

Validazione dei metodi quantitativi bioanalitici in spettrometria di massa: regole e protocolli sperimentali allineati alla regolamentazione europea 2017/746

Bioanalytical method validation of quantitative mass spectrometry-based assays: rules and experimental protocols following 2017/746 regulation

AUTORI: Antonio D'Avolio¹, Viviana Izzo^{2,3}, Bruno Charlier^{2,3}, Jessica Biasizzo⁴, Alessia Cafaro⁵, Sara Cairolì⁶, Giuliana Cangemi⁵, Chiara Capelli⁷, Albino Coglianesi⁸, Ugo De Grazia⁹, Amedeo De Nicolò¹, Rossana Domenis⁴, Claudia Giugliano¹⁰, Bianca Maria Goffredo⁶, Marta Leporati¹¹, Antonello Nonnato¹², Federica Pigliasco⁵, Marco Cantù^{#10}

#. Gruppo di Studio SIBIOC (Società Italiana di Biochimica Clinica e Biologia Molecolare Clinica): La spettrometria di massa: applicazioni e innovazioni diagnostiche

- 1 Laboratory of Clinical Pharmacology and Pharmacogenetics, University of Turin, Department of Medical Sciences, Amedeo di Savoia Hospital, 10149, Turin, Italy
- 2 Unità Operativa Complessa di Farmacologia Clinica - AOU San Giovanni di Dio e Ruggi d'Aragona Salerno, Italia
- 3 Dipartimento di Medicina, Chirurgia e Odontoiatria - Università degli Studi di Salerno, Italia
- 4 Institute of Clinical Pathology, Academic Hospital "Santa Maria della Misericordia", Azienda Sanitaria Universitaria Friuli Centrale, Udine, Italy
- 5 Biochemistry, Pharmacology and Newborn Screening Unit, Central Laboratory of Analyses, IRCCS Istituto Giannina Gaslini, Genoa, Italy
- 6 Laboratorio di Malattie Metaboliche e Biologia del Farmaco I.R.C.C.S. Ospedale Pediatrico Bambino Gesù
- 7 Pharmacotoxicology Unit, Centro Diagnostico Italiano C.D.I., 20147 Milan, Italy
- 8 Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, 84081 Salerno, Italy
- 9 Laboratorio Patologia Clinica e Sostanze d'abuso; Centro Diagnostico Italiano C.D.I., 20147 Milan, Italy

10 Laboratory of Clinical Biochemistry and Pharmacology, Clinic of Laboratory Medicine, Institute for Integrated Diagnostics of Italian Switzerland (IDISI), Ente Ospedaliero Cantonale, Bellinzona, Switzerland

11 Settore Chimica Analitica e Calcolosi Renale - SC Laboratorio Analisi, AO Ordine Mauriziano di Torino, Largo Turati 62, 10128 Torino, Italy

12 Clinical Biochemistry Unit, Dept. Diagnostic Laboratory, A.O.U. "Città della Salute e della Scienza" Hospital, Turin, Italy

Corresponding author: Prof. Antonio D'Avolio (BSc, MSc, SM); Laboratory of Clinical Pharmacology and Pharmacogenetics*. University of Torino, Department of Medical Sciences, Pad. Q; Amedeo di Savoia Hospital, Corso Svizzera 164 - 10149 Turin (ITALY).

Fax: +39.011.4393996; TEL. +39.011.4393841, e-mail: antonio.davolio@unito.it.

*UNI EN ISO 9001 and 13485 Certified Laboratory;

Parole chiave: LC-MS/MS, LC-MS, validazione metodi bioanalitici, molecole endogene, molecole esogene, IVD, IVDR, therapeutic drug monitoring

Indice delle abbreviazioni

CE-IVD: Conformité Européenne-In Vitro Diagnostic

CE-IVDR: Conformité Européenne-In Vitro Diagnostic Regulation

CLSI: Clinical and Laboratory Standard Institute

CQ: Controllo di Qualità

CRC: Composto di Riferimento Cromatografico

CRM: Certified Reference Material

CSF: liquido cerebrospinale

DBS: dried blood spots

DPS: dried plasma spots

EM-nEM: Effetto Matrice – Effetto Matrice normalizzato dallo standard interno

EMA: European Medicines Agency

FDA: Food and Drug Administration

Fn: Falso Negativo

Fp: Falso Positivo

GTFI: Gruppo Tossicologi Forensi Italiani

HRMS: Spettrometria di Massa ad Alta Risoluzione

ISO: International Organization for Standardization

ISS: Istituto Superiore di Sanità
IVD: In Vitro Diagnostics
IVDR: In Vitro Diagnostic Regulation
LC: Cromatografia Liquida
LC-MS: Cromatografia Liquida accoppiata alla Spettrometria di Massa
LC-MS/MS: Cromatografia Liquida accoppiata alla Spettrometria di Massa tandem
LDT: Laboratory Developed Test
LLOQ: Lower Limit Of Quantification
LOD: Limit Of Detection
LOQ: Limit Of Quantification
OR: Odds Ratio
RA: Abbondanza Relativa
RRT: Tempo di Ritenzione Relativo
RT: Tempo di Ritenzione
SCAN: Scansione a spettro completo
SI: Standard Interno
SIM: Selected Ion Monitoring
SRM: Selected Reaction Monitoring
ULOQ: Upper Limit Of Quantification
VEQ: Valutazione Esterna di Qualità
VIM: Vocabolario Internazionale di Metrologia
Vn: Vero Negativo
Vp: Vero Positivo
VPN: Valore Predittivo Negativo
VPP: Valore Predittivo Positivo

Abstract (ENG)

Liquid chromatography coupled to mass spectrometry (LC-MS) has rapidly extended the spectrum of clinical chemistry methods during the past decade, thus becoming a standard tool for the evaluation of endogenous and exogenous compounds in many research and clinical laboratories. To date, numerous guidelines for bioanalytical method validation have been published; however, there remains a clear need for harmonized experimental procedures that can be broadly adopted by the scientific community engaged in the extensive clinical use of mass spectrometry. Most importantly, the new Regulation (EU) 2017/746 has sparked an active debate on the evaluation of both analytical and clinical performance of laboratory-developed tests (LDTs) and commercially available diagnostic kits, with clear implications also for LC-MS-based methodologies. The main aim of this document is to suggest specific experimental procedures for an exhaustive evaluation of analytical and clinical performances of bioanalytical LC-MS-based methods, which are inspired by current international guidelines and integrated with practical tips aimed at achieving better performance of methods used in routine clinical practice. Here we have introduced concepts such as "normalized matrix effect" (nME) and protocols for method validation using "pathological samples" and "altered samples" as the most commonly not-normal samples retrieved in laboratory routine. Rules and procedures described in this work are not aimed at replacing international guidelines; they are instead intended to better elucidate to LC-MS users several methodological aspects related to the new 2017/746 European regulation and to underline further experimental steps to be performed during each method/kit validation. Moreover, chapter VI of the EU regulation details several aspects in terms of clinical evidence, which should be able to demonstrate compliance of a bioanalytical method with the requirements of Annex I. The clinical evidence corroborates the declared intended use of the method and should be able to demonstrate (by reference to the state of the art) that clinical benefits will be effectively achieved. The main aim of the second part of our work is to shed light on the requirements for the clinical performance evaluation, including tips and advice to correctly set up the clinical performance evaluation plan and the clinical evaluation report. Several aspects concerning the scientific validity and the clinical performance of a bioanalytical method are also discussed, which involve diagnostic sensitivity and specificity, negative and positive predictive values, both for drugs and endogenous compounds.

1. Introduzione

Sono trascorsi alcuni anni dalla pubblicazione sulla rivista Biochimica Clinica del primo documento dal titolo “Validazione dei metodi quantitativi bioanalitici in spettrometria di massa: regole e protocolli sperimentali” a cura del gruppo di studio in spettrometria di massa della SIBioC¹.

L'entrata in vigore del Regolamento europeo 2017/746 (IVD-R) nel maggio 2022² ha stimolato gli autori ad un importante aggiornamento, con l'intento di guidare i professionisti alla corretta messa in pratica delle linee guida (nell'ambito della normativa stessa) per la verifica delle prestazioni cliniche ed analitiche dei metodi di laboratorio basati sulle tecniche accoppiate alla spettrometria di massa. Uno dei concetti chiave attorno al quale ruota la normativa 2017/746 riguarda l'importanza della valutazione dell'evidenza clinica di una metodica bioanalitica, sia essa commercialmente disponibile o sviluppata “*in-house*”, la quale ha lo scopo fondamentale di dimostrare che il beneficio clinico e la sicurezza di una metodica siano raggiunti in conformità con lo stato dell'arte nel campo clinico di interesse. L'evidenza clinica si basa su tre pilastri principali: la “validità scientifica”, ovvero l'evidenza che l'analita, il marcatore o la molecola studiata sia rilevante per l'aspetto fisiologico o condizione patologica di interesse; le “prestazioni analitiche”, che forniscono indicazioni sulla capacità del dispositivo di misurare accuratamente ed in maniera riproducibile l'analita di interesse, ed infine, la “prestazione clinica”, ovvero la capacità di una metodica bioanalitica di produrre risultati correlati al quadro clinico reale a seconda della popolazione target e dell'utente. Ognuno di questi tre aspetti dovrebbe essere sviluppato nel dettaglio e descritto separatamente in quello che la normativa definisce come il “*performance evaluation report*”, che include il report relativo a tutti e tre gli aspetti (validità scientifica, prestazione analitica e prestazione clinica), e che andrebbe aggiornato con cadenza regolare dallo sperimentatore.

Il presente documento, che include la valutazione delle prestazioni analitiche e di quelle cliniche di una metodica, e che rappresenta un importante aggiornamento rispetto al documento precedente, tiene conto in parte anche delle raccomandazioni per l'accreditamento dei laboratori clinici secondo la norma ISO 15189 che, insieme alla certificazione IVDR, ha introdotto nuove importanti specifiche di performance³.

Il documento è stato elaborato da professionisti con comprovata esperienza in Medicina di Laboratorio, membri del Gruppo di Studio di Spettrometria di Massa della SIBioC. Il loro intento, oltre a quello di revisione e aggiornamento dello stato dell'arte, è stato quello di armonizzare le raccomandazioni contenute nelle linee guida per la valutazione delle

performance analitiche dei metodi bioanalitici emanate da enti scientifici istituzionali riconosciuti internazionalmente quali:

- a) Istituto Superiore di Sanità (ISS)
- b) European Medicines Agency (EMA)
- c) Food and Drug Administration (FDA)
- d) Clinical and Laboratory Standard Institute (CLSI)
- e) Gruppo Tossicologi Forensi Italiani (GTFI)
- f) World Anti-Doping Agency (WADA)

Il documento è rivolto a tutti i professionisti di Medicina di Laboratorio, con lo scopo di diffondere e promuovere la cultura analitico-strumentale attraverso l'esperienza del Gruppo di Studio di Spettrometria di Massa della SIBioC. Inoltre, si è cercato di mettere a disposizione uno strumento di riferimento per un corretto approccio analitico che garantisca uno standard di qualità improntato all'armonizzazione e alla confrontabilità dei risultati, fornendo allo stesso tempo le indicazioni pratiche sulla gestione dei processi di valutazione delle performance analitica e clinica in riferimento alla normativa 2017/746.

2. Valutazione della performance analitica del metodo

Come evidenziato da diverse linee guida di recente pubblicazione, qualsiasi metodo analitico, indipendentemente dal fatto che sia sviluppato ex novo o derivi in parte o totalmente da dati di letteratura, deve essere completamente validato, soprattutto quando utilizzato in ambito clinico.

Lo scopo della valutazione delle prestazioni analitiche è quello di dimostrare l'affidabilità di un metodo nella determinazione della concentrazione di un analita (o di più analiti) in una matrice biologica specifica, come il sangue, il plasma, il siero, le urine, il fluido cerebrospinale o la saliva, considerando anche i nuovi sistemi di campionamento quali quelli su carta o filtro, come i *dried blood spots* (DBS) e i *dried plasma spots* (DPS)⁴⁻¹⁰. Generalmente, una validazione completa dovrebbe essere effettuata per ciascun tipo di analita e di matrice considerati, e valutando anche la presenza di diversi additivi quali ad esempio i diversi anticoagulanti presenti nelle provette di raccolta¹¹. Tutti i passaggi richiesti per una validazione completa dei metodi bioanalitici devono fare riferimento alle linee guida

internazionali attualmente disponibili^{4,11,12}, le quali devono però essere riesaminate criticamente in funzione delle specifiche caratteristiche della tecnologia LC-MS.

2.1. Selettività – Specificità analitica

La LC-MS è una metodologia analitica ottimale in termini di selettività e specificità sia rispetto alle analisi classiche in LC che ai metodi immunometrici, in quanto combina la separazione LC con la capacità di identificazione della spettrometria di massa attraverso il monitoraggio delle transizioni SRM dei singoli analiti. Tuttavia, anche utilizzando la spettrometria di massa possono essere presenti interferenze causate da composti endogeni, esogeni, molecole isobare, metaboliti, impurezze, cross-talk (interferenza di segnale) con altri analiti o SI co-eluenti e contaminanti derivanti dal pretrattamento del campione che possono influire su selettività, specificità, accuratezza e precisione.

Secondo le indicazioni EMA, la selettività è la capacità di un metodo analitico di differenziare e misurare l'analita d'interesse in presenza di potenziali sostanze interferenti nella matrice biologica. Se la presenza di tali interferenti è nota ed attesa, il loro impatto deve essere preso in considerazione durante le fasi di sviluppo e validazione del metodo, tenendo conto sia del loro peso molecolare che delle caratteristiche cromatografiche. Il metodo cromatografico deve essere in grado di differenziare nella matrice biologica sia gli analiti d'interesse sia lo SI dai composti endogeni o da altri componenti presenti nel campione del paziente.

La selettività e la specificità devono essere valutate sia per gli analiti di interesse che per i relativi SI, analizzando individualmente almeno sei diverse “matrici “bianche” (assenza di analiti e SI, o, nel caso di composti endogeni, matrici surrogate “*neat*”).

In particolare, se vengono analizzate molecole esogene, possono essere usati sei campioni di pazienti che non contengono le molecole d'interesse. Inoltre, è necessario valutare eventuali interferenti comunemente riscontrabili in campioni alterati, come ad esempio campioni plasmatici emolizzati, lipemici o itterici. Solo in caso di uso di matrici rare (quali essudati, fluido cerebrospinale, tessuti solidi, etc..) e per applicazioni non routinarie (es. per ricerca) il numero dei campioni potrebbe essere ridotto, anche se è un'opzione non preferibile. L'assenza d'interferenti in matrici bianche può essere accettata quando la risposta media per gli analiti ai relativi tempi di ritenzione è al di sotto del 20% del loro LLOQ e del 5% della risposta dello/gli SI¹¹.

La valutazione della specificità è fondamentale soprattutto in ambiti legali e forensi, quali la tossicologia, dove l'identificazione dell'analita deve essere garantita. Un analita risulta correttamente identificato in un campione quando sono rispettati i **criteri minimi di**

identificazione^{13, 14}. L'identificazione di un analita mediante LC-MS/MS si basa sul confronto tra il RT (criterio cromatografico) e la RA degli ioni diagnostici dell'analita, quei frammenti ionici specifici e caratteristici che si formano durante la frammentazione di una molecola e che servono ad identificarla (criterio spettrometrico) rilevati in un campione rispetto a quelli di un campione di riferimento analizzato nella stessa seduta analitica. Solitamente, si sceglie come ione diagnostico quello più abbondante, che verrà usato come riferimento per il RT e per il calcolo della RA degli ioni¹³. È importante precisare che ciò non costituisce un criterio assoluto, poiché se lo ione più abbondante è al contempo meno stabile, o vi è una chiara evidenza di interferenze sulla sua transizione (*ion ratio* variabile), esso potrebbe non essere scelto come ione diagnostico di riferimento.

2.1.1 Criteri cromatografici

Il RT del picco cromatografico dell'analita nel campione non deve differire (ΔRT) di più dell'1% o di $\pm 0,1$ minuti da quello dello stesso analita in un campione di riferimento analizzato nello stesso batch analitico¹³. In alternativa, il Laboratorio può scegliere di utilizzare come criterio di accettabilità il RRT, dove viene misurato il RT del picco di interesse rispetto ad un CRC. Se il CRC non è l'analita marcato isotopicamente, l'RRT dell'analita nel campione non deve differire di oltre $\pm 1\%$ da quello dello stesso analita in un campione di riferimento analizzato nello stesso batch analitico. Se invece il CRC è l'analita marcato isotopicamente, l'RRT dell'analita nel campione non deve differire di oltre $\pm 0,5\%$ da quello dello stesso analita in un campione di riferimento analizzato nello stesso batch analitico¹³. Secondo altri riferimenti bibliografici, indipendentemente dal fatto che il CRC sia marcato isotopicamente o meno, la tolleranza per RRT è più ampia ($\pm 2\%$)¹⁴.

2.1.2 Criteri spettrometrici

Per quanto riguarda l'identificazione mediante spettrometria di massa, i criteri minimi di identificazione cambiano in base alla tipologia di acquisizione utilizzata (SCAN, SIM, SRM, HRMS-SCAN). In questo documento si fa riferimento all'acquisizione in SRM, poiché maggiormente utilizzata in ambito clinico.

In SRM devono essere acquisite almeno due transizioni (ione precursore/ione prodotto) per ogni analita o SI.

Le RA di una qualsiasi delle transizioni considerate non devono differire di più della quantità specificata nella Tabella 1 dalle corrispondenti RA delle stesse transizioni acquisite dal campione di riferimento¹³.

Tabella 1. Calcolo della abbondanza relativa (RA) delle transizioni acquisite in SRM.

RA nel cromatogramma di riferimento* (% ione diagnostico di riferimento)	Tolleranza massima per RA nel cromatogramma del campione	Esempio	
		RA (% ione diagnostico di riferimento**)	Finestra di tolleranza (% diagnostico di riferimento)
>50	± 10 (assoluto)	100**	110 – 90
		60	50 – 70
>25	± 20 (relativo)	40	32 – 48
1 – 25	± 5 (assoluto)***	10	5 – 15
		3	>0***

* Analizzato nello stesso batch analitico

** Deve essere valutata anche la finestra massima di tolleranza per lo ione diagnostico di riferimento

*** Gli ioni diagnostici devono essere sempre rivelati nel campione (S/N>3).

Anche secondo le linee guida GTFI devono essere acquisite almeno due transizioni ione precursore/ione prodotto per ogni analita e SI. Per garantire l'identificazione del picco, si raccomanda di non utilizzare ioni prodotto derivanti da frammentazioni aspecifiche (es. perdita di H₂O)¹⁴.

2.1.3 Valutazione delle interferenze tra SI marcati con isotopi stabili e gli analiti

Congiuntamente alle altre valutazioni di selettività e specificità, si suggerisce inoltre di valutare eventuali interferenze causate dalla presenza dello SI sia in termini di cross-talk che di effetto matrice. Per cross-talk si intende il segnale generato al tempo di ritenzione dell'analita d'interesse causato da un'interferenza spettrale dello SI o di altri misurandi o dalla presenza di impurezze nello stesso. Per valutare questo fenomeno si suggerisce di aggiungere ad una matrice bianca (priva di analita) la quantità di standard interno prevista dal metodo e di valutare il segnale al RT dell'analita, che deve essere inferiore al 20% del segnale al suo LLOQ. Analogamente, si consiglia di valutare il segnale derivante dal cross-talk dell'analita sul suo standard interno, considerando il segnale derivante da alte concentrazioni di analita (ULOQ) in un campione privo di SI: in questo caso, il cross-talk deve essere inferiore al 5% del segnale medio dello SI.

Inoltre, si suggerisce di valutare l'effetto matrice dato dallo SI verso il suo analita e viceversa, attraverso valutazione post-estrazione (si veda paragrafo 2.8.1).

2.1.4 Valutazione delle interferenze date da analiti e/o metaboliti comunemente riscontrabili nella matrice

Durante lo sviluppo di un metodo bioanalitico è sempre necessario valutare quali composti, comunemente presenti nei campioni per i quali il metodo è stato sviluppato, possano interferire con la quantificazione degli analiti. Ad esempio, un metodo sviluppato per quantificare dei farmaci nel plasma dei pazienti dovrebbe tener conto delle possibili interferenze di farmaci comunemente co-somministrati, dei metaboliti e di eventuali altri composti riscontrabili in tale matrice^{15, 16}. Per queste valutazioni, è possibile fare riferimento alle linee guida EMA. Nello specifico, dovrebbe essere valutata l'eventuale esistenza di interferenze causata da:

- metaboliti conosciuti del(gli) analita(i) di interesse e di composti isobari;
- interferenze causate da prodotti di degradazione generati durante la preparazione di campioni, ad esempio nel caso di preparazioni complesse (che richiedono idrolisi, derivattizzazioni, etc..)¹⁷.

Per le molecole esogene, se sono presenti isomeri (in particolare se tali isomeri hanno un valore biologico/diagnostico differente), la separazione cromatografica dovrebbe essere totale¹⁸. In generale, se si sospetta che specifiche molecole possano interferire con l'analisi, si suggerisce di analizzare un campione bianco (privo di analita e di SI) ed un campione fortificato con analita e SI a cui siano stati aggiunti i possibili interferenti alle concentrazioni attese nella matrice biologica analizzata. Nel caso di analisi di composti endogeni si useranno matrici surrogate "neat", ossia una matrice priva degli analiti di interesse. In campo clinico, al fine di valutare i principali contaminanti dei prelievi ematici, devono essere valutati sia campioni lipemici che emolitici nei quali, analogamente ai casi precedenti, siano assenti sia gli analiti d'interesse che gli SI. In particolare, si suggerisce di testare almeno due matrici con grado elevato di lipemia e almeno due matrici con elevato grado di emolisi.

Come sopra descritto, l'assenza di interferenti è accettata quando la risposta media per gli analiti è al di sotto del 20% (e comunque mai superiore al 25%) del LLOQ e del 5% per lo SI. Qualora uno dei due campioni alterati abbia comunque generato una risposta superiore ai limiti indicati, si suggerisce di testare ulteriori campioni (sei in totale). Se la selettività/specificità per i campioni emolizzati e lipemici non risultasse accettabile sarà necessario provvedere a modificare il metodo analitico oppure, se possibile, verificare fino a

che grado di emolisi e lipemia la selettività non viene inficiata. Campioni caratterizzati da emolisi e lipemia superiori a tale grado andrebbero rigettati come non idonei. Si suggerisce inoltre di considerare i principali contaminanti che potrebbero essere presenti nella popolazione dello studio (es. popolazioni soggette ad ittero, etc..) testando sempre almeno due campioni biologici bianchi secondo le modalità precedentemente indicate¹⁹.

2.2 Carry-over

Il fenomeno del carry-over deve essere studiato attentamente e minimizzato durante lo sviluppo di un metodo LC-MS, prima ancora della validazione. Durante la validazione il carry-over dovrebbe essere valutato iniettando dei campioni bianchi (o *neat*, per gli analiti endogeni) dopo il calibratore a più alta concentrazione di analita o dopo un campione con una concentrazione a livello dell'ULOQ. In alcune particolari situazioni (esempio: analisi tossicologiche su matrice urinaria) la concentrazione attesa per i pazienti può essere notevolmente superiore a quella dell'ultimo punto di calibrazione (es. metadone nelle urine di pazienti in terapia). In questi casi si suggerisce di valutare il carry-over anche dopo l'analisi di questi campioni ad elevata concentrazione. La più alta concentrazione di analita alla quale non si osserva il carry-over (sopra il LOD del metodo) è considerata la concentrazione alla quale il metodo non è influenzato da carry-over. Questa concentrazione deve essere valutata in tripla analisi.

Come riportato dalle linee guida, il carry-over nel bianco (o nel *neat*) dopo un campione ad alta concentrazione non deve essere maggiore del 20% del LLOQ e del 5% per lo SI.

Se possibile, la procedura analitica deve essere modificata per rimuovere ogni carry-over; quando non vi è la possibilità di eliminare completamente il carry-over, le procedure devono riportare come questo possa essere mitigato¹⁹.

2.3. Sensibilità analitica. Limite di rilevabilità e limite di quantificazione

La sensibilità analitica di un metodo si riferisce al cambiamento della risposta dello strumento in corrispondenza del cambiamento della grandezza misurata e correla la concentrazione di un analita con il gradiente della curva di risposta. Il prefisso “analitico” viene utilizzato per evitare confusione con il concetto di “sensibilità diagnostica” ed il suo utilizzo come sinonimo di limiti di rilevamento, seppur spesso ancora utilizzato, è comunque sconsigliato dal VIM²⁰.

Per le misurazioni effettuate a basse concentrazioni atte a definire i limiti di rilevamento²¹, bisogna innanzitutto definire tre concetti generali:

- stabilire un valore del risultato che si ritiene indicare un livello di analita significativamente diverso da zero (valore critico);
- conoscere la più bassa concentrazione di analita che può essere rivelata dal metodo ad un livello di fiducia specificato. In altre parole, definire la concentrazione reale alla quale si potrà affermare se il valore critico definito è stato superato. Per questo concetto sono utilizzati termini quali LOD;
- stabilire il livello più basso a cui la prestazione analitica è accettabile per quella tipologia di applicazione entro i limiti di confidenza definiti in seguito. Questo terzo concetto è comunemente definito LOQ.

Il LOD è definito come la concentrazione alla quale il rivelatore restituisce un segnale dell'analita almeno tre volte superiore al segnale di base di un campione bianco (rumore di fondo)¹. Il LOD permette dunque di definire la minima concentrazione rilevata da cui è possibile dedurre con ragionevole certezza statistica la presenza dell'analita. È possibile calcolare il LOD effettuando da sei a quindici ripetizioni (di solito ne sono raccomandate 10) di campioni bianchi, cioè matrici che non contengano quantità rivelabili di analita oppure $n=10$ misurazioni ripetute di campioni di prova con basse concentrazioni di analita (vicine o al di sotto del LOD stimato). In alternativa, è possibile effettuare misurazioni ripetute di bianchi reagente (materiale usato nel processo di analisi che non contiene l'analita di interesse) a cui sono state aggiunte basse concentrazioni di analita. Nel caso in cui i bianchi reagente non siano trattati seguendo l'intera procedura di misura e siano sottoposti direttamente alla misurazione strumentale, il calcolo fornirà il LOD strumentale. Va sottolineato in questo contesto che il limite di rilevabilità strumentale è diverso dal limite di rilevabilità del metodo (che tiene conto di parametri come la resa di estrazione, l'effetto matrice etc.). Il LOD del metodo deve essere preferibilmente basato sull'analisi di campioni che siano stati sottoposti all'intero processo analitico utilizzando risultati calcolati con la stessa curva di calibrazione²².

Il LOD corrisponde al triplo della deviazione standard della media delle determinazioni ripetute dei campioni sopra descritti, ovvero $LOD = 3 \times S'_0$ (questo numero deriva dall'intervallo di confidenza di un test t di Student a due code).

Il LOQ rappresenta la più bassa concentrazione di analita, misurata su dieci replicati per almeno cinque sedute analitiche che fornisce un picco identificabile con precisione del 20% e una inesattezza pari al 15%²². Il LOQ può essere utilizzato come punto più basso della curva

di calibrazione¹¹; tuttavia, nei metodi bioanalitici è possibile non includerlo e selezionare punti a concentrazione più elevata qualora il LOQ non sia rilevante per la definizione dell'intervallo di interesse del metodo. Come per il LOD, è possibile calcolare il LOQ effettuando 10 misurazioni ripetute di campioni bianchi o bianchi reagente addizionati con basse concentrazioni di analita, o misurazioni ripetute di campioni di prova con basse concentrazioni di analita. Qualora questi bianchi reagente non siano trattati seguendo l'intera procedura di misura e siano sottoposti direttamente allo strumento, il calcolo fornirà il LOQ strumentale.

Per quanto riguarda il calcolo del LOQ, se l'esattezza e la precisione sono costanti in un intervallo di concentrazione prossimo al LOD, il LOQ è numericamente pari al decuplo della deviazione standard della media delle determinazioni in matrice $LOQ = 10 \times S'_0$. Per metodi che coinvolgono l'utilizzo di diverse matrici, è necessario stabilire LOD e LOQ per ciascuna matrice.

La determinazione dei LOQ attraverso il calcolo delle deviazioni standard di campioni ripetuti vicini allo zero strumentale (valore critico) è ampiamente utilizzato, ma talvolta si possono considerare altri approcci nella validazione dei metodi bioanalitici. Ad esempio, negli studi di validazione, la determinazione del cut-off strumentale può essere ricavata direttamente dalla curva di calibrazione²³ oppure inserendo dei controlli al di sotto e al di sopra del limite identificato e valutando la percentuale di falsi negativi (< 5%) ottenuta per ciascun punto.

2.4. Intervallo di misurazione e linearità

La curva di calibrazione descrive la relazione tra la concentrazione nominale di uno specifico analita e la risposta della piattaforma analitica utilizzata, e viene allestita utilizzando degli opportuni standard di calibrazione o calibratori. Questi ultimi vengono ottenuti addizionando la matrice biologica in esame con quantità note di uno o più analiti. Sarebbe preferibile usare matrici provenienti da almeno sei soggetti diversi ad eccezione delle situazioni che prevedono la preparazione dei calibratori in matrici rare. In questo contesto va sottolineato, tuttavia, che nella maggior parte dei casi l'aggiunta dello standard di calibrazione alla matrice non permette all'analita aggiunto di legarsi alla matrice tanto fortemente quanto sarebbe avvenuto se l'analita fosse stato naturalmente presente nel campione. Ciò potrebbe inficiare la stima dei campioni "reali" soprattutto in caso di forte legame tra la matrice analizzata e la molecola in esame. Durante la validazione dovrebbero essere valutati sia l'intervallo di misura strumentale che l'intervallo di misura del metodo. Per valutare l'intervallo di misura strumentale e confermarne l'idoneità allo scopo occorre utilizzare campioni di taratura con concentrazioni

che vadano oltre ($\pm 10\%$ o anche $\pm 20\%$) l'intervallo di concentrazione atteso e riportare i segnali in un grafico; le concentrazioni scelte dovrebbero essere distribuite in maniera approssimativamente equidistante all'interno dell'intervallo di misura.

La valutazione iniziale dell'intervallo di misura strumentale si effettua mediante un'ispezione visiva della curva di risposta²². La fase successiva è la conferma della relazione tra la concentrazione e la risposta strumentale attraverso l'esame dei parametri statistici della regressione e del diagramma dei residui per il modello scelto. L'intervallo di misura del metodo deve essere definito per ogni matrice prevista nel campo di applicazione. Le interferenze attribuibili alla matrice, infatti, possono influenzare la linearità della risposta, mentre la resa delle diverse fasi di estrazione/recupero dell'analita può variare al variare della matrice del campione.

Il materiale utilizzato per la preparazione delle curve di calibrazione deve essere del massimo grado di purezza disponibile. I calibratori possono essere acquistati, se commercialmente disponibili, oppure preparati dal laboratorio. Quando è possibile si raccomanda di utilizzare CRM (sciolti in matrice o anche in solvente) per validare i calibratori utilizzati, in quanto il valore dei CRM è assegnato mediante metodi di riferimento²⁰.

L'intervallo di calibrazione può essere definito tra i valori di LLOQ e ULOQ, ma come già sottolineato precedentemente, è possibile utilizzare un primo punto di calibrazione con valore più alto del LLOQ se quest'ultimo è lontano dall'intervallo di interesse clinico.

La curva di calibrazione include un campione bianco, un campione "zero" (campione bianco arricchito con SI) e almeno 6-8 livelli di calibrazione¹⁹.

La risposta strumentale (segnale, area/altezza del picco) deve essere rappresentata in un diagramma cartesiano e plottata contro il valore del relativo standard di calibrazione. La curva di calibrazione può essere rappresentata da un semplice modello di regressione o da altri modelli matematici. L'equazione che generalmente descrive la curva di calibrazione nell'intervallo del metodo è quella lineare, in cui la relazione tra la grandezza X (variabile indipendente) e quella Y (variabile dipendente) è descritta dall'equazione della retta $Y=aX+b$ dove a rappresenta la pendenza della retta e b è l'intercetta sull'asse delle Y .

Ciascuno standard di calibrazione dovrebbe essere analizzato in triplicato, e i dati provenienti da tutti i replicati accettabili dovrebbero essere utilizzati nell'analisi di regressione²⁴. Devono essere considerate tutte le curve accettabili ottenute durante la convalida della metodica, sulla base di un minimo di tre esecuzioni indipendenti nell'arco di diversi giorni, riportando i parametri della curva di calibrazione ovvero i valori di R^2 (ideale > 0.995), pendenza e intercetta nel caso di modello lineare.

L'assunzione che viene fatta per la curva di calibrazione è che l'errore di misura sia lo stesso e uniformemente distribuito per ogni campione. Se questo tipo di assunzione non è applicabile bisogna procedere con un'analisi estesa o con l'analisi dei minimi quadrati pesati. Le caratteristiche di distribuzione dei residui dovrebbero essere infatti normali o centrate intorno al valore dello zero (test Kolmogorov-Smirnov). In caso contrario si dovrebbe optare per un modello di regressione non lineare che richiede un maggior numero di standard di calibrazione²⁵.

L'esattezza delle concentrazioni retro-calcolate di ciascuno standard di calibrazione dovrebbe essere entro $\pm 20\%$ della concentrazione nominale al LLOQ ed entro $\pm 15\%$ a tutti gli altri livelli. Questo criterio deve essere soddisfatto da almeno il 75% dei punti di calibrazione se la curva è processata una sola volta e dal 50% dei punti di calibrazione se la curva è processata due volte durante la seduta analitica¹⁹.

Nel caso in cui si utilizzino repliche, i criteri (entro $\pm 15\%$ o $\pm 20\%$ per LLOQ) dovrebbero essere soddisfatti anche per almeno il 50% degli standard di calibrazione testati per livello di concentrazione. Nel caso in cui uno standard di calibrazione non soddisfi questi criteri, questo campione di standard di calibrazione dovrebbe essere rifiutato e la curva di calibrazione senza tale standard di calibrazione dovrebbe essere rivalutata, ripetendo al contempo l'analisi di regressione.

I campioni relativi alle curve di calibrazione possono essere congelati e utilizzati in futuro, se si è dimostrata la stabilità dell'analita a quelle condizioni²⁴.

Nel caso di molecole endogene è possibile utilizzare diversi approcci in base alla disponibilità di materiali:

- se sono disponibili in commercio, si raccomanda di utilizzare curve di calibrazione certificate preparate in matrice;
- se non sono disponibili standard commerciali, aggiungere quantità note di standard ad una "matrice surrogata";
- calibrazione con doppia marcatura. Se disponibili standard con diverso numero di isotopi (es D_x/D_y con $x \geq 3$ e $y-x \geq 3$), preparare la curva di calibrazione aggiungendo quantità note di standard D_x (contenente x deuteri ad esempio) alla matrice e utilizzare lo standard D_y come standard interno;
- nel caso in cui non fosse possibile preparare una "matrice surrogata" con le caratteristiche necessarie, è possibile preparare le curve di calibrazione in solvente dopo averle accuratamente confrontate con quelle preparate in matrice eseguita con il

metodo delle aggiunte. La differenza delle pendenze e delle intercette deve essere inferiore al 5%¹.

2.5. Valutazione dell'accuratezza: esattezza e precisione

Nel corso degli anni, a seconda dei documenti e delle linee guida consultate, vi è stata una forte ambiguità nella definizione dei parametri che definiscono lo scostamento di misura nell'ambito dei metodi bioanalitici. Tuttora, l'utilizzo del termine "accuratezza" da parte di riferimenti importanti, come le linee guida EMA e FDA, come "grado di vicinanza medio al valore vero o nominale" è frutto di questa ambiguità tradizionale.

A tal proposito, la normativa ISO 5725²⁶ ed il VIM si sono espressi a favore dell'utilizzo del termine "esattezza" per definire tale concetto (e quindi di inesattezza come "scostamento sistematico"), indicando invece "accuratezza" come la capacità del metodo di fornire risultati insieme "esatti" e "precisi", quindi come una grandezza non direttamente misurabile, che viene definita congiuntamente da esattezza e precisione. Di seguito, definiremo i protocolli sperimentali per la valutazione di questi parametri secondo la definizione della ISO 5725 e del VIM.

2.5.1. Scelta dei livelli dei controlli

Tutte le linee guida (EMA, CLSI e FDA) sono concordi nel suggerire, in fase di validazione del metodo, l'utilizzo di campioni o materiali di riferimento (idealmente identici ai campioni biologici analizzati) da utilizzarsi come CQs, che coprano l'intero range di calibrazione considerato durante la validazione. In particolare, le linee guida EMA suggeriscono che la valutazione di esattezza e precisione sia effettuata su cinque diversi livelli: 1) al livello del LLOQ, 2) ad una concentrazione bassa (inferiore a 3 x LLOQ), 3) ad una concentrazione media tra LLOQ e ULOQ, 4) ad una concentrazione alta (circa l'80% del ULOQ) e 5) al livello del ULOQ.

2.5.2. Esattezza

Secondo VIM e ISO 5725, l'esattezza o inesattezza (bias) rappresenta il grado di concordanza tra il valore medio ottenuto a partire da diverse determinazioni dello stesso misurando e un valore di riferimento accettato (detto anche valore vero del misurando). Essa è spesso espressa come bias (inesattezza), ossia come scostamento sistematico rispetto al valore vero (la differenza tra il valore ottenuto e il valore vero) che può quindi essere positivo oppure

negativo. Qualora il valore “vero” non possa essere conosciuto (nella maggior parte dei casi), si utilizza come riferimento il valore nominale.

L'esattezza della misura è fortemente influenzata dalla calibrazione. La matrice utilizzata per la preparazione dei calibratori (solvente, urine, plasma, albumina, ecc.), la concentrazione degli analiti e altri fattori contribuiscono a determinare un errore che incide sull'esattezza della misura. È pertanto preferibile utilizzare calibratori tracciabili, caratterizzati da un grado di incertezza definito.

L'esattezza può essere valutata utilizzando materiali che abbiano un valore di riferimento assegnato come ad esempio CQs interni e/o esterni o materiale di riferimento certificato, come ad esempio i materiali certificati del National Institute of Standards & Technology.

Esistono diversi protocolli di valutazione dell'esattezza. Nel caso di protocolli di validazione è necessario poter includere il maggior numero di fattori influenti sul metodo nella validazione (lotti o differenti preparazioni fatte in casa di reattivi/calibratori, differenti lotti/preparazioni delle fasi mobili, etc.). Maggiore è l'estensione del protocollo sulla scala temporale (aumento dei giorni di valutazione), maggiore sarà la presenza di variabili che potrebbero essere monitorate, pertanto, un protocollo minimo 5x3 (“replicati x giorni” rispettivamente, come suggerito da EMA) o 5x5 spesso non è sufficiente ad includere tutte le variabili. Dal punto di vista del monitoraggio a lungo termine dell'esattezza, potrebbe essere più propriamente consigliato un protocollo 2x20 o maggiore²⁷.

L'esattezza così misurata può essere espressa in diversi modi, intercambiabili ma che non devono essere confusi:

Esattezza %: $[(\text{valore medio delle misurazioni} / \text{valore vero}) * 100]$

Inesattezza (Bias)%: $[(\text{valore medio delle misurazioni} - \text{valore vero}) / \text{valore vero}] * 100$.

Il bias% per ogni livello di concentrazione dei materiali analizzati non deve essere superiore al 15% (esattezza 85 – 115 %), ad eccezione del bias al livello del LLOQ, che non deve essere superiore al 20% (esattezza 80 – 120 %). Questi range di accettabilità possono essere adattati qualora esistano linee guida per lo specifico parametro che definiscano dei valori più stringenti (es. criteri basati sulla variabilità biologica)²⁸.

2.5.3 Precisione

La precisione di un metodo analitico è definita come il grado di concordanza fra misurazioni indipendenti di concentrazione di un analita e deve includere tutto il processo di analisi, dalla raccolta e conservazione del campione, alla preparazione (estrazione e purificazione) fino all'analisi strumentale (2-5). Per una valutazione adeguata della precisione devono essere preparate e conservate, nelle stesse e opportune condizioni (previa verifica della stabilità dell'analita nelle condizioni di conservazione e di lavoro), un numero di aliquote di campioni di CQ, alle concentrazioni indicate (si veda paragrafo 2.5.1 "scelta dei livelli dei controlli"). Come specificato per l'esattezza, è possibile utilizzare materiali di riferimento disponibili commercialmente. Sarà importante verificare che tra i materiali di controllo e i campioni dei pazienti non esistano delle discrepanze di matrice e, qualora queste siano inevitabili, sarà fondamentale aver effettuato preventivamente una valutazione adeguata dell'effetto matrice. Tuttavia, con il termine "precisione" si intendono due livelli di concordanza: quello tra replicati all'interno della stessa seduta analitica e nella stessa giornata ("intra-run" e/o "intra-day", anche detta "ripetibilità") e quello tra diverse sessioni e/o giornate ("inter-run" e/o "inter-day", anche detta "riproducibilità").

In entrambi i casi, un metodo si ritiene preciso qualora il "coefficiente di variazione" (CV%, misura di imprecisione) dei replicati "intra-day" e quello "inter-day" siano entro il 15% per ogni livello di concentrazione, ad eccezione di quello del LLOQ, in cui è considerato accettabile un CV% del 20% ^{4, 28, 29}. Il CV% può essere calcolato secondo la seguente formula:

$$\%CV = (DS \text{ tra i replicati} / \text{Media dei replicati}) \times 100.$$

Sebbene nelle linee guida EMA si parli di "intra-run" ed "inter-run", è consigliabile effettuare la valutazione della precisione come "inter-day", ovvero valutarla in giornate diverse, utilizzando operatori, calibratori e reagenti differenti. In accordo con le linee guida EMA è necessario, come precedentemente indicato per quanto riguarda l'esattezza, l'utilizzo di un protocollo di valutazione minimo di 5x3 (replicati x giorni); ciononostante, tenendo conto dell'importanza della valutazione sul medio periodo, come suggerito in letteratura, è auspicabile un'estensione minima del protocollo a 5x5. Dal punto di vista del monitoraggio a medio termine della precisione, potrebbe essere più propriamente consigliato un protocollo 2x20 o maggiore²⁷.

Infine, la valutazione dei valori di CV% accettabili dovrebbe tener conto del contesto diagnostico: in alcuni contesti (raramente) valori d'imprecisione lievemente superiori al 15%

potrebbero essere biologicamente accettabili mentre, in altri contesti in cui un'eccellente ripetibilità e riproducibilità siano vitali dal punto di vista diagnostico (eg. monitoraggio di marcatori tumorali in corso di terapia), sarà consigliabile considerare valori limite di accettazione per il CV% nettamente inferiori.

Tali considerazioni sono particolarmente vere qualora esistano documenti riconosciuti validi come linee guida nazionali o internazionali specifiche per una determinata sostanza o classe di sostanze (es. farmaci), oppure criteri universalmente riconosciuti basati sulla variabilità biologica.

2.6. Recupero

L'elevato recupero dell'analita dalla matrice è un risultato desiderabile della preparazione del campione, ed è quindi una caratteristica importante della procedura di estrazione. La valutazione del recupero è suddivisa in recupero assoluto, recupero relativo e recupero dell'estrazione.

Il *recupero assoluto* è il rapporto percentuale della risposta del rivelatore ottenuto da un campione di matrice “bianca” fortificata con l'analita, trattata seguendo l'intera procedura analitica, e una soluzione standard direttamente iniettata nello strumento, che rappresenta il recupero del 100%.

Quando non è disponibile una matrice “bianca”, si consiglia di utilizzare almeno 6 campioni con bassa concentrazione di analita misurata prima della procedura di recupero.

$$\text{Recupero assoluto} = [\text{risposta dell'analita inserita in campione "bianco" (processato)} / \text{risposta dell'analita dello standard puro (non processato)}] * 100$$

Il recupero assoluto dovrebbe essere effettuato a 3 concentrazioni diverse (per esempio seguendo i livelli di riferimento per i CQ, che a loro volta dovrebbero rientrare nel range di interesse dell'analita in questione), per 5 replicati.

Il *recupero relativo* è il rapporto percentuale della risposta del rivelatore ottenuto da un campione bianco con l'aggiunta dell'analita e una soluzione standard alla stessa concentrazione, entrambi trattati secondo l'intera procedura analitica.

$$\text{Recupero relativo} = [\text{risposta dell'analita inserita in campione bianco (processato)} / \text{risposta dell'analita nella soluzione standard (processata)}] * 100$$

Per entrambi, il recupero dell'analita non deve essere necessariamente al 100%, i criteri di accettazione sono soddisfatti quando i risultati sono tra 90-110%. Il *recupero dell'estrazione* è il rapporto percentuale della risposta del rivelatore ottenuta da un campione di matrice bianca fortificata con l'analita prima del trattamento e lo stesso campione bianco fortificato con l'analita, alla stessa concentrazione nominale, dopo essere stato trattato secondo l'intera procedura analitica.

$$\text{Recupero dell'estrazione} = [\text{risposta dell'analita spiked prima dell'estrazione (processato)} / \text{risposta dell'analita spiked dopo l'estrazione}] * 100$$

Il recupero dell'estrazione di un analita e del suo standard interno devono essere coerenti e riproducibili (criteri di accettazione 60-140%), possibilmente con un errore standard relativo inferiore al 10%.

Una percentuale minore o più elevata di recupero (<60% e >140%) potrebbe essere accettata solo se l'errore standard relativo è inferiore al 5% per entrambi (analita e SI), indicando un'elevata riproducibilità e robustezza del metodo. Il recupero <60% e >140% deve comunque essere valutato attentamente e giustificato. Per valutare se le perdite di analita nell'estrazione sono dovute a effetti matrice (vedere dopo) o proprio a scarsa estrazione, tutti gli esperimenti del recupero dovrebbero essere eseguiti durante lo sviluppo del metodo. Gli esperimenti relativi al recupero dovrebbero essere effettuati e determinati per almeno 6 campioni di pazienti o 6 diversi lotti di matrice commerciale, per almeno 10 determinazioni nell'intervallo delle concentrazioni previste, utilizzando anche matrici bianche con lipidi e matrici emolizzate, o campioni di pazienti (in particolare, l'aggiunta di concentrazioni note di analiti deve essere eseguita e valutata [lo stesso per valutare il recupero] con i campioni bianchi e per i pazienti con iperlipidemia ed emolisi). Il bias accettabile è < del 15%. Quando non è disponibile un campione "bianco" (per analiti endogeni) si dovrebbe utilizzare una matrice alternativa; ad esempio, matrice trattata appositamente o matrice surrogata, come siero bovino o umano in tampone fosfato (phosphate buffered saline, PBS).

2.7 Test della integrità della diluizione

In alcuni casi può essere necessario diluire il campione: per esempio quando si ha una concentrazione sopra l'ULOQ, quando sono disponibili solo piccole quantità di campione, oppure dopo la preparazione del campione per adattarsi alle condizioni cromatografiche (caratteristiche del loop dell'autocampionatore, saturazione della colonna, etc.).

La valutazione della “*dilution integrity*” o test della integrità della diluizione deve far parte anche della validazione parziale in quanto può influenzare l’accuratezza, il *bias* e la precisione. Il test della “*dilution integrity*” deve essere effettuato considerando i fattori (da 3 a 5) di diluizione affinché le concentrazioni dell’analita possano ricadere nel range di calibrazione della metodica. Per ciascun fattore occorre eseguire almeno cinque replicati nella stessa seduta analitica. L’inaccuratezza o il bias che ne derivano devono essere al di sotto del 15%. Se la validazione della diluizione non viene effettuata, campioni con delle concentrazioni al di sopra del più alto punto della curva di calibrazione, devono essere riportati sul referto come “al di sopra del punto di calibrazione più elevato” o come “maggiore di ULOQ”. Qualora la diluizione venga effettuata nell’ambito dell’analisi, è auspicabile riportare tale passaggio nel referto.

2.8 Effetto Matrice, interferenze endogene ed esogene

L’EM è uno dei più importanti parametri da investigare durante lo sviluppo di un metodo in spettrometria di massa; tale effetto può dare origine ad una soppressione o, più raramente, ad un incremento del segnale³⁰. L’EM è collegato all’eluizione di componenti della matrice del campione in analisi, quali i fosfolipidi, gli acidi organici, sali, ecc., capaci di influenzare lo stato di carica all’interno della sorgente ionica o di causare una co-precipitazione in sorgente durante la desolvatazione. Questo effetto può condizionare altri parametri quali LOD, LLOQ, ma soprattutto la precisione e l’esattezza del metodo. Per questo motivo l’EM dovrebbe essere attentamente studiato durante la validazione del metodo.

I campioni utilizzati per lo studio dell’EM devono rispecchiare la tipologia di campioni per cui il metodo è stato sviluppato partendo dalle medesime matrici biologiche (plasma, sangue intero, saliva, urine, biopsie, etc.); occorre valutare specifiche condizioni clinico/patologiche (campioni lipemici, disfunzioni metaboliche, etc.), campioni post mortem, etc. Nel caso di matrici derivanti da sangue (sangue intero, plasma, siero) è altamente consigliata la valutazione dell’EM dei principali contaminanti (inclusi emolisi, iperlipidemia, iperbilirubinemia). Ogni valutazione sull’EM deve essere effettuata utilizzando lo stesso tipo di provetta da prelievo (con il medesimo anticoagulante) che verrà utilizzata in fase di routine; nel caso in cui si vogliano utilizzare più tipi di provette, bisogna effettuare la valutazione dell’EM per ogni tipo di agente anticoagulante o additivo presente nelle provette da prelievo che si vogliono validare. Infine, qualora si vogliano misurare le concentrazioni di farmaci, che possono essere somministrati insieme ad eccipienti che possono avere un impatto sul EM, sarebbe opportuno valutare il loro effetto aggiungendo tali sostanze nelle matrici “bianche”

utilizzate per la valutazione dell'EM, prima dell'estrazione. Idealmente, durante lo sviluppo del metodo, bisognerebbe puntare ad ottenere un EM% medio più basso possibile ed il più possibile stabile tra i diversi lotti di matrice (vedi prossimi paragrafi), lavorando con una curva di calibrazione preparata in matrice. Qualora si rilevi un EM% potenzialmente impattante sulle performance del metodo, dovrebbero essere considerate modifiche al metodo cromatografico, alla preparazione del campione, allo SI (considerando come opzione ideale l'uso di SI marcati con isotopi stabili) o anche semplicemente al volume d'iniezione o alla diluizione del campione, ove la sensibilità strumentale lo consenta.

Inoltre è importante precisare che queste valutazioni sperimentali sull'EM, particolarmente in vista della nuova normativa, devono essere considerate estremamente specifiche per la configurazione strumentale del sistema LC-MS utilizzato per la validazione, e non esportabili su altre piattaforme: in caso di cambiamento del set-up strumentale, sarà necessario effettuare nuove prove di EM. Questo punto risulta particolarmente importante, poiché diverse tecniche di ionizzazione o anche solamente diverse architetture nella sorgente possono avere un effetto completamente differente sull'EM, portando a risultati molto diversi da strumento a strumento. Qualora l'EM sia difficile da eliminare, in quanto generato da fattori intrinseci (es. pH, sali, altri fattori), il laboratorio deve verificarne l'impatto su LLOQ, precisione ed esattezza dell'metodo. In particolare, a questo proposito è importante che il laboratorio valuti la capacità dello SI di eliminare/contenere l'impatto dell'EM residuo sui risultati, attraverso la valutazione dell'effetto matrice normalizzato dallo SI (nEM), come descritto nei prossimi paragrafi.

2.8.1 Valutazione Post-Estrazione

La valutazione post-estrazione dell'EM è il metodo di eccellenza, in quanto ne permette una misurazione diretta sia sugli analiti che sugli standard interni.

I campioni privi di analita che verranno fortificati post-estrazione devono essere preparati da almeno 6 lotti di campione differenti (es. plasma) e/o, se possibile, devono derivare da donatori differenti e dovrebbero essere analizzati in triplicato. La matrice priva di analita deve essere trattata secondo la metodica estrattiva prevista ed, in seguito, all'estratto viene aggiunta una quantità nota di analita e del suo SI. L'area dell'analita e dello SI ottenute da questo campione devono essere comparate con l'area ottenuta da una soluzione di analita in solvente (privo di interferenti di matrice biologica) con una concentrazione uguale al campione fortificato. Nel caso non sia disponibile una matrice priva dell'analita, è possibile eventualmente effettuare un'analisi su 6 lotti differenti di matrice addizionati con isotopi

stabili (^2H , ^{13}C , ^{15}N , etc.) dell'analita in oggetto, verificando accuratamente che i tempi di ritenzione di queste molecole siano identici a quelli degli analiti. In questo caso, prima di procedere con la valutazione dell'EM, è inoltre necessario verificare che l'analita non dia interferenze con l'isotopo stabile o vice-versa; questo per evitare che la presenza dell'analita nella matrice possa alterare i risultati del test. Per ogni lotto di matrice (o paziente) devono essere analizzate un minimo di 2 differenti concentrazioni rappresentative della distribuzione dei valori biologici/patologici o terapeutici/tossici ed a livelli alti e bassi della curva di calibrazione. Ad esempio, potrebbe essere utilizzato un valore pari a 3x il LLOQ (eg. pari al valore del CQ basso), uno in un punto intermedio della curva (o mediano tra LLOQ e ULOQ) ed uno vicino, ma al di sotto, del punto più alto della curva di calibrazione (eg. pari al valore del CQ alto) o all' ULOQ.

L'effetto matrice percentuale è calcolato indipendentemente per ogni lotto di matrice/paziente e per ogni analita nel seguente modo³⁰:

$$EM\%: [(A_{x \text{ matrice}} / A_{x \text{ solvente}}) - 1] * 100$$

$A_{x \text{ matrice}}$ = Area dell'analita "x" addizionato al campione con matrice dopo estrazione

$A_{x \text{ Solvente}}$ = Area dell'analita "x" in solvente

Si suggerisce di applicare il fattore “-1” esclusivamente a fini descrittivi, al fine di rappresentare in modo intuitivo la direzionalità dell'EM: ad esempio, un rapporto tra segnali del 98% viene descritto come -2% di effetto matrice medio. È tuttavia importante non utilizzare tale fattore “-1” per la valutazione della variabilità dell'effetto fra le diverse matrici testate (CV%), allo scopo di evitare di rendere il parametro matematicamente non rappresentativo della reale dispersione dei dati (forte sovrastima).

La media degli EM% è comunemente utilizzata per la valutazione della presenza di EM. Per alcune linee guida, quali FDA, Scientific Working Group for Forensic Toxicology, Clinical & Laboratory Standards Institute C62-A³¹, la valutazione dell'EM% medio è l'unica valutazione richiesta in tema di EM. Tuttavia, poiché la quantificazione di un analita viene effettuata valutando il rapporto tra l'area dello stesso e quella del suo SI, è facile comprendere come l'EM delle due molecole debbano essere messi in relazione tra loro per valutare l'effettivo impatto dell'EM sulle performance del metodo. Su questo punto, solo le linee guida EMA forniscono una indicazione in merito, senza però entrare nello specifico su come effettuare la valutazione a livello operativo. La valutazione della relazione degli EM di analita e SI è particolarmente importante nel caso in cui si utilizzino SI non marcati con isotopi stabili o marcati con isotopi stabili la cui differente idrofilia rispetto all'analita (ad es. SI marcati con

deuteri ^2H) potrebbe dare una sufficiente separazione cromatografica tale da non sovrapporre perfettamente i cromatogrammi di SI ed analita^{32, 33 34}. Per questo motivo la valutazione di un EM normalizzato dall'SI (nEM, anche chiamato “fattore matrice” nelle linee guida EMA) diventa importante al fine di valutare tutti gli analiti coinvolti nel metodo quantitativo (analita e SI) e, soprattutto, valutare l'appropriatezza dell'SI scelto per ogni misurando. La formula per il calcolo del nEM è così indicata:

$$nEM\%: [(A_{x \text{ Matrice}}/A_{SI \text{ Matrice}})/(A_{x \text{ Solvente}}/A_{SI \text{ Solvente}}) - 1] * 100$$

$A_{x \text{ matrice}}$ = Area dell'analita "x" addizionato al campione con matrice dopo estrazione

$A_{x \text{ Solvente}}$ = Area dell'analita "x" in solvente

$A_{SI \text{ matrice}}$ = Area dell'SI dell'analita "x" addizionato nel campione con matrice dopo estrazione

$A_{SI \text{ Solvente}}$ = Area dell'SI dell'analita X in solvente

Come per l'EM%, anche in questo caso, il fattore “-1” viene utilizzato unicamente per determinare la direzionalità della media del nEM%, ma non deve essere applicato per il calcolo del CV%, allo scopo di evitare una sovrastima notevole della sua variabilità.

L'nEM medio dovrebbe essere preferibilmente compreso nel range $\pm 15\%$, in caso contrario è importante valutare che questo non abbia un impatto significativo sugli altri parametri di validazione del metodo (in particolare la sensibilità e l'esattezza). Tuttavia, il più importante parametro da tenere in considerazione per una corretta valutazione dei possibili errori nella quantificazione causati dall'EM, come correttamente sottolineato dalle linee guida EMA, è il coefficiente di variazione percentuale (CV%) di tutti i singoli SI-nEM%, calcolato su almeno 6 lotti diversi di matrice (oltre a 2 lotti, analizzati in triplicato, per eventuali matrici “alterate”, come ad esempio in caso di plasma con emolisi, lipemico, etc.). Il CV% dell'nEM% è il parametro che rappresenta realmente la robustezza del metodo in termini di EM nel suo insieme, particolarmente quando si lavora con curve di calibrazione e controlli di qualità realizzati in matrice (“gold-standard”), che sono in grado di compensare l'nEM% medio. Per tale motivo, il CV% dovrebbe essere contenuto possibilmente entro il 10% (in particolare per le matrici normali) e, comunque, non superare il 15% (comprendendo anche matrici alterate, più critiche) Il report finale di validazione del metodo deve quindi riportare i valori medi di SI-nEM% ed il suo CV%].

2.8.2 Valutazione con Infusione Post-Colonna

In aggiunta al metodo dell'addizione post-estrattiva, un'altra tecnica che può essere utile per la valutazione dell'EM è l'infusione post-colonna dell'analita³⁵. Tale tecnica si basa sull'infusione di una concentrazione costante di analita in sorgente, combinata con il flusso di fasi mobili provenienti dalla colonna: in tal modo, il segnale dell'analita (idealmente costante) varierà durante la corsa cromatografica a seconda di quanto proviene dalla colonna cromatografica. Ne deriva una valutazione quali-quantitativa combinata dell'EM causato dalle fasi mobili stesse e da eventuali composti co-eluenti provenienti dalla matrice biologica. Per poter "isolare" l'EM dovuto ai soli composti provenienti dalla matrice biologica bisognerà, dunque, confrontare le tracce derivanti dall'iniezione dello stesso volume di solvente puro e di estratti di matrice, in entrambi i casi privi di analiti ed SI. Tale metodo non è, tuttavia, da intendersi come "alternativo" al precedente, poiché non fornisce una valutazione quantitativa dell'EM, ma definisce solo dei parametri qualitativi. Infatti, il suo utilizzo può fornire informazioni complementari al metodo dell'addizione post-estrattiva, evidenziando la posizione temporale delle aree di alto EM rispetto ai tempi di ritenzione degli analiti e degli SI. Contemporaneamente, può offrire indicazioni sulla robustezza del metodo nel tempo, ad esempio quando i tempi di ritenzione risultano distanti dalle aree di alto EM sia per l'analita sia per il SI, o quando vi è una buona concordanza tra i due. Queste informazioni si rivelano particolarmente utili durante le fasi di sviluppo del metodo.

A livello operativo, un campione di solvente privo di analita e 6 differenti lotti di matrice devono essere preparati secondo il metodo in validazione, ma senza l'aggiunta dello standard interno, ed iniettati nel sistema LC-MS. Contemporaneamente, una concentrazione nota di analita (o SI) viene infusa mediante una pompa siringa direttamente nel flusso di eluente, appena prima della sorgente ionica. La matrice influenzerà la ionizzazione dell'analita dispensato dalla pompa siringa producendo un cromatogramma con possibili picchi sia positivi che negativi. Tutti i campioni preparati dovrebbero essere testati a 3 differenti concentrazioni (vedi paragrafo precedente), rappresentative del range di calibrazione.

Poiché la valutazione dell'EM post-colonna è una valutazione qualitativa, all'interno del report deve essere inserita una dettagliata descrizione di tutti gli analiti coinvolti (SI incluso) in termini di profilo cromatografico, tempi di ritenzione ed intensità dell'EM osservato in relazione ai tempi di ritenzione dell'analita e del suo SI con osservazioni rispetto all'eventuale variazione dell'EM in base alla concentrazione dell'analita infuso. Non essendo un metodo prettamente quantitativo, non è possibile definire dei valori di accettabilità; pertanto in caso di EM presenti ai tempi di ritenzione dell'analita e/o dello SI si consiglia di modificare le

condizioni cromatografiche, il metodo di estrazione e di utilizzare SI marcati con isotopi stabili, in particolare, se in UHPLC, con ^{13}C o ^{15}N .

2.8.3 Comparazione dei coefficienti angolari delle curve di calibrazione

Quando l'applicazione delle tecniche di addizione post-estrattiva e di infusione post-colonna risulta difficoltosa, ad esempio per l'analisi di composti endogeni in cui non è possibile ottenere matrici prive di analita o evitare l'EM generato dal composto endogeno, e la calibrazione viene effettuata con uno standard marcato, una terza opzione consiste nell'utilizzo della tecnica del confronto dei coefficienti angolari delle curve di calibrazione^{36,37}. Tale tecnica prevede la preparazione di curve di calibrazione direttamente in solvente ed in almeno 6 lotti diversi della matrice da valutare (eventualmente comprensive di matrici "alterate"), sottoposti ad estrazione ed analisi. La presenza di nEM sarà evidenziata come una differenza tra le pendenze (coefficienti angolari o "slope") della curva preparata in solvente e quelle in matrice. Inoltre, il CV% dei coefficienti angolari delle curve in matrice sarà interpretabile come una buona stima del CV% del nEM. Anche in questo caso, idealmente la deviazione media tra i coefficienti angolari ottenuti per le curve in matrice ed in solvente (stima del nEM% medio) dovrebbe rientrare nel $\pm 15\%$, ed il CV% dovrebbe essere contenuto entro il 10% e comunque sempre inferiore al 15%.

Tuttavia, sebbene tale approccio sia pratico e molto popolare, esso valuta l'EM dal punto di vista complessivo (di "efficienza di processo"), ovvero una combinazione della variabilità apportata dal nEM, dal recupero dell'analita e dell'SI in fase estrattiva. Di conseguenza, se emergono problematiche (ad esempio un CV superiore al 10%), è necessaria una valutazione approfondita mediante metodiche alternative (infusione post-colonna e/o addizione post-estrattiva) al fine di chiarire le cause e applicare le opportune modifiche metodologiche.

2.9. Stabilità

La stabilità di un analita è funzione delle sue proprietà fisico-chimiche, delle condizioni di conservazione e della matrice in cui esso è presente.

Lo studio della stabilità è una parte importante del metodo di validazione al fine di garantire una corretta gestione di tutti i reagenti e dei campioni durante l'intero processo di analisi^{4, 11, 12, 38}. Le valutazioni di stabilità devono garantire che ogni fase del processo (preparazione, conservazione, elaborazione e analisi del campione) non influisca sulla concentrazione dell'analita in esame. Lo studio della stabilità deve valutare tutte le soluzioni, reagenti e campioni durante:

- raccolta dei campioni (prelievo)
- stoccaggio dei campioni a breve termine
- stoccaggio dei campioni a lungo termine
- cicli di congelamento/scongelamento
- stoccaggio dopo processamento/estrazione (prima dell'analisi)

I parametri e le condizioni da valutare dovrebbero eccedere o al limite coprire i normali tempi delle procedure di campionamento ed analisi (come stimato dalla normale gestione del laboratorio). Dovrebbero riguardare le condizioni del campione previste prima del ricevimento, all'arrivo e durante l'analisi in laboratorio. In generale le condizioni minime da valutare sono: stabilità a differenti tempi di conservazione ed a differenti temperature, stabilità alla luce (indiretta ed eventualmente diretta), stabilità a ripetuti cicli di congelamento/scongelamento. La convalida dei tempi di conservazione deve essere eseguita su controlli di stabilità che siano stati conservati per un tempo uguale o superiore ai periodi di conservazione dei campioni di studio.

La stabilità dell'analita deve essere determinata analizzando 3 aliquote indipendenti ad almeno 2 concentrazioni differenti per ogni condizione di conservazione ipotizzata, preferibilmente includendo un controllo a bassa concentrazione (ad esempio circa 3 volte il LLOQ, basso QC) e uno ad alta concentrazione (vicino all'ULOQ, alto QC).

I controlli di stabilità devono essere analizzati utilizzando una curva di calibrazione ottenuta da standard appena preparati (tempo "zero") con i relativi CQ o con CQ per i quali sia stata già dimostrata la stabilità. Un campione è considerato "stabile" ad una specifica condizione se la concentrazione media di ogni livello rimane entro $\pm 15\%$ della concentrazione determinata al tempo zero¹¹. Una valutazione della stabilità basata solo su dati pubblicati non può essere considerata completa se non sufficientemente descritta e verificata. Se le condizioni di conservazione cambiano o l'analisi del campione viene modificata andando al di fuori delle condizioni di conservazione stabilite, la stabilità deve essere verificata alle nuove condizioni⁴.

2.9.1 Stabilità delle soluzioni e dei reagenti

La stabilità delle soluzioni standard di stoccaggio e di lavoro, sia degli analiti che degli SI, delle soluzioni di standard in matrice con e senza analiti, delle soluzioni di reagenti preparate in laboratorio (miscele di solventi, tamponi, soluzioni saline, etc.) devono essere valutate rispetto a soluzioni preparate al momento (da standard in polveri o liquidi). Per le condizioni si veda il paragrafo 2.9.4 "stabilità a lungo termine".

La stabilità delle soluzioni di stoccaggio e di lavoro dell'analita e dello standard interno deve essere determinata nelle condizioni di conservazione utilizzate durante l'analisi dei campioni in studio, usando le concentrazioni più basse e più alte di queste soluzioni. La stabilità delle soluzioni di stoccaggio e di lavoro deve essere testata con una diluizione adeguata, tenendo conto della linearità della curva di calibrazione della metodica analitica utilizzata. Nel caso in cui la stabilità vari con la concentrazione si deve valutare la stabilità di tutte le concentrazioni delle soluzioni di stoccaggio e delle soluzioni di lavoro. Le soluzioni di stoccaggio non dovrebbero essere ottenute da materiali di riferimento prossimi alla data di scadenza. La pratica di realizzare soluzioni di stoccaggio e soluzioni di lavoro a partire da materiali di riferimento al fine di estendere la data di scadenza per l'uso dello standard di riferimento non è accettabile⁴.

La stabilità di un analita in una particolare matrice è rilevante solo per quella matrice e non può essere estrapolata per altre matrici. Per le soluzioni allestite in laboratorio, ciascuna soluzione deve essere etichettata riportando le seguenti informazioni: nome del reagente preparato, data di preparazione, data di validità, operatore, condizioni di conservazione, materiale di partenza. La verifica della stabilità delle soluzioni preparate in laboratorio si basa su prove di riproducibilità dei CQ fino a variazioni del CV% del 15%¹¹.

2.9.2 Stabilità dopo cicli di congelamento/scongelamento

Nel caso sia necessaria una conservazione di reattivi, campioni, CQ o calibratori congelati, in particolar modo se in forma liquida, confezionati in quantitativi maggiori che per il singolo utilizzo, è necessario effettuare uno studio di stabilità ai ripetuti cicli di congelamento/scongelamento. I CQ devono essere scongelati e analizzati secondo le stesse procedure dei campioni di studio. I controlli di stabilità devono essere conservati congelati per almeno 12 ore tra i cicli di scongelamento, effettuando un minimo di tre cicli. Per ogni ciclo i CQ devono essere posti a scongelare a temperatura ambiente (o alla temperatura di processamento) fino al suo completo scongelamento^{4,11}.

2.9.3 Stabilità a breve termine e gestione del prelievo

Al fine di valutare la stabilità del campione primario (es. tubo di prelievo o del campionamento), bisogna aggiungere allo stesso, immediatamente dopo il prelievo (o appena possibile), una quantità nota di analita e conservarlo a temperatura ambiente (es. 25±5°C) per almeno 24 ore. Altri periodi di tempo possono essere valutati compatibilmente con il tempo atteso di arrivo del campione dalla sede del prelievo alla corretta sede di

stoccaggio/processamento in laboratorio. In condizioni speciali, ad esempio nel caso in cui il campione debba essere spedito e/o nel caso la stabilità a 25°C sia troppo ridotta rispetto ai tempi di consegna del campione al laboratorio è necessaria una valutazione anche a 35±5 °C e/o 4/8 °C e/o a -20 °C. Si raccomanda di eccedere notevolmente rispetto al tempo stimato, in quanto i campioni che non arriveranno in laboratorio in un tempo utile rispetto ai tempi di cui è stata valutata la stabilità, saranno da considerare non conformi e quindi non utilizzabili per l'analisi. L'analisi deve essere effettuata su una sola tipologia di prelievo; l'uso di differenti coadiuvanti (es. tubi litio-eparina, citrato, provette senza gel, etc.) o conservanti (es. riducenti) richiede una rivalutazione della stabilità dell'analita; inoltre, nel caso si debbano usare differenti coadiuvanti o conservanti, deve essere riconsiderato l'impatto sugli altri parametri di validazione. Il report di validazione deve dichiarare le condizioni di luce a cui viene condotto lo studio (es. conservazione al buio, luce non diretta, luce diretta). Lo studio deve essere condotto tutto alle stesse condizioni di luce per l'intera durata dello studio a lungo termine; eventuali altre condizioni di luce possono essere valutate e dichiarate nel report ^{4, 11}.

2.9.4 Stabilità a lungo termine

Tutti i reagenti, le soluzioni, i campioni, i CQ e i calibratori devono essere sottoposti ad una valutazione di stabilità a lungo termine^{11,39}. I CQ devono essere conservati nel congelatore nelle medesime condizioni di conservazione e almeno con la stessa durata dei campioni di studio. Questa valutazione può essere condotta, in base al tipo di soluzione testata, da 15 giorni (es. stabilità a lungo termine dei campioni) fino ad 1 anno o più (reagenti, soluzioni, controlli di qualità) a discrezione del laboratorio.

La stabilità a lungo termine deve essere anche testata a 25±5 °C. Sulla base dei risultati della stabilità a breve termine o dei risultati preliminari ottenuti a 25±5 °C, ulteriori studi possono essere effettuati al fine di valutare temperature più basse (4/8 °C, -20 °C, -80 °C).

Nel caso in cui la matrice biologica sia sangue intero, bisogna fare particolare attenzione alla stabilità dell'analita subito dopo la raccolta del campione e prima della preparazione per la conservazione per garantire che le concentrazioni misurate riflettano quelle presenti al momento del prelievo¹¹.

Nel caso in cui vengano utilizzati reattivi, campioni, controlli di qualità o calibratori liofilizzati, lo studio di stabilità a lungo termine deve essere condotto anche sui rispettivi risospesi/solubilizzati.

Il report di validazione deve dichiarare le condizioni di luce a cui viene condotto lo studio (es. conservazione al buio, luce non diretta, luce diretta). Lo studio deve essere condotto tutto alle

stesse condizioni di luce per l'intera durata dello studio a lungo termine; eventuali altre condizioni di luce possono essere valutate e dichiarate nel report.

2.9.5 Stabilità dopo processamento/estrazione (prima dell'analisi)

Una volta processato/estratto il campione, questo può richiedere tempo prima di poter essere analizzato. Tale tempo può essere il tempo di attesa nell'autocampionatore così come nel caso sia necessario rianalizzare gli stessi campioni il giorno successivo o a causa di un momentaneo malfunzionamento strumentale. Lo studio viene solitamente condotto a 4–8 °C, corrispondenti alla temperatura operativa tipica di un autocampionatore, per un minimo di 48 ore. Può comunque essere effettuato a temperature differenti, in base alle esigenze del laboratorio. Nel caso sia prevista la conservazione previa congelamento, è necessario condurre uno studio di stabilità sui cicli di congelamento e scongelamento del campione processato o estratto. Come per gli studi a breve e a lungo termine, anche in questo caso il report di validazione deve dichiarare le condizioni di luce adottate durante lo studio; eventuali altre condizioni di luce possono essere valutate e dichiarate nel report¹¹.

3. Utilizzo clinico di un metodo bioanalitico

La validità scientifica e la valutazione delle prestazioni cliniche, che potrebbero essere incluse nel ben più noto concetto di “validazione clinica”, implicano, nell'intenzione della normativa, una raccolta ed una attenta analisi di dati per stabilire e verificare l'effettivo beneficio clinico dell'applicazione di una metodica bioanalitica in ambito clinico, ovvero l'impatto positivo di un test correlato alla sua funzione, come quella di screening, monitoraggio, diagnosi o aiuto alla diagnosi dei pazienti, o l'impatto positivo che il test può avere in generale sulla gestione dei pazienti o sulla salute pubblica⁴⁰.

La validità scientifica si riferisce all'associazione tra un analita ed una condizione clinica o stato fisiologico. Può essere dimostrata utilizzando evidenze di letteratura, linee guida cliniche fornite da associazioni professionali, studi di *proof of concepts*, studi di performance clinica, o opinioni di *consensus expert*. Nella stesura del report relativo alla validità scientifica sarebbe consigliabile dimostrare, quando possibile ovviamente, che sono state considerate sia pubblicazioni a favore che pubblicazioni contrarie presenti nella letteratura scientifica internazionale³. Secondo diversi autori⁴¹, la performance clinica tende invece a focalizzare la propria attenzione sulla specifica popolazione in esame, sulla condizione clinica o sullo stato fisiologico. I dati possono essere ottenuti da studi clinici, da evidenze di letteratura, da

esperienza pubblicata ottenuta da testing diagnostico di routine e/o da prove valutative effettuate su una specifica popolazione di riferimento (pediatrica, geriatrica, con comorbidità) o da VEQ³.

In ogni caso, e tenendo conto dell'intenso dibattito attualmente presente in letteratura su questo specifico argomento⁴²⁻⁴⁴, gli autori non hanno ritenuto opportuno scindere le due categorie di "validità scientifica" e "prestazione clinica", ritenendo indispensabile compilare un report durante la validazione di una metodica bioanalitica basata sulla spettrometria di massa che includa tutti gli aspetti sopracitati, ivi compreso la definizione degli intervalli di riferimento, della variabilità intra- e inter-individuale e la valutazione delle variazioni che possono essere collegate ai diversi stati fisiologici e che possono essere specifici di una popolazione in esame. Ciò, nell'intenzione degli autori, diventa ad esempio di fondamentale importanza nel confronto tra metodiche sviluppate *in house* in laboratori diversi che oltre al confronto in termini di valutazione delle prestazioni analitiche, possono e debbono essere confrontate soprattutto sul piano delle potenziali diverse performance cliniche.

3.1 Raccolta e manipolazione di campioni clinici

La raccolta del campione biologico necessario all'analisi deve avvenire seguendo protocolli specifici per l'analisi in esame, utilizzando provette o contenitori idonei.

Il personale di laboratorio deve assicurarsi che i campioni ricevuti siano integri e conformi all'analisi richiesta. Ogni campione deve essere etichettato in modo univoco e deve essere attuato un sistema di tracciabilità di ciascun campione dall'arrivo all'analisi. Le condizioni di conservazione dei campioni devono essere definite in un protocollo analitico o istruzione di lavoro, il rispetto di tali condizioni deve essere controllato e documentato in modo appropriato.

A seconda dell'analisi che si deve svolgere, il campione raccolto può provenire da diverse matrici biologiche. Di seguito un elenco delle matrici coinvolte nei processi di analisi quantitativa in spettrometria di massa: sangue intero, plasma, siero, urine, DBS, DPS, CSF, biopsie, liquido di broncolavaggio, cellule, capelli, saliva, feci, latte materno. La raccolta e la manipolazione dei campioni biologici è un processo che deve essere supportato da specifiche procedure operative rese disponibili a tutti gli operatori del laboratorio, secondo le Politiche UNI EN ISO 15189³.

Di seguito vengono riportate alcune RACCOMANDAZIONI di carattere generale:

- tutti i campioni devono essere trattati come potenzialmente infetti;

- indossare sempre guanti monouso (o dispositivi di protezione individuale specifici indicati in procedura);
- il contenitore primario (es. provetta) per essere trasportato deve essere inserito all'interno di un contenitore idoneo per il trasporto (recipiente secondario) con simbolo di rischio biologico.

Ogni tipo di campione biologico deve essere preparato secondo modalità standardizzate, al fine di isolare ed identificare le frazioni del campione biologico stesso necessarie all'analisi richiesta.

3.2 Ri-analisi di campioni clinici

Generalmente, le linee guida inquadrano la ri-analisi dei campioni clinici soprattutto nell'ambito degli studi clinici. Diversa è l'esigenza nella pratica routinaria di un laboratorio di chimica clinica e/o farmacologia. In quest'ultimo caso, compatibilmente con la stabilità dell'analita, la ri-analisi di campioni clinici può essere di aiuto nel monitoraggio delle performance complessive del metodo in spettrometria di massa.

Infatti, i calibratori e i CQ utilizzati durante la validazione potrebbero non essere sufficienti per mimare i campioni clinici "reali" in quanto potrebbero esserci differenze che a loro volta potrebbero inficiare l'accuratezza e la precisione dell'analita (ad es. differenze nel legame con le proteine, conversione in metaboliti o presenza di co-medieazioni nei campioni reali).

È perciò necessario valutare l'accuratezza nei campioni reali ri-analizzando i campioni clinici in sedute di lavoro differenti, in giorni differenti, utilizzando curve di calibrazione preparate *ex-novo*¹¹.

In linea di massima, per gli studi clinici, si raccomanda di ri-analizzare il 10% dei campioni nel caso in cui il numero totale dei campioni sia meno di 1000 e il 5% del numero dei campioni oltre i 1000, mentre per le attività di routine clinica potrebbe essere consigliato ri-analizzare 1-5% del numero dei campioni settimanali/mensili, in funzione del volume di analisi, della stabilità e della relativa conservazione dei campioni e dell'analita. Tale attività potrebbe essere ulteriormente utile dopo il fallimento di una seduta di analisi per problemi tecnici e analitici.

È inoltre consigliabile analizzare campioni che coprano tutto il range delle concentrazioni attese e degli eventuali cut-off clinici decisionali, e la loro selezione dovrebbe evitare bias di selezione, rendendola più casuale possibile (es. sesso, età, giorno del prelievo, ecc.).

Generalmente per gli studi, la concentrazione ottenuta dalla prima analisi (originale) e quella ottenuta nella ri-analisi (ripetizione) dovrebbero essere all'interno del 20% della loro media (o

entro l'obiettivo riferito all'accuratezza dichiarata nel metodo) per almeno il 67% delle ripetizioni effettuate. Questo tipo di approccio è consigliabile anche per la ri-analisi in ambito della routine clinica, prendendo in carico e valutando gli eventuali campioni che fuoriescono dai criteri di accettabilità.

La differenza percentuale dei risultati è determinata dalla seguente equazione:

$$\text{Differenza \%} = (\text{Ripetizione} - \text{Originale}) * 100 / \text{Media}$$

3.3 Sensibilità diagnostica

La sensibilità di un test diagnostico è la proporzione di malati sulla popolazione totale presa in considerazione che vengono classificati in maniera corretta, ed è dunque l'indicatore che ci permette di rispondere alla domanda: qual è la probabilità che il test diagnostico dia risultato positivo ad un individuo malato?

Ovvero: $\text{SENSIBILITÀ} = \text{VERI POSITIVI} / \text{TOT MALATI}$

Questa definizione per i test diagnostici in LC-MS non è sempre applicabile, in quanto la tecnologia è, nella maggior parte delle volte, utilizzata per quantificare un analita in un definito range di concentrazioni. Per i farmaci la condizione clinica da verificare è l'assunzione del farmaco da parte del paziente, rispetto alla capacità del metodo di quantificarlo (in riferimento al LLOQ analitico).

Per i farmaci la definizione della Sensibilità Diagnostica dovrebbe essere la seguente:

la proporzione di pazienti che assumono il farmaco e con concentrazioni superiori al LLOQ che il test è in grado di identificare e quantificare, rispetto a tutta la popolazione arruolata che assume il farmaco in maniera corretta.

Ovvero: $\text{SENSIBILITÀ} = \text{VERI POSITIVI (correttamente classificati dal test, ovvero [farmaco] > LLOQ)} / \text{TOT MALATI (pazienti arruolati che assumono il farmaco in maniera corretta)}$.

Rimane sottinteso che la sensibilità analitica del test oggetto di valutazione clinica deve essere congrua rispetto alle concentrazioni terapeutiche attese. Quando vengono analizzati analiti per la potenziale tossicità, dobbiamo fare riferimento alle concentrazioni o cut-off clinici tossici identificati dalle linee guida e/o dalla letteratura scientifica. La condizione clinica da verificare è quindi lo stato di tossicità del paziente clinicamente riconosciuto, rispetto alla capacità del metodo di quantificarlo (in riferimento al LOD e/o al LLOQ).

Ovvero: $\text{SENSIBILITÀ} = \frac{\text{VERI POSITIVI (correttamente classificati dal test, ovvero [tossico] > \text{cut-off tossicologico})}}{\text{TOT MALATI (pazienti arruolati certamente intossicati)}}$

Quando vengono analizzati analiti endogeni, dobbiamo fare riferimento ai range di concentrazione di popolazione o ai cut-off clinici identificati dalle linee guida e/o dalla letteratura scientifica. La condizione clinica da verificare è in questo caso lo stato di malattia (condizione patologica) del paziente clinicamente riconosciuto, rispetto alla capacità del metodo di quantificarlo (in riferimento al range di riferimento e/o ai cut-off).

Ovvero: $\text{SENSIBILITÀ} = \frac{\text{VERI POSITIVI (correttamente classificati dal test, ovvero [sostanza endogena] > o < \text{del cut-off clinico o fuori dal range di riferimento})}}{\text{TOT MALATI (pazienti arruolati con una determinata condizione clinica patologica oggetto del test)}}$

3.4 Specificità diagnostica

La specificità di un test diagnostico è la proporzione di sani sulla popolazione totale presa in considerazione che vengono classificati in maniera corretta, ed è dunque l'indicatore che ci permette di rispondere alla domanda: qual è la probabilità che il test diagnostico dia risultato negativo ad un individuo sano?

Ovvero: $\text{SPECIFICITÀ} = \frac{\text{VERI NEGATIVI}}{\text{TOT SANI}}$

Anche in questo caso, questa definizione per i test diagnostici in LC-MS non è sempre di facile applicazione, in quanto la tecnologia è, come detto per la sensibilità diagnostica, per la maggior parte delle volte, utilizzata per quantificare un analita in un definito range di concentrazioni. Per i farmaci la condizione clinica da verificare è l'assenza del farmaco nei campioni di pazienti, rispetto alla capacità del metodo di non rilevarlo (in riferimento al Limite di rilevazione analitica LOD e/o al LLOQ).

Per i farmaci la definizione della Specificità Diagnostica dovrebbe essere la seguente:

è la proporzione di pazienti che non assumono il farmaco e che il test è in grado di identificare con concentrazioni < di LOD e/o LLOQ, rispetto a tutta la popolazione arruolata che non assume il farmaco.

Ovvero: $\text{SPECIFICITÀ} = \frac{\text{VERI NEGATIVI (correttamente classificati dal test, ovvero [farmaco] < \text{LOD e/o LLOQ})}}{\text{TOT SANI (pazienti arruolati che non assumono il farmaco)}}$

Rimane sottinteso che la specificità analitica del test oggetto di valutazione clinica deve essere congruo alle concentrazioni terapeutiche attese.

Quando vengono analizzati analiti per la potenziale tossicità, dobbiamo fare riferimento alle concentrazioni o cut-off clinici tossici identificati dalle linee guida e/o dalla letteratura scientifica. La condizione clinica da verificare è quindi lo stato assenza di tossicità del

paziente clinicamente riconosciuto, rispetto alla capacità del metodo di quantificarlo (in riferimento al cut-off tossicologico).

Ovvero: $\text{SPECIFICITÀ} = \frac{\text{VERI NEGATIVI (correttamente classificati dal test, ovvero [tossico] < cut-off tossicologico)}}{\text{TOT SANI (pazienti arruolati certamente non intossicati)}}$

Quando vengono analizzati analiti endogeni, dobbiamo fare riferimento ai range di concentrazione di popolazione o ai cut-off clinici identificati dalle linee guida e/o dalla letteratura scientifica. La condizione clinica da verificare è in questo caso lo stato di malattia del paziente clinicamente riconosciuto, rispetto alla capacità del metodo di quantificare l'analita endogeno (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

Ovvero: $\text{SPECIFICITÀ} = \frac{\text{VERI NEGATIVI (correttamente classificati dal test, ovvero [sostanza endogena] > o < del cut-off clinico o nel range di riferimento)}}{\text{TOT SANI (pazienti arruolati senza condizione clinica patologica oggetto del test)}}$

3.5 Valore predittivo di un test

Il valore predittivo di un test diagnostico esprime la probabilità che un paziente abbia, o non abbia, una certa malattia conoscendo il risultato del test. Il valore predittivo non è una caratteristica intrinseca del test, al contrario della sensibilità e della specificità, ma dipende dalla prevalenza della malattia. **La prevalenza** è il rapporto tra il numero di soggetti malati e il totale degli individui (malati + sani) presenti in una certa popolazione a un dato istante. In termini probabilistici indica quanto è diffusa la malattia nella popolazione (**FIGURA 1**).

Anche in questo caso, come per la specificità e sensibilità diagnostica, il valore predittivo dipende dalla condizione clinica in analisi e dalla tipologia di analita. Pertanto, dovremo affrontare la valutazione dei valori predittivi positivo e negativo del test con un approccio diversificato in funzione che si stia analizzando un farmaco (o altro xenobiotico), una sostanza per la sua tossicità o un analita endogeno. Ovvero cambia la definizione di soggetto malato o meno, ed il concetto di positività del test.

	D+	D-			D+ = malati; D- = sani
T+	Vp	Fp errore α	tot test+	Vpp: $Vp/Vp+Fp$	T+ = test positivo; T- = test negativo
T-	Fn errore β	Vn	tot test -	Vpn: $Vn/Vn+Fn$	Vp = Veri positivi; Fp = Falsi positivi; Vn = Veri negativi; Fn = Falsi negativi.
	tot malati	tot sani			Sn= sensibilità; Sp= specificità; Vpp= valore predittivo positivo; Vpn= valore predittivo negativo.
	Sn: $Vp/Vp+Fn$	Sp: $Vn/Fp+Vn$			

Figura 1. Impatto della prevalenza sulla performance clinica del test diagnostico. La figura illustra come la variazione della prevalenza di malattia modifichi il valore predittivo positivo (VPP) e negativo (VPN), pur mantenendo costanti sensibilità e specificità. Il VPP cresce all'aumentare della prevalenza, mentre il VPN diminuisce, riflettendo il diverso peso relativo dei veri positivi e dei veri negativi nella popolazione analizzata.

3.6 Valore predittivo positivo

Il **valore predittivo positivo (VPP)** misura la probabilità condizionale di essere malati essendo risultati **positivi** al test diagnostico; rappresenta, cioè, la proporzione di soggetti **positivi** al test che hanno realmente la malattia.

3.6.1 VPP per i Farmaci

Per i farmaci la condizione clinica da verificare è la presenza del farmaco nei campioni di pazienti, rispetto alla capacità del metodo di rilevarlo (in riferimento al LOD e/o al LLOQ).

Il **VPP** per i farmaci misura la probabilità di essere in terapia farmacologica (**Vp**) essendo **positivi** al test diagnostico rispetto a tutta la popolazione risultata positiva al test diagnostico (**Vp+Fp**).

$$VPP = Vp / (Vp + Fp)$$

Vp = pazienti che assumono il farmaco e risultano positivi alla presenza dello stesso al test (in riferimento al LOD e/o al LLOQ).

Fp = pazienti che non assumono il farmaco ma risultano positivi alla presenza dello stesso al test (in riferimento al LOD e/o al LLOQ).

3.6.2 VPP per le sostanze tossiche

Quando vengono analizzati analiti per la potenziale tossicità, dobbiamo fare riferimento alle concentrazioni o cut-off clinici tossici identificati dalle linee guida e/o dalla letteratura scientifica.

La condizione clinica da verificare è quindi lo stato di tossicità del paziente clinicamente riconosciuto, rispetto alla capacità del metodo di quantificarlo (in riferimento al LOD e/o al LLOQ).

Il **VPP**, per gli analiti con potenziale tossicità, misura la probabilità di essere intossicati (**Vp**) essendo **positivi** al test diagnostico (in riferimento al LOD e/o al LLOQ, ed anche alla concentrazione di cut-off tossicologico) rispetto a tutta la popolazione risultata positiva al test diagnostico (**Vp+Fp**).

$$\mathbf{VPP=Vp/(Vp+Fp)}$$

Vp = pazienti che manifestano lo stato di tossicità e risultano positivi alla presenza dell'analita tossico al test (in riferimento al LOD e/o al LLOQ, ed anche alla concentrazione di cut-off tossicologico).

Fp = pazienti che non manifestano lo stato di tossicità ma risultano positivi al test per la presenza dell'analita tossico (in riferimento al LOD e/o al LLOQ, ed anche alla concentrazione di cut-off tossicologico).

3.6.3 VPP per le sostanze endogene

Quando vengono analizzati analiti endogeni, dobbiamo fare riferimento ai range di concentrazione di popolazione o ai cut-off clinici identificati dalle linee guida e/o dalla letteratura scientifica.

La condizione clinica da verificare è in questo caso lo stato di malattia del paziente clinicamente riconosciuto, rispetto alla capacità del metodo di quantificare l'analita endogeno (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

Il **VPP** per gli analiti endogeni misura la probabilità di essere malato (**Vp**) essendo **positivi** al test diagnostico (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ) rispetto a tutta la popolazione risultata positiva al test diagnostico (**Vp+Fp**).

$$\mathbf{VPP=Vp/(Vp+Fp)}$$

Vp = pazienti che manifestano lo stato di malattia e risultano positivi al test per la presenza dell'analita (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

Fp = pazienti che non manifestano lo stato di malattia ma risultano positivi al test per la presenza dell'analita (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

3.7 Valore predittivo negativo

Il **valore predittivo negativo** (VPN) misura la probabilità condizionale di essere sani essendo risultati **negativi** al test diagnostico; rappresenta la proporzione di soggetti **negativi** al test che non hanno la malattia.

3.7.1 VPN per i Farmaci

Per i farmaci la condizione clinica da verificare è l'assenza del farmaco nei campioni di pazienti, rispetto alla capacità del metodo di rilevarlo (in riferimento al LOD e/o al LLOQ).

Il **VPN** per i farmaci misura la probabilità di non essere in terapia farmacologica (**Vn**) essendo **negativi** al test diagnostico rispetto a tutta la popolazione risultata negativa al test diagnostico (**Vn+Fn**).

$$\text{VPN} = \text{Vn} / (\text{Vn} + \text{Fn})$$

Vn = pazienti che non assumono il farmaco e risultano negativi alla presenza dello stesso al test (in riferimento al LOD e/o al LLOQ).

Fn = pazienti che assumono il farmaco ma risultano negativi alla presenza dello stesso al test (in riferimento al LOD e/o al LLOQ).

3.7.2 VPN per le sostanze tossiche

Quando vengono analizzati analiti per la potenziale tossicità, dobbiamo fare riferimento alle concentrazioni o cut-off clinici tossici identificati dalle linee guida e/o dalla letteratura scientifica.

La condizione clinica da verificare è quindi l'assenza di tossicità del paziente clinicamente riconosciuta, rispetto alla capacità del metodo di quantificare l'analita tossico (in riferimento al LOD e/o al LLOQ).

Il **VPN**, per gli analiti con potenziale tossicità, misura la probabilità di non essere intossicati (**Vn**) essendo **negativi** al test diagnostico (in riferimento al LOD e/o al LLOQ, ed anche alla concentrazione di cut-off tossicologico) rispetto a tutta la popolazione risultata negativa al test diagnostico (**Vn+Fn**).

$$\text{VPN} = \text{Vn} / (\text{Vn} + \text{Fn})$$

Vn = pazienti che non manifestano lo stato di tossicità e risultano negativi alla presenza dell'analita tossico al test (in riferimento al LOD e/o al LLOQ, ed anche alla concentrazione di cut-off tossicologico).

F_n = pazienti che manifestano lo stato di tossicità ma risultano negativi al test per la presenza dell'analita tossico (in riferimento al e/o al LLOQ, ed anche alla concentrazione di cut-off tossicologico).

3.7.3 VPN per le sostanze endogene

Quando vengono analizzati analiti endogeni, dobbiamo fare riferimento ai range di concentrazione di popolazione o ai cut-off clinici identificati dalle linee guida e/o dalla letteratura scientifica.

La condizione clinica da verificare è in questo caso l'assenza di malattia del paziente clinicamente riconosciuta, rispetto alla capacità del metodo di quantificare l'analita endogeno (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

Il **VPN** per gli analiti endogeni misura la probabilità di non essere malato (**V_n**) essendo **negativo** al test diagnostico (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ) rispetto a tutta la popolazione risultata negativa al test diagnostico (**V_n+F_n**).

$$\text{VPN} = \text{V}_n / (\text{V}_n + \text{F}_n)$$

V_n = pazienti che non manifestano lo stato di malattia e risultano negativi al test per la presenza dell'analita (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

F_n = pazienti che manifestano lo stato di malattia ma risultano negativi al test per la presenza dell'analita (in funzione del range di riferimento e/o ai cut-off clinici, oltre che in riferimento al LOD e/o al LLOQ).

3.8 Rapporto di probabilità

L'OR ha lo scopo di determinare se esiste una associazione tra l'esposizione a certi fattori e l'insorgenza di una malattia⁴⁵ o la sua manifestazione clinica. L'esposizione a certi fattori è di fatto, ciò che misura il nostro test diagnostico (tra cui i test in spettrometria di massa) in funzione dei valori di riferimento di normalità, efficacia e tossicità clinica.

L'OR è il rapporto di probabilità che si manifesti o non si manifesti la condizione clinica e/o malattia nei casi in cui il valore misurato sia al di fuori dei valori di riferimento (di normalità, efficacia e tossicità clinica) rispetto alla probabilità che si manifesti la condizione clinica e/o malattia nei casi in cui il valore misurato sia all'interno dei valori di riferimento (di normalità, efficacia e tossicità clinica).

3.9 Valori attesi in popolazioni sane ed affette da patologie

Un primo distinguo in riferimento ai valori attesi di un test dovrebbe essere effettuato ragionando su un test diagnostico ben conosciuto, ovvero consolidato nei suoi range di normalità, efficacia e/o tossicità, rispetto ad un test di nuova introduzione dove i valori dei range di normalità, efficacia e/o tossicità non siano sufficientemente descritti e/o consolidati dalla letteratura scientifica. Nello specifico, per i test consolidati è possibile fare riferimento ai valori già individuati e validati clinicamente. Tale utilizzo è vincolato però ad una valutazione della tecnica analitica e della popolazione di riferimento (esempio: etnia, sesso, età, ecc.) utilizzate e citate nei documenti a cui si vuol fare riferimento.

Nel caso in cui il test non sia ancora sufficientemente descritto rispetto ai suoi valori di riferimento, gli sviluppatori del test diagnostico devono effettuare una valutazione clinica sulla propria popolazione target e definire quelli che sono i valori attesi nei pazienti “sani” e “malati”.

3.10 Numerosità campionaria

Nell’ambito delle validazioni cliniche la numerosità dei pazienti arruolati negli studi clinici di validazione dei metodi analitici per diagnostici può essere variegata.

Sarebbe sempre utile confrontarsi con l’ente accreditatore e/o organismo notificante di riferimento (ad esempio ISS, ecc.) per concordare i numeri minimi relativamente ai pazienti “sani” e “malati” da includere nello studio di validazione.

È consuetudine arruolare un numero minimo di pazienti pari a 100 “sani” e 100 “malati”; in particolari ambiti clinici o in presenza di indicazioni precise (linee guida, norme europee, ecc.), il numero può e deve cambiare in base al disegno dello studio ed alla prevalenza dello stato patologico oggetto del test, con un’attenta valutazione della potenza statistica.

4. Validazione parziale, kit commerciali e cross-validazione

I Kit IVD e IVDR devono essere progettati, validati e prodotti seguendo la normativa europea di riferimento del maggio 2017, ove, tra le varie novità, è stata introdotto l’obbligo di effettuare una valutazione delle performance cliniche dei Kit diagnostici². La normativa prefigura 3 situazioni in cui è necessario procedere almeno ad una validazione parziale e/o verifica delle performance:

1. introduzione di un nuovo Kit (primo acquisto) IVD/IVDR. Le verifiche da effettuare, relativamente a ciascun punto sopra riportato, sono legate alle normative in vigore allo stato attuale;
2. modifica e/o adattamento di un Kit IVD/IVDR (es. nuovo strumento, cambio software, ecc...);
3. modifica di un metodo LDT già validato internamente al laboratorio.

Specificatamente per il punto 1, l'operatore finale dovrebbe essere affiancato da un "operatore esperto", sino a quando l'operatore non raggiunge i livelli di performance qualitativi dichiarati dal produttore² del Kit CE-IVD/IVDR.

È importante sottolineare che una validazione parziale del metodo andrebbe effettuata ogni volta che si adotti un nuovo kit commerciale presso il proprio laboratorio.

Tali kit (in particolare quelli IVDR) potrebbero essere validati dal produttore su strumentazioni differenti che potrebbero presentare una performance variabile in termini di range del test, linearità, riproducibilità etc.. In questo caso decade la marchiatura CE-IVDR, ed è indispensabile effettuare almeno una validazione parziale o totale (in funzione delle differenze tra la strumentazione disponibile e quella utilizzata per la validazione IVD/IVDR da parte del produttore) e verificare le performance, come richiesto dall'ISO 15189³.

In riferimento al punto 2, modifiche o cambiamenti a metodi bioanalitici che rientrano in questa categoria possono includere, per esempio:

- trasferimento di metodi bioanalitici tra laboratori o operatori;
- modifiche della metodologia analitica (es. cambio di componenti del rivelatore o di altro componente strumentale e/o configurazione, inclusa la piattaforma software);
- cambio di anticoagulante (es. nel prelievo venoso da eparina ad EDTA);
- cambio di matrice e/o campione biologico (es. passaggio da urine a plasma);
- modifiche nelle procedure di processamento del campione o delle modalità di campionamento (es. impiego di dispositivi per matrici essiccate);
- modifiche nel volume di campione utilizzato (es. riduzione del volume in pediatria).

Lo scopo della validazione parziale è quello di dimostrare che le modifiche applicate al metodo non inficiano le performance del metodo stesso.

Le validazioni parziali possono variare dalla sola ri-valutazione di accuratezza e precisione intra-assay/inter-assay, ad una validazione quasi completa. La decisione su quali caratteristiche del metodo richiedano una validazione aggiuntiva va presa in base a considerazioni logiche su quelli che, presumibilmente, sono i parametri che sono stati inficiati dalle modifiche attuate al metodo. Per esempio, una modifica alla fase stazionaria della

colonna analitica o alla composizione della fase mobile potrebbero modificare la linearità e/o la selettività del metodo e/o l'effetto matrice, che andrebbero quindi opportunamente rivalutate. Per quanto concerne il punto 3 si ribadisce che per validare un test LDT un laboratorio deve essere obbligatoriamente accreditato UNI EN ISO 15189.

La tipologia di ri-validazione del metodo LDT deve essere effettuata valutando, come per i punti 1 e 2, le differenze tecniche e/o metodologiche introdotte rispetto al metodo originario validato, e considerando le relative ricadute sulle prestazioni del metodo stesso.

Pertanto, anche in questo caso, la decisione su quali caratteristiche del metodo richiedano una validazione aggiuntiva va presa in base a considerazioni logiche su quelli che, presumibilmente, sono i parametri che sono stati inficiati dalle modifiche attuate al metodo.

4.1 Cross validazione

Per meglio comprendere gli argomenti descritti in questa sezione, fare riferimento in particolare alle linee guida EMA. La cross-validazione è una comparazione dei parametri di validazione tra più metodi bioanalitici, o tra diversi laboratori, che deve essere effettuata nei seguenti casi:

- quando due o più metodi bioanalitici vengono utilizzati per generare dati all'interno dello stesso studio e/o attività di laboratorio
- quando lo stesso metodo bioanalitico viene utilizzato da due laboratori differenti per generare dati all'interno dello stesso studio e/o attività di laboratorio
- quando vengano effettuate modifiche “minori” ad un metodo esistente (per es. cambio di componenti del rivelatore o di altro componente strumentale e/o configurazione, inclusa la piattaforma software; cambio di anticoagulante; etc.)

Per effettuare la cross-validazione bisogna analizzare uno stesso set/lotto di CQ (ad alta, media e bassa concentrazione, in triplicato) o di campioni clinici (in numero superiore a 30, se disponibili) utilizzando entrambi i metodi da confrontare⁴⁶.

Il confronto tra i metodi deve essere valutato attraverso congrui metodi statistici, come per esempio il test di Bland Altman e/o la regressione Passing-Bablok⁴⁷.

Per i CQ l'accuratezza media ottenuta per entrambi i metodi deve essere entro il $\pm 15\%$. Per i campioni clinici la differenza tra i due valori ottenuti deve essere all'interno del 20% della media per almeno il 67% dei test effettuati²⁹

Questi valori devono essere inquadrati nell'ambito del test diagnostico, del suo significato clinico e linee guide specifiche del singolo test, quindi i criteri di accettabilità potrebbero variare (generalmente più stringenti).

5. Discussione e Conclusioni

Le analisi in LC-MS sono diventate uno strumento primario nei test di medicina di laboratorio. La comprensione della terminologia utilizzata, attraverso i regolamenti e le direttive provenienti da diversi paesi, è diventata una sfida per gli scienziati e gli operatori coinvolti nei laboratori. Fino ad oggi, infatti, sono state spesso utilizzate nella validazione di metodi bioanalitici diverse terminologie non univoche e chiare, e sono stati pubblicati in letteratura dei lavori che utilizzano protocolli per la validazione LC-MS con differenti livelli di accettazione. Questo causa una perdita di armonizzazione dei risultati nonostante la tecnologia LC-MS sia tra le più sensibili e specifiche a disposizione.

Attualmente, non sono ancora state pubblicate in Italia delle linee guida univoche e condivise che siano esclusivamente dedicate alla validazione di metodi bioanalitici clinici basati sulla spettrometria di massa. Inoltre, chi vuole sviluppare e validare un metodo quantitativo in LC-MS o LC-MS/MS deve fare riferimento alle più importanti linee guida internazionali, che però, in alcuni casi non sono abbastanza specifiche per la LC-MS. Mancano ad esempio indicazioni specifiche per gli analiti endogeni^{4,12,28,29} oppure, in alcune linee guida, non sempre viene considerato che i campioni dei pazienti possono presentare delle “alterazioni”, come nel caso di prelievi lipemici, iperbilirubinemici o emolizzati. Queste possibili alterazioni devono essere assolutamente considerate nel processo di validazione di un LDT sia di un metodo commerciale, ed i risultati devono essere pubblicati o opportunamente riportati nel report di validazione e/o nel “data sheet” del kit diagnostico IVDR². Nell’ambito del gruppo di studio “La spettrometria di massa: applicazioni e innovazioni diagnostiche” della SIBIOC, abbiamo cercato con il presente lavoro di colmare questa lacuna illustrando brevemente gli step metodologici da eseguire durante il processo di valutazione delle performance cliniche e analitiche.

Nel caso dei kit (CE-IVDR o meno) usati per le analisi LC-MS in biochimica clinica, questi non sono sempre validati su tutti gli strumenti attualmente sul mercato (e sarebbe impossibile considerando le innumerevoli configurazioni e combinazioni strumentali possibili oggi e nel prossimo futuro). Questi kit (inclusi quelli con marchio CE-IVDR) necessitano però almeno di una validazione parziale sugli strumenti disponibili in ciascun laboratorio, per rivalutare le performance e ridefinirne le potenziali limitazioni analitiche.

In conclusione, le nuove regole ed i protocolli riportati in questo lavoro non si sostituiscono in alcun modo alle attuali linee guida, ma hanno come finalità quella di fornire una guida

concreta a chi si approccia allo sviluppo e alla valutazione delle prestazioni di metodi LC-MS, sia essa totale o parziale, raccomandando alcuni ulteriori passaggi da effettuare durante la validazione, con l'obiettivo finale di fornire al clinico e al paziente i migliori risultati analitici e clinici possibili.

Acknowledgment / Ringraziamenti

Si ringraziano la Dott.ssa Cinzia Benagli (Ente Ospedaliero Cantonale - Bellinzona) e la Dott.ssa Debora Pensi (Università di Torino) per la loro lettura critica del manoscritto.

Conflict of interest / Conflitti di Interesse

- AD è socio di CoQua Lab srl
- VI è socia di MaTOX srl

BIBLIOGRAFIA

1. Antonio D'Avolio, Marco Cantù, Jacopo Gervasoni, Carlo Artusi, Mariela Marinova, Antonello Nonnato, Giuliana Cangemi, Silvia Persichilli. Validazione dei metodi quantitativi bioanalitici in spettrometria di massa: regole e protocolli sperimentali. *Biochimica Clinica*, 2018, *Biochimica Clinica*, 2018, vol. 42, n. 1. DOI: 10.19186/BC_2018.004].
2. UE. IVD - Regolamento (Ue) 2017/746 Del Parlamento Europeo e Del Consiglio Del 5 Aprile 2017 <http://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32017R0746&from=EN> (26/05/2017).
3. International Organization for Standardization. (2022). Medical laboratories — Requirements for quality and competence (ISO 15189:2022).
4. FDA. Guidance for Industry: Bioanalytical Method Validation (Rev. 1). <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM368107.pdf> (09/09/2014).
5. [Baietto L, Simiele M, D'Avolio A. How effective is the use of DBS and DPS as tools to encourage widespread therapeutic drug monitoring? *Bioanalysis* 2014; 6: 425-7.].
6. [D'Avolio A, Simiele M, Siccardi M et al. HPLC-MS method for the quantification of nine anti-HIV drugs from dry plasma spot on glass filter and their long term stability in different conditions. *J Pharm Biomed Anal* 2010; 52: 774-80.].
7. [Zimmer JS, Christianson CD, Johnson CJ et al. Recent advances in the bioanalytical applications of dried matrix spotting for the analysis of drugs and their metabolites. *Bioanalysis* 2013; 5: 2581-8.].
8. McCloskey LJ, Yoo JH, Stickle DF. Interpatient distributions of bloodspot area per fixed volume of application: comparison between filter paper and non-cellulose dried matrix spotting cards. *Clin Chim Acta* 2014; 437: 187-90].
9. [Cangemi G, Barco S, Castagnola E et al. Development and validation of UHPLC-MS/MS methods for the quantification of colistin in plasma and dried plasma spots. *J Pharm Biomed Anal* 2016; 129: 551-7].
10. Mensitieri, F.; Coglianese, A.; Giudice, V.; Charlier, B.; De Rosa, F.; Filippelli, A.; Dal Piaz, F.; Izzo, V. Effects of selected preanalytical variables on Dried Blood Spot (DBS)

and Volumetric Adsorptive Microsampling (VAMS) based bioanalytical methods for the determination of four β -lactam antibiotics. *BIOCHIMICA CLINICA*. Vol. 46. Pag.141-153. ISSN:0393-0564. (DOI): 10.19186/BC_2022.013].

11. EMA. Guideline on bioanalytical method validation. (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf) 2011.

12. SWGTOX. Standard practices for method validation in forensic toxicology. (<http://www.swgtox.org/documents/Validation3.pdf>) 2013.

13. WADA Technical Document – TD2021IDCR Version Number: 1.0. Minimum criteria for chromatographic-mass spectrometric confirmation of the identity of analytes for doping control purposes.

14. Revisione n. 5 del 29 maggio 2017 a cura della Commissione Qualità dell'Associazione Scientifica "Gruppo Tossicologi Forensi Italiani" (GTFI). Linee guida per le strutture dotate di laboratori per gli accertamenti di sostanze d'abuso con finalità tossicologico-forensi e medico-legali su campioni biologici prelevati da vivente.

15. D'Avolio A, Sciandra M, Siccardi M et al. A new assay based on solid-phase extraction procedure with LC-MS to measure plasmatic concentrations of tenofovir and emtricitabine in HIV infected patients. *J Chromatogr Sci* 2008; 46: 524-8].

16. D'Avolio A, Siccardi M, Sciandra M et al. HPLC-MS method for the simultaneous quantification of the new HIV protease inhibitor darunavir, and 11 other antiretroviral agents in plasma of HIV-infected patients. *J Chromatogr B Analyt Technol Biomed Life Sci* 2007; 859: 234-40].

17. [Vogeser M, Seger C. Pitfalls associated with the use of liquid chromatography-tandem mass spectrometry in the clinical laboratory. *Clin Chem* 2010; 56: 1234-44.

18. [D'Avolio A, De Nicolo A, Agnesod D et al. A UPLC-MS/MS method for the simultaneous plasma quantification of all isomeric forms of the new anti-HCV protease].

19. ICH guideline M10 on bioanalytical method validation, 13 March 2019 EMA/CHMP/ICH/172948/201.

20. [International vocabulary of metrology – Basic and general concepts and associated terms (VIM) 3rd edition JCGM 200:2012].

21. B. Magnusson and U. Örnemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0].

22. B. Magnusson and U. Örnemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0.

23. ISO 11843-2:2007 Capability of detection – Part 2: Methodology in the linear calibration case, ISO Geneva.

24. Huidong Gu, Guowen Liu, Jian Wang, Anne-Françoise Aubry, and Mark E. Arnold. Selecting the Correct Weighting Factors for Linear and Quadratic Calibration Curves with Least-Squares Regression Algorithm in Bioanalytical LC-MS/MS Assays and Impacts of Using Incorrect Weighting Factors on Curve Stability, Data Quality, and Assay Performance. *Anal. Chem.* 2014, 86, 18, 8959–8966].

25. Seyed Mojtaba Moosavi and Susan Ghassabian. Linearity of Calibration curves for analytical methods: a review of criteria for assessment of method reliability. IntechOpen, book chapter, doi: 10.5772/intechopen.72932.

26. International Organization for Standardization. (2023). Accuracy (trueness and precision) of measurement methods and results (ISO 5725-1:2023 [Available from: <https://www.iso.org/standard/69418.html>].

27. [Vidali, M., Tronchin, M., & Dittadi, R. (2016). Protocollo per la comparazione di due metodi analitici di laboratorio. *Biochim. Clin*, 40, 129-142.
28. CLSI. C62-A Liquid Chromatography-Mass Spectrometry Methods; Approved Guideline. https://clsi.org/media/1346/c62a_sample.pdf (01/01/2014).
29. EMA. Guideline on bioanalytical method validation. (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf) 2011].
30. De Nicolò, A., Cantù, M., & D'Avolio, A. (2017). Matrix effect management in liquid chromatography mass spectrometry: the internal standard normalized matrix effect. *Bioanalysis*, 9 14, 1093-1105.
31. Scientific Working Group for Forensic Toxicology (SWGTOX), Clinical & Laboratory Standards Institute (CLSI) C62-A.
32. Iyer SS, Zhang ZP, Kellogg GE et al. Evaluation of deuterium isotope effects in normal-phase LC-MS-MS separations using a molecular modeling approach. *J Chromatogr Sci* 2004; 42: 383-7].
33. Jemal M, Schuster A, Whigan DB. Liquid chromatography/tandem mass spectrometry methods for quantitation of mevalonic acid in human plasma and urine: method validation, demonstration of using a surrogate analyte, and demonstration of unacceptable matrix effect in spite of use of a stable isotope analog internal standard. *Rapid Commun Mass Spectrom* 2003; 17: 1723-34].
34. Wang S, Cyronak M, Yang E. Does a stable isotopically labeled internal standard always correct analyte response? A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *J Pharm Biomed Anal* 2007; 43: 701-7.
35. Bonfiglio R, King RC, Olah TV, Merkle K. The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds. *Rapid Commun Mass Spectrom*. 1999 Jun;13(12):1175-1185. doi: 10.1002/(SICI)1097-0231(19990630)13:12<1175::AID-RCM639>3.0.CO;2-0. PMID: 10407294.
36. Matuszewski BK, Constanzer ML, Chavez-Eng CM. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Anal Chem*. 2003 Jul 1;75(13):3019-30. doi: 10.1021/ac020361s. PMID: 12964746].
37. [Matuszewski BK. Standard line slopes as a measure of a relative matrix effect in quantitative HPLC-MS bioanalysis. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2006 Jan 18;830(2):293-300. doi: 10.1016/j.jchromb.2005.11.009. Epub 2005 Nov 23. PMID: 16310419.
38. CLSI. C62-A Liquid Chromatography-Mass Spectrometry Methods; Approved Guideline 2014 [Available from: https://clsi.org/media/1346/c62a_sample.pdf.
39. UE. IVD - Regolamento (Ue) 2017/746 Del Parlamento Europeo e Del Consiglio Del 5 Aprile 2017 [Available from: <http://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32017R0746&from=EN>.
40. IVDR 2017/746, Allegato XIII.
41. MedTech Europe. Clinical Evidence Requirements under the IVD Regulation. 3rd ed. Brussels: MedTech Europe; 2023] [Leefflang MMG, Reitsma JB, Bruns DE, Stayman MR, Linnet K. How to: evaluate a diagnostic test. *Clin Microbiol Infect*. 2019;25(5):520-528.
42. Bogduk N, Smith E, Jones M. On understanding the validity of diagnostic tests. *Diagn Pathways*. 2022;7(1):12-20.
43. Clinical and Laboratory Standards Institute. Framework for Developing Evidence of Clinical Validity of Medical Laboratory Test Methods. CLSI guideline EP49. Wayne, PA: CLSI; 2025.

44. Mantra Systems. Scientific Validity Reports – Foundation of the IVDR Performance Evaluation. Mantra Systems white paper; 2023.
45. RISCHIO RELATIVO E ODDS RATIO: COME MISURARE IL RISCHIO [Available from: <https://www.med4.care/rischio-relativo-odds-ratio-come-misurare-il-rischio/>].
46. ICH M10 on bioanalytical method validation - Scientific guideline [Available from: <https://www.ema.europa.eu/en/ich-m10-bioanalytical-method-validation-scientific-guideline>].
47. Vidali, Tronchin, Dittadi (2016). Protocollo per la validazione e comparabilità dei metodi. Biochim Clin 40:129–142.