<http://www.dsa.unipr.it/soliani/soliani.html>