

Mise en œuvre du profil d'exactitude

Max Feinberg¹

Résumé : Cet article a pour objet d'expliquer comment la procédure de construction du profil d'exactitude peut être utilisée pour valider une méthode d'analyse quantitative. Le profil d'exactitude est applicable aussi bien à la validation interne dans un laboratoire d'essais qu'à l'interprétation de données collectées dans le cadre d'une étude inter-laboratoires conçue selon les recommandations de la norme ISO 5725-2. Elle peut aussi s'appliquer aux essais basés sur des énumérations, comme les méthodes microbiologiques, au prix d'un certain nombre de modifications qui ne seront pas présentées ici. L'approche proposée doit permettre une validation interne des méthodes dans les domaines suivants :

- les méthodes non normalisées ;
- les méthodes conçues/développées par le laboratoire ;
- les méthodes normalisées employées en dehors de leur domaine d'application prévu ;
- les amplifications ou modifications de méthodes normalisées ;
- et aussi une méthode alternative par rapport à une méthode de référence.

Mots-clés : Validation, profil d'exactitude, justesse, fidélité, incertitude, méthodes d'analyse quantitatives

Introduction

Le profil d'exactitude ne s'applique qu'aux méthodes complètement développées et mises au point. En particulier, la sélectivité/spécificité doit avoir été étudiée ainsi que le domaine d'application de la méthode à valider, en termes de types de matrice et de niveaux de concentrations. Le domaine d'application du profil d'exactitude s'inspire directement de la norme ISO 17025:2005 qui propose que

« Le laboratoire doit valider les méthodes non normalisées, les méthodes conçues/développées par le laboratoire, les méthodes normalisées employées en dehors de leur domaine d'application prévu, ainsi que les amplifications ou les modifications de méthodes normalisées, afin de confirmer que les méthodes sont aptes à l'emploi prévu. La validation doit être aussi étendue que l'impose la réponse aux besoins pour l'application ou le domaine d'application donné ».

Bien sûr, de nombreux développements statistiques ne sont pas présentés mais sont disponibles à la référence (Feinberg M., 2009).

¹ UR1204 Met@risk - Méthodologie d'analyse de risque alimentaire- INRA – F-75231 Paris cedex 05

☎ 01 44 08 16 52 ✉ max.feinberg@paris.inra.fr

1. Profil d'exactitude en 10 étapes

1.1 Définir la quantité mesurée

À partir du mode opératoire de la méthode, on définit la quantité mesurée en précisant les formules d'expression du résultat final et la procédure pour établir ce résultat. En particulier, il faut faire la distinction entre deux types de méthodes.

- Les méthodes « indirectes ou rationnelles » qui exigent un étalonnage préalable pour calculer la concentration des échantillons inconnus. Pour ces méthodes, on procède en deux étapes : d'abord construire la courbe d'étalonnage en utilisant le même principe physico-chimique que celui qui sert pour les échantillons ; ensuite faire des mesures sur les échantillons inconnus et calculer leurs concentrations à l'aide d'un modèle d'étalonnage ;
- Les méthodes « directes ou empiriques » dont l'analyte est défini par la méthode elle-même. On mesure alors la concentration des échantillons inconnus à l'aide d'une autre méthode de mesure, comme une pesée ou une titrimétrie. Parmi ces méthodes, dites critères ou de type I dans le Codex Alimentarius, on trouve l'azote organique total de Kjeldahl, l'extrait sec, les lipides totaux, etc.

Il est fondamental que la quantité mesurée lors de la validation soit bien la même que celle qui sera mesurée en routine. En particulier, si le mode opératoire prévoit d'exprimer le résultat final à partir d'un seul mesurage, il ne faut pas exprimer le résultat d'un essai de validation comme la moyenne de plusieurs répétitions.

1.2 Préciser des objectifs de la validation

Choisir le domaine de validation

Définir le domaine de validation de la méthode, sous la forme d'une gamme de concentrations absolues ou relatives, par exemple, entre 1 g et 60 g ou entre 1 et 500 µg/l. Pratiquement, c'est à travers les choix effectués pour construire les plans d'expériences, en termes de gamme de concentrations pour un ou plusieurs types de matrice, qu'on démontre dans quel domaine la méthode est effectivement valide, et si elle est capable de fournir des résultats acceptables. Le domaine où cette démonstration est faite est appelé **domaine de validité** et peut être plus petit que le **domaine d'application** défini *a priori*.

Le choix du domaine d'application peut correspondre à une obligation légale. Par exemple, dans le domaine pharmaceutique, il faut prendre en compte les domaines minimums suivants :

- pour les substances pures ou les spécialités pharmaceutiques, le domaine doit au moins couvrir entre 80 et 120 % de leurs valeur nominales ;
- pour une vérification de concentration continue, le domaine doit s'étendre entre 70 et 130 % de la concentration attendue.

Si, lors de l'utilisation en routine de la méthode on rencontre des échantillons dont la concentration n'est pas incluse dans le domaine de validité, il est interdit d'extrapoler. Il faut alors procéder à une étude complémentaire pour étendre ce domaine ou procéder à une dilution dans la mesure où on a démontré qu'elle n'entraîne pas d'effet. Une mauvaise définition du domaine d'application peut avoir des conséquences importantes sur la validité, surtout en termes de spécificité, d'interférences ou de réactions croisées

Les limites d'acceptabilité sont notées $\pm\lambda$, lorsqu'elles sont exprimées de façon absolue dans la même unité que le mesurande, ou $(1 \pm \lambda) \times 100$ de façon relative. Dans la mesure du possible, on va définir les limites d'acceptabilité, en se référant à un document, une pratique de la profession ou une exigence scientifique. La liste des approches possibles qui sont présentées ci-dessous n'est pas exhaustive et n'a qu'une valeur d'exemple. Cependant, il n'entre pas dans le champ d'application de ce texte de fixer des valeurs pour λ .

- **Législation.** Plusieurs textes réglementaires fournissent des valeurs qui peuvent servir de base pour fixer λ . Par exemple, la Décision de la commission 2002/657 indique dans son tableau 2 la justesse minimale des méthodes quantitatives pour les médicaments et résidus de médicaments vétérinaires dans les produits d'origine animale (Décision de la commission 2002/657 (CE) du 12 août 2002). D'autres documents publiés par la Commission du Codex Alimentarius définissent une « approche critère » pour différents analytes et matrices (FAO, 1997 ; CX/MAS 09/30/7) qui devrait clarifier la situation au niveau international.
- **Pratiques de la profession.** Certaines professions ont défini des limites d'acceptabilité sous la forme d'écart maximum acceptable (EMA), comme cela se pratique pour le contrôle de qualité ou la métrologie légale. C'est une approche classique en biologie médicale pour les paramètres du sang ou dans l'industrie pharmaceutique pour le contrôle des spécialités pharmaceutiques.
- **Modèle empirique.** Il existe des modèles mathématiques qui relient des performances analytiques à la concentration. Le plus connu est celui développé par Horwitz vers 1980 ; il est repris dans de nombreux textes et permet de prédire la valeur attendue de l'écart-type de reproductibilité s_R en fonction d'un niveau de concentration X (en particulier : décision de la commission 2002/657 (CE) du 12 août 2002 ; CX/MAS 09/30/7). Pour une concentration inférieure à 10^{-7} kg.kg⁻¹, il s'exprime sous la forme,

$$s_R = 0,02 \cdot X^{0,85} \quad (1)$$

- **À partir de performances connues.** Par ailleurs, si on considère que la reproductibilité limite représente un écart maximal acceptable (EMA) entre 2 mesures faites en condition de reproductibilité, pour un niveau de confiance de 95%, celui-ci vaut $|EMA| \leq 2,83s_R$ et il vaudra $|EMA| \leq 4,24s_R$ pour un niveau de 99%. En combinant cette définition à l'équation (1) on peut définir la limite d'acceptabilité à partir de l'EMA et du modèle Horwitz :

$$|EMA| \leq 0,057 \times X^{0,85} \text{ ou } |EMA| \leq 0,085 \times X^{0,85}$$

ou exprimé en %.

$$EMA\% \leq \frac{0,057 \times X^{0,85}}{X} \times 100 \text{ ou } EMA\% \leq \frac{0,085 \times X^{0,85}}{X} \times 100$$

Dans le cas où on dispose d'une estimation expérimentale de l'écart-type de reproductibilité – par exemple si la méthode a été soumise à une étude inter-laboratoires – on peut affiner les limites d'acceptabilité en utilisant la valeur observée de s_R .

- **À partir de la limite de quantification.** Lorsque l'objectif de la méthode est l'application d'une norme sanitaire, il est possible de déduire des limites d'acceptabilité, à partir de la limite de quantification calculée avec le profil d'exactitude (Feinberg M., 2010c). La décision de validation se fera alors en fonction du risque encouru, obtenu *a posteriori*.

Il est souvent plus commode que les limites d'acceptabilité soient exprimées sous la forme d'un pourcentage mais ce n'est pas une obligation. De même, pour simplifier, on choisit une valeur constante pour tout le domaine d'application. Mais, si le domaine d'application est large, on peut utiliser des valeurs variant en fonction du niveau de concentration.

Dans le cas où les valeurs de référence sont connues avec une incertitude élevée, le choix des limites d'acceptabilité devra être compatible avec cette incertitude.

1.3 Sélectionner les échantillons de validation

Choisir la ou les matrices

Les essais qui servent à calculer les éléments du profil d'exactitude doivent être réalisés sur des échantillons clairement identifiés, appelés **échantillons de validation**. Ce sont des matériaux qui doivent être le plus représentatif possible du domaine d'application de la méthode. Ils peuvent ainsi inclure différentes matrices pour les méthodes multi-matrices. Mais surtout on doit être capable de leur attribuer une **valeur de référence** qui représente la teneur attendue. Penser à choisir des matériaux stables et homogènes en quantité suffisante pour réaliser l'ensemble des essais définis par le plan d'expérience de validation.

Établir les valeurs de référence des échantillons de validation

Pour estimer la justesse de la méthode, il faut disposer d'échantillons de validation dont la concentration est connue le plus exactement possible avec une incertitude connue. Cette concentration correspond à la **valeur de référence** assignée à l'échantillon de validation et doit être fixée indépendamment de la méthode à valider, on la note X . Il existe plusieurs approches possibles pour établir la valeur de référence assignée à un échantillon de validation, parmi lesquelles :

- utiliser des matériaux de référence certifiés (MRC), externes (MRE) ou internes (MRI) ; la traçabilité de ces matériaux décroît en fonction de leur nature ;
- réaliser des ajouts dosés à partir d'une molécule étalon de pureté connue ;
- préparer des échantillons dopés ou synthétiques avec la matrice choisie ;
- utiliser une méthode de référence, appliquée en interne ou en externe par un laboratoire pair. Si la valeur de référence est estimée par des répétitions de la méthode de référence, il conviendra de sélectionner la valeur la plus représentative, comme la moyenne ou la médiane. Cette approche doit être préférentiellement appliquée dans le cas de la validation d'une méthode alternative par rapport à une méthode de référence ;
- utiliser de la dilution isotopique lorsqu'on dispose de l'analyte sous une forme marquée.

Pour la préparation d'échantillons de validation par la méthode des ajouts dosés de niveaux variés, veiller à réduire le plus possible la dépendance entre les différentes préparations. Les ajouts dosés peuvent être effectués sous une forme liquide (par ajout d'une solution étalon) ou solide (par pesée). Si on procède par des dilutions successives (1/2, 1/4, 1/8...) à partir d'une même solution mère, on crée une corrélation entre les mesures et on augmente les conséquences d'une erreur de justesse ; il est donc déconseillé de procéder de cette façon. Cette remarque s'applique aussi à la préparation des solutions étalons.

Le choix raisonné des valeurs de référence revêt un aspect fondamental pour l'application du profil d'exactitude car il conditionne la justesse de la méthode. En fonction de la technique analytique choisie, toute approche visant à cet objectif est acceptable, aussi longtemps qu'une justification scientifique fondée est apportée.

Lorsqu'on est dans l'incapacité d'assigner une valeur de référence à un échantillon de validation, il est possible de la remplacer par la moyenne ou la médiane des concentrations obtenues directement sur cet échantillon. La valeur de référence ainsi obtenue n'est plus indépendante de la méthode en cours de validation et interdit l'évaluation du biais de justesse. Cette approche très restrictive doit être clairement justifiée dans le dossier de validation.

1.4 Planifier les essais de validation

Organiser les essais de validation

Le plan d'expérience de validation sert à estimer, dans les conditions où le mode opératoire sera appliqué en routine, quelles seront les performances en routine de la méthode. Dans ce contexte, un essai consiste à faire un mesurage sur un échantillon de validation, de préférence dont la valeur de référence a été établie avec une incertitude connue. Les notations sont rappelées à la Partie 3.

Pour réaliser ce plan, prévoir :

- I séries de mesures ($1 \leq i \leq I$) ;
- pour chaque série, effectuer J répétitions ($1 \leq j \leq J$) ;
- K niveaux de concentration ($1 \leq k \leq K$) couvrant le domaine d'application de la méthode. Dans ce contexte, on entend par un niveau une valeur de référence.

Construire un tableau qui rassemblera les données brutes selon le modèle du **tableau 1**. Les valeurs de référence X_k des échantillons de validation sont exprimées en quantités absolues (mg, g, log,...) ou relatives (mg/kg, µg/kg, mg/l,...). Pour les méthodes indirectes, la réponse Y est notée dans une unité correspondant à celle de la méthode instrumentale (surface du pic, hauteur de pic, absorbance,...). Pour les méthodes directes, elle est directement exprimée dans la même unité que celle des valeurs de référence.

Tableau 1 : organisation des mesures du plan de validation

Niveaux	Séries (Jours)	Concentrations des échantillons de validation (valeurs de référence) X	Mesurages (Réponses instrumentales) Y			
			1	2	...	J
1	1	x_{11}	y_{111}	y_{121}	...	y_{1J1}
			
	I		...			
2	1					
	...					
	I					
...				
K	1					
				
	I	x_{IK}				y_{IJK}

Choisir le nombre de séries, de répétitions et de niveaux de validation

L'influence du nombre de séries I est discutée au § 6.3.2 de la norme ISO 5725-1 à propos du choix du nombre de laboratoires pour une étude inter-laboratoires (NF ISO 5725-2:1994). D'ailleurs, à partir de ces données, l'International Union for Pure and Applied Chemistry (IUPAC) a décidé que I doit être supérieur ou égal à 8 pour qu'une étude inter-laboratoires soit reconnue comme recevable au plan scientifique. Dans le cas d'une validation interne, il semble irréaliste pour des raisons économiques de maintenir une telle exigence. C'est pourquoi, on peut choisir les exigences minimales suivantes :

- le nombre de séries I égal ou supérieur à 3. Une série peut être représentée par un jour mais aussi par une combinaison de diverses sources d'incertitude, comme plusieurs appareils, plusieurs opérateurs et plusieurs jours.
- le nombre **constant** de répétitions par série et par niveau J égal ou supérieur à 2.
- le nombre de niveaux de concentration K égal ou supérieur à 3. Il est indispensable d'avoir $K \geq 3$ car on pourra alors vérifier la linéarité entre les concentrations des valeurs de référence et les concentrations retrouvées : on a besoin de 3 niveaux pour conduire cette vérification. Mais, lorsqu'il est nécessaire de valider la méthode à proximité de sa limite de quantification LQ , il est conseillé de choisir K égal ou supérieur à 4 (Feinberg M., 2010c).

Si on a choisi $I = 3$, $J = 2$ et $K = 3$, le plan de validation comporte 18 essais. À part ces limites minimales, une grande flexibilité est laissée dans le choix du nombre d'essais.

Mais il faut se rappeler que plus le nombre d'essais est élevé, meilleures sont les estimations des critères de validation, les intervalles de tolérance resserrés et la confiance qu'on peut avoir dans la validation augmentée. Cependant, pour améliorer les choses, il est généralement préférable d'augmenter le nombre de séries plutôt que le nombre de répétitions ou de niveaux.

1.5 Planifier l'étalonnage (pour les méthodes indirectes)

Organiser les essais d'étalonnage

Le plan d'étalonnage a pour objectif de permettre l'estimation des coefficients du modèle de la courbe d'étalonnage². Il n'est obligatoire que pour les méthodes indirectes qui requièrent un étalonnage pour réaliser la quantification. Dans le cas des méthodes directes (comme la cellulose brute) la quantification se fait sans étalonnage, souvent par une pesée, l'analyte étant défini par la méthode. Le dispositif expérimental utilisé (nombre de niveaux d'étalonnage, de répétitions par niveau et l'étendue du domaine d'étalonnage) doit refléter le plus exactement possible ce qui sera fait en routine. L'étendue de la gamme d'étalonnage n'est pas obligatoirement confondue avec le domaine d'application de la méthode. Pour réaliser ce plan, on va choisir les paramètres minimums suivants :

- prévoir I séries ($1 \leq i \leq I$). Ce nombre doit absolument être le même que pour le plan de validation. Par exemple, si le jour se confond avec la série, les mesurages d'étalonnage doivent être effectués les mêmes jours que les mesurages de validation. On calculera ainsi un modèle d'étalonnage pour chaque jour et non pas un modèle global ;
- pour chaque niveau, effectuer J' mesurages répétés ($1 \leq j \leq J'$). Le nombre de répétitions J' peut être différent du nombre J choisi pour le plan de validation ;

² Sous l'influence de la littérature anglo-saxonne, de nombreux analystes français utilisent le terme « calibration » à la place d'étalonnage.

- pour chaque série prévoir K' niveaux ($1 \leq k \leq K'$) étalons de concentrations connues, couvrant la gamme d'étalonnage. Le nombre de niveaux K' peut être différent du nombre K choisis pour le plan de validation. Par exemple, on peut choisir $K' = 1$ si on est sûr que la courbe d'étalonnage est une droite qui passe par le 0.

Construire le **tableau 2** qui résume l'ensemble des essais et mesurages à effectuer. Dans ce tableau les concentrations des étalons sont exprimées en quantités absolues (mg, g,...) ou relatives (mg/kg, $\mu\text{g}/\text{kg}$, mg/l,...). Après réalisation, les réponses sont notées dans les unités de la méthode instrumentale (surface du pic, hauteur de pic, absorbance,...).

Choisir le nombre de séries, de répétitions et de niveaux d'étalonnage

La valeur souhaitable pour K' dépend du type de fonction de réponse instrumentale attendue, comme l'indique le **tableau 3**. La règle préconisée consiste à prendre au moins un niveau de plus qu'il n'y a de coefficients dans le modèle d'étalonnage. Cette contrainte est liée à l'utilisation d'une méthode de régression aux moindres carrés pour estimer les coefficients du modèle (Azais J.M., Bardet J.M., 2006 - Huet S. *et al.*, 2004). Cependant, on peut aussi utiliser la valeur minimale qui figure dans le même tableau. Les K' étalons doivent permettre une bonne estimation statistique des paramètres de la fonction de réponse. La méthode des moindres carrés, qui sert aux calculs, est très sensible au dispositif expérimental et il faut veiller à ce point.

Tableau 2 : organisation des essais du plan d'étalonnage

Niveaux des étalons	Séries (Jours)	Concentrations des échantillons de validation (valeurs de référence) X	Mesurages (Réponses instrumentales) Y			
			1	2	...	J'
1	1	x_{11}	y_{111}	y_{121}	...	$y_{1J'1}$
			
	I		...			
2	1					
	...					
	I					
...				
K'	1					
				
	I	$x_{IK'}$				$y_{IJ'K'}$

Le nombre de répétitions J' pour chaque étalon dans chaque série est laissé au choix. Cependant, il est recommandé de prendre $J' \geq 2$. Lors de l'utilisation en routine, on pourra modifier le mode opératoire normalisé de la méthode en fonction des conditions démontrées comme étant capables de fournir des résultats acceptables.

Si la méthode nécessite un étalonnage, il faut veiller à ce qu'aucune des réponses y obtenues sur les échantillons de validation ne dépasse la réponse la plus élevée ou la plus petite, recueillie dans le tableau 1. En d'autres termes, il ne faut pas extrapoler le modèle en-dehors des concentrations où il a été construit.

1.6 Réaliser les essais

Il est impératif de réaliser les mesures pour le plan de validation **en appliquant la méthode telle qu'elle sera utilisée en routine**. En particulier, appliquer au mieux les conditions d'étalonnage même si celles-ci pourront être modifiées au vu des résultats obtenus par le profil d'exactitude. Par contre, le nombre de « répétitions » utilisées pour exprimer un résultat final doit être fidèlement respecté. Par exemple, si chaque résultat final est exprimé comme la moyenne de 2 répétitions, chaque essai doit être fait selon cette méthode.

Dans le cas où des effets de matrice ont été mis en évidence lors du développement de la méthode, on peut être amené à réaliser plusieurs plans de validation pour calculer un facteur de correction qui sera appliqué en routine (sauf interdiction réglementaire). Il faut alors effectuer deux plans de validation sur deux échantillons indépendants et utiliser les résultats du premier pour calculer le facteur de correction qui sera appliqué au second plan (Feinberg M., 2010c)

Le rôle du profil d'exactitude est d'estimer, à partir des résultats obtenus lors de la validation, quelle garantie aura l'utilisateur que la méthode utilisée en routine fournira des résultats acceptables. C'est pourquoi, la réalisation des plans doit respecter les conditions suivantes :

- **Exigence d'être en condition de fidélité intermédiaire ou de reproductibilité.** Les mesurages doivent être réalisés en condition de fidélité intermédiaire ou de reproductibilité, avec pour objectif de prendre en compte le plus de sources d'incertitude possible. Comme ces sources d'incertitude dépendent de la méthode et des matrices analysées, il est impossible de définir une stratégie unique. Il est recommandé de construire un diagramme causes/effets pour choisir les conditions les plus globales. Dans la plupart des cas, on pourra décider de faire les mesurages sur plusieurs jours dans la mesure où on prend ainsi en compte les incertitudes dues à l'appareillage (réglages, et préparation des réactifs), à l'étalonnage, au personnel, à la conservation de l'échantillon... Mais pour des échantillons très instables, on pourra réaliser les mesurages sur un seul jour avec plusieurs opérateurs ou instruments.
- **Exigence de couvrir le domaine d'application.** Une étude de validation doit couvrir l'ensemble du domaine d'application de la méthode. Pour satisfaire ces deux contraintes, plusieurs choix sont possibles et, dans tous les cas, avant de commencer il faut disposer de quantités suffisantes d'échantillons pour faire tous les mesurages :
 - utiliser un seul échantillon dont on sait faire varier le niveau de concentration, par exemple au moyen d'ajouts dosés (approche de choix pour une méthode indirecte) ;
 - utiliser plusieurs échantillons (ou même plusieurs matrices proches) présentant des niveaux de concentration variés (généralement le cas pour une méthode directe) ;
 - préparer des matrices synthétiques (matières premières ou produits chimiques) ;
 - utiliser des matériaux de référence externes ou certifiés...
- **Exigence de synchroniser les plans.** Dans le cas, où deux plans sont nécessaires – étalonnage et validation – il est indispensable de réaliser les essais dans les mêmes conditions de fidélité intermédiaire. Par exemple, si le jour est la source de variation choisie, il faut réaliser des mesurages d'étalonnage et de validation pour un même jour donné. Les données d'étalonnage de ce jour serviront à prédire les concentrations à partir des données de validation faites le même jour.

1.7 Calculer les concentrations prédites inverses [méthodes indirectes]

Avant tout traitement numérique, il est recommandé de procéder à une illustration graphique préalable et à un examen visuel des données afin de détecter une erreur flagrante, comme une donnée mal enregistrée. Ces graphiques peuvent être mis en annexe du dossier d'évaluation. La présence d'une donnée anormale amène à suspecter une mauvaise mise au point de la méthode et peut invalider la procédure de validation.

Calculer les modèles d'étalonnage

Pour les méthodes directes, cette étape de calcul n'existe pas.

Pour les méthodes indirectes, il est nécessaire d'exprimer la réponse instrumentale Y en fonction des concentrations x des étalons, à l'aide d'un modèle mathématique f de la forme :

$$Y = f(x) \quad (2)$$

Les fonctions f classiquement utilisées sont regroupées au **tableau 3** mais cette liste n'est pas limitative. Les paramètres a_1, a_2, \dots sont appelés les paramètres du modèle.

Calculer les paramètres du modèle d'étalonnage à partir des données recueillies pour chaque série k , de façon à obtenir k ensembles de valeurs des paramètres. Le même type de modèle doit être utilisé pour l'ensemble des données, quelle que soit la série, mais les valeurs des paramètres peuvent être différentes d'une série à l'autre. Cette approche permet de prendre en compte les variations inter-séries (inter-jours) observées.

Le calcul des estimations des coefficients du modèle d'étalonnage peut faire appel aux diverses techniques statistiques classiques détaillées dans la littérature :

1. Régression par la méthode des moindres carrés (AFNOR, 1996)
2. Régression pondérée (Azaïs J.M., Bardet J.M., 2006)
3. Régression non linéaire (Huet S., et al., 2004)

Tableau 3 : principales fonctions de réponse utilisables pour un étalonnage

Type	Équation	Souhaitable pour K'	Minimum pour K'
Droite passant par l'origine	$Y = a_1x$	2	1
Droite	$Y = a_0 + a_1x$	3	2
Fonction quadratique	$Y = a_0 + a_1x + a_2x^2$	4	3
Fonction logistique à 4 paramètres	$Y = a_0 + \frac{a_3 - a_0}{1 + \left(\frac{a_2}{x}\right)^{a_1}}$	5	5

La méthode la plus accessible est celle des moindres carrés, car elle est disponible sur les tableurs. La régression pondérée s'applique dans les cas où les variances des réponses ne sont pas homogènes entre niveaux. Les estimations des paramètres du modèle pour chaque série sont regroupées dans le **tableau 4**.

Si le type de modèle d'étalonnage a été clairement établi lors du développement de la méthode, on pourra choisir le nombre minimum pour K' . Cependant, on aura intérêt à prendre le nombre souhaitable ou plus afin de tester si d'autres modèles plus complexes ne sont pas

mieux adaptés. En effet, à partir des données du plan d'étalonnage, il est possible d'estimer plusieurs modèles d'étalonnage puis de construire ainsi plusieurs profils d'exactitude. On pourra alors retenir le modèle qui fournit le profil le plus favorable. Il conviendra alors de modifier le mode opératoire en fonction du modèle choisi, sans pour autant modifier les limites d'acceptabilité λ .

Tableau 4 : estimations des paramètres du modèle en fonction de la série

Série	1	...	I
a_0			
a_1			
...			

Calculer les concentrations retrouvées par prédiction inverse

Pour les méthodes indirectes, les modèles d'étalonnage servent à calculer les concentrations retrouvées, à partir des données au plan de validation (Tableau 2), en utilisant la fonction inverse du modèle d'étalonnage, selon le modèle mathématique suivant :

$$\hat{x} = z = f^{-1}(Y) \quad (3)$$

La fonction inverse est appelée équation de prédiction inverse et les valeurs z ainsi obtenues sont appelées concentrations retrouvées. Le **tableau 5** fournit ces équations selon la fonction de réponse choisie au tableau 3.

Tableau 5 : équations de prédiction inverse selon le type de la fonction de réponse

Fonction de réponse	Fonction inverse de calcul de la concentration retrouvée
Droite passant par l'origine	$z = \frac{Y}{a_1}$
Droite	$z = \frac{Y - a_0}{a_1}$
Fonction quadratique	$z = \frac{-a_1 + \sqrt{a_1^2 - 4a_2(a_0 - Y)}}{2a_2}$
Logistique à 4 paramètres	$z = \frac{a_2}{\left(\frac{a_3 - a_0}{Y - a_0} - 1\right)^{\frac{1}{a_1}}}$

Utiliser les données du plan de validation et les paramètres des modèles d'étalonnage estimés à partir du plan d'étalonnage pour calculer les concentrations retrouvées dans les échantillons de validation. Effectuer ces calculs par série, avec le modèle de la série correspondante. Ce mode de calcul souligne le fait qu'il est indispensable que les essais des plans d'étalonnage et de validation soient synchronisés c'est-à-dire réalisés le même jour et/ou par le même opérateur et/ou par le même laboratoire.

1.8 Calculer les critères de validation

Calculer la justesse par série

Rassembler l'ensemble des données dans un tableau selon le modèle du Tableau 6. Au cas où différents modèles d'étalonnage sont testés, construire autant de tableaux qu'il y a de modèles. Utiliser les valeurs de z pour calculer les écarts-types de répétabilité, inter-séries et de fidélité intermédiaire (ou de reproductibilité). Ce calcul s'effectue indépendamment pour chaque niveau de concentration k (avec $1 \leq k \leq K$) selon le principe de la norme ISO 5725-2, comme il est décrit dans Feinberg M. (2010c). Pour certaines mesures exprimées en logarithmes, comme les comptages microbiologiques ou unité formant colonie (ufc), exprimés en \log_{10} (ufc), il n'est pas utile de calculer les valeurs relatives.

Tableau 6 : concentrations retrouvées par prédiction inverse obtenues à partir des données du plan de validation

Niveau	Série	Répétition	Concentration théorique X	Concentration retrouvée Z	Biais absolu	Biais relatif
1	1	1	x_{111}	z_{111}	b_{ijk}	$b_{ijk}\%$
1	1	2				
1		...				
...		J				
1	I	1				
1	I	2				
1	I	...				
1	I	J				
K	I	J	x_{IJK}	z_{IJK}	b_{IJK}	

Dans le **tableau 6**, il est possible de calculer un ou plusieurs critères exprimant la justesse (sous la forme d'un biais). Ces critères sont les suivants :

$$\text{Biais absolu : } b_{ijk} = z_{ijk} - x_{ijk} \quad (4)$$

$$\text{Biais relatif : } b_{ijk}\% = \frac{z_{ijk} - x_{ijk}}{x_{ijk}} \times 100 \quad (5)$$

$$\text{Recouvrement} = \frac{z_{ijk}}{\bar{x}_k} \times 100 \quad (6)$$

Calculer la justesse et la fidélité par niveau

Rassembler avec les critères de justesse, calculés pour un niveau et non plus pour un mesurage individuel, sous la forme d'un tableau, comme le **tableau 7**.

Dans les cas très particuliers où la justesse n'est pas exigée, la valeur de référence moyenne peut être fixée comme étant égale à la concentration retrouvée moyenne. Il est alors inutile de calculer un biais moyen ou un taux de recouvrement moyen.

Tableau 7 : critères de fidélité et de justesse par niveau

Critères	Symbole	Niveaux		
		1	...	K
Valeur de référence moyenne	$\bar{\bar{x}}_k$			
Concentration retrouvée moyenne	$\bar{\bar{z}}_k$			
Écart-type de répétabilité	s_{kr}			
Écart-type inter-séries	s_{kB}			
Écart-type de fidélité intermédiaire	$s_{kFI} = \sqrt{s_{kr}^2 + s_{kB}^2}$			
Coefficient de variation de la fidélité intermédiaire	$\frac{s_{kFI}}{\bar{\bar{x}}_k} \times 100$			
Biais moyen absolu	$\bar{\bar{z}}_k - \bar{\bar{x}}_k$			
Biais moyen relatif	$\left(\frac{\bar{\bar{z}}_k - \bar{\bar{x}}_k}{\bar{\bar{x}}_k} \right) \times 100$			
Taux de recouvrement moyen	$\frac{\bar{\bar{z}}_k}{\bar{\bar{x}}_k} \times 100$			

Calculer les intervalles de tolérance

La méthode de calcul proposée par Mee (1984) est celle qui a été choisie pour cette procédure. Elle a aussi été adoptée par une commission de la Société des sciences et techniques pharmaceutiques (SFSTP - 2003, 2006). Le calcul se fait à partir des données du **tableau 7**, indépendamment pour chaque niveau de concentration k . Pour des raisons de simplification des formules, l'indice k est omis dans les formules suivantes.

Exprimer l'intervalle de tolérance comme un intervalle symétrique autour de la concentration retrouvée moyenne $\bar{\bar{z}}$ du niveau :

$$\bar{\bar{z}} \pm k_{tol} \times s_{IT} \quad (7)$$

Calculer l'écart-type de l'intervalle de tolérance s_{IT} selon la formule (8) ; il servira aussi à l'estimation de l'incertitude de mesure. Il est obtenu grâce aux formules suivantes :

$$s_{IT} = s_{FI} \left(\sqrt{1 + \frac{1}{I \times J \times B^2}} \right) \quad (8)$$

$$B = \sqrt{\frac{R+1}{J \times R+1}} \quad (9)$$

$$R = \frac{s_B^2}{s_r^2} \quad (10)$$

La quantité k_{tol} est appelé **facteur de couverture de l'intervalle de tolérance** et vaut :

$$k_{tol} = t_{v; \frac{1+\beta}{2}} \quad (11)$$

$t_{v; \frac{1+\beta}{2}}$ est le quantile de la distribution t de Student pour v degrés de liberté et β la probabilité du contenu de l'intervalle de tolérance. Le nombre de degrés de liberté v est calculé selon méthode d'approximation et donne :

$$v = \frac{(R+1)^2}{\left(R + \frac{1}{J}\right)^2 + \frac{1 - \frac{1}{J}}{IJ}} \quad (12)$$

Pour chaque niveau k , calculer l'écart-type de l'intervalle de tolérance s_{IT} , le facteur de couverture k_{tol} et les limites basse et haute de l'intervalle de tolérance. La valeur choisie pour β doit être au moins de 80%. L'ensemble des calculs est rassemblé dans un tableau de la forme du **tableau 8**.

Tableau 8 : résumé des différents éléments qui serviront au profil d'exactitude, incluant les limites des intervalles de tolérance par niveau

Critères	Symbole	Niveaux		
		1	...	K
Valeur de référence moyenne	$\bar{\bar{x}}$			
Limite de tolérance basse	$\bar{\bar{z}} - k_{tol} \times s_{IT}$			
Limite de tolérance haute	$\bar{\bar{z}} + k_{tol} \times s_{IT}$			
Limite de tolérance basse relative	$\left(\frac{\bar{\bar{z}} - k_{tol} \times s_{IT}}{\bar{\bar{x}}}\right) \times 100$			
Limite de tolérance haute relative	$\left(\frac{\bar{\bar{z}} + k_{tol} \times s_{IT}}{\bar{\bar{x}}}\right) \times 100$			
Limite d'acceptabilité basse relative	$(1 - \lambda) \times 100$			
Limite d'acceptabilité haute relative	$(1 + \lambda) \times 100$			

Le nombre v (Eq.(12)) est rarement un nombre entier et il convient d'utiliser des tables de la loi de Student qui acceptent des nombres de degrés de liberté non entiers. Cependant, il est possible d'approcher la valeur du quantile par interpolation linéaire entre les 2 nombres de degré de liberté entiers qui encadrent v . Ce calcul approché est illustré dans la feuille de calcul (dans les cellules B30 à B32) présentée par Max Feinberg (2010c).

Le rapport $R = \frac{s_B^2}{s_r^2}$ entre la variance inter-séries et la variance de répétabilité intervient dans

plusieurs formules. Il traduit l'importance relative de l'effet de la série. Par exemple, dans le cas où l'effet jour est l'effet série, si la méthode est très stable d'un jour à l'autre et est capable de fournir des résultats très proches pour un même échantillon, R est proche de 1 et ne jouera qu'un rôle secondaire. Par contre, si ce rapport augmente le nombre de degrés de

liberté est diminué et, plus le nombre de degrés de liberté est petit, plus le quantile de la distribution t de Student est élevé et élargit l'intervalle de tolérance.

1.9 Construire le profil d'exactitude

Le profil d'exactitude peut être construit de différentes façons, en fonction du type de données traité. La méthode la plus classique, lorsqu'on a à faire à des concentrations relatives, est celle présentée à la **figure 1** où les performances sont exprimées de façon relative, par un taux de recouvrement. Mais si les données sont des comptages exprimés en logarithmes, il vaut mieux exprimer les performances, comme des différences entre 2 logarithmes, ce qui est équivalent à un rapport.

Pour construire le profil d'exactitude, sélectionner dans le **tableau 8**, les lignes suivantes :

- à reporter sur l'axe horizontal
 - 1) Les valeurs de référence moyennes.
- à reporter sur l'axe vertical
 - 2) Les limites de tolérance basses relatives ;
 - 3) Les limites de tolérance hautes relatives ;
 - 4) Les taux de recouvrement moyens ;
 - 5) Les limites d'acceptabilité basses relatives ;
 - 6) Les limites d'acceptabilité hautes relatives.

Reporter ces données sur un graphique en utilisant les valeurs de référence moyennes pour dessiner l'axe des abscisses, comme l'illustre la **figure 1**.

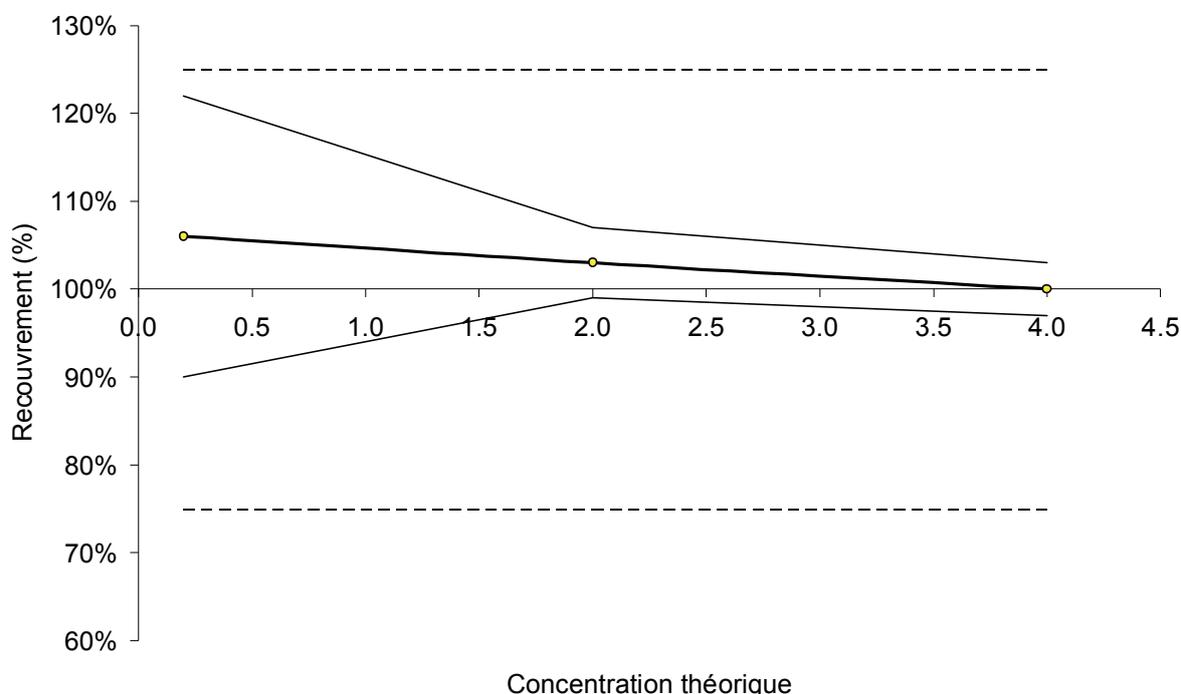


Figure 1 : profil d'exactitude exprimé par le taux de recouvrement

1.10 Interpréter le profil d'exactitude pour la validation

Pour utiliser le profil d'exactitude en vue de valider une méthode, il faut avoir fixé les deux critères de décision suivants :

- **les limites d'acceptabilité $\pm\lambda$.** Elles servent à traduire les objectifs pratiques des utilisateurs. Elles délimitent un intervalle autour de la valeur de référence. Le plus souvent, ces limites sont réglementaires ou issues de la réglementation. Mais dans le cas où il n'existe pas de référence établie, il convient de prendre en compte les attentes des utilisateurs finaux, comme une LQ donnée (voir § 1.2) ;
- **la proportion β .** Elle représente la proportion de futurs résultats qui seront en moyenne compris dans les intervalles de tolérance. La valeur choisie pour β dépend largement du champ d'application (contrôle sanitaire, contrôle de fabrication, etc.). Il est évident que plus β est petit, par exemple 70 %, plus la méthode risque de produire des résultats qui ne correspondent pas aux spécifications annoncées. C'est pourquoi, dans la méthode du profil d'exactitude cette proportion a été fixée à 80 %, au moins.

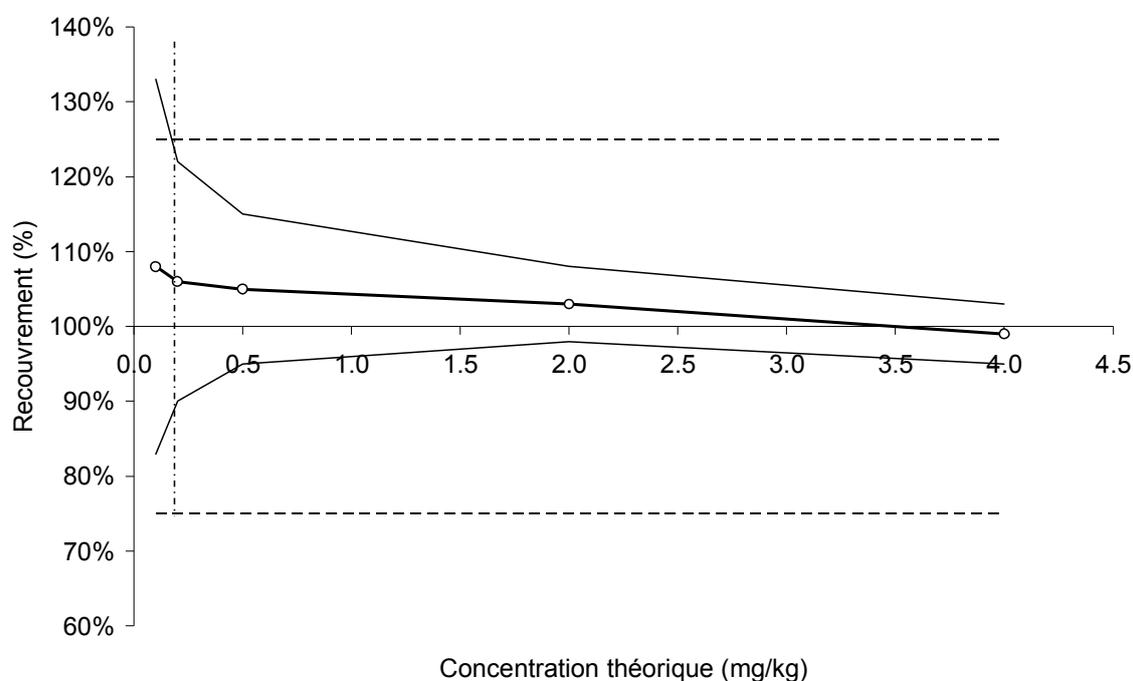


Figure 2 : exemple d'une méthode validée sur une zone réduite du domaine de validation

Règles de décision

Reporter les limites d'acceptabilité sur le graphique du profil d'exactitude pour permettre une interprétation visuelle directe des résultats. La **figure 2** illustre une situation classique où la majeure partie des intervalles de tolérance est comprise dans l'intervalle d'acceptabilité. Dès que l'intervalle de tolérance sort de l'intervalle d'acceptabilité, on peut conclure que la méthode n'est plus capable de fournir suffisamment de résultats acceptables, en fonction des choix faits au départ de l'étude. Par exemple, pour $\beta = 80\%$ et une limite d'acceptabilité de $\pm 25\%$, on peut conclure que la méthode est valide entre environ 0,2 mg/kg et 4,0 mg/kg. En outre, ce graphique fournit d'autres indications. En particulier, sur la **figure 2**, il apparaît que

la justesse varie avec la concentration. Le taux de recouvrement qui traduit la justesse est d'environ 108 % aux basses concentrations pour être proche de 100 % aux concentrations élevées. À l'aide des résultats statistiques regroupés dans le tableau construit sur le modèle du tableau 7, on peut trouver différents éléments qui expliquent le comportement de ce biais de la méthode.

Dans le cas illustré par la **figure 3**, la méthode présente un biais systématique élevé : le taux de recouvrement est seulement d'environ 70 % ; dû par exemple à un effet de matrice non contrôlé. Il faut alors conclure que la méthode n'est pas valide dans le domaine étudié. Si les pratiques de la profession ou la réglementation le permettent, il est alors possible d'appliquer un facteur de correction qui viendra corriger ce biais. La procédure à appliquer pour calculer et valider ce facteur de correction est décrite par Max Feinberg (2010c).

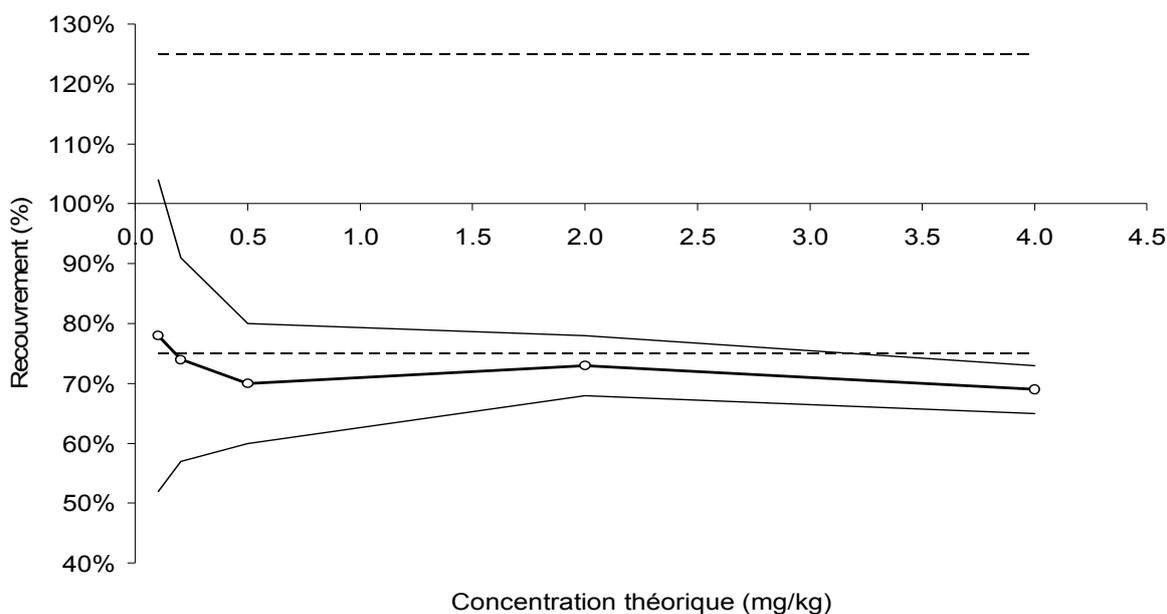


Figure 3 : exemple d'un profil d'exactitude pour une méthode présentant un biais systématique important, caractérisé par un taux de recouvrement d'environ 70 %

Définir du domaine de validité

Le domaine de validité est déterminé par la zone du domaine de validation dans laquelle la méthode fournit une proportion de résultats acceptables au moins égale à β . Il est limité par une borne supérieure qui équivaut à la limite de quantification inférieure et une limite supérieure qui correspond à la limite de quantification supérieure.

Choisir la procédure d'étalonnage pour la routine

Dans la mesure où avec les méthodes indirectes on dispose de données d'étalonnage, il est possible de construire plusieurs profils d'exactitude avec le même jeu de données mais en utilisant différents modèles d'étalonnage, par exemple un modèle passant par le zéro et construit avec une seule solution étalon qui est comparé à un modèle linéaire complet calculé à partir de 3 solutions étalons. On peut alors sélectionner le profil le plus favorable ou la procédure d'étalonnage la plus simple qui, malgré tout, permet d'atteindre l'objectif fixé.

Identifier les points atypiques

La méthode de calcul proposée dans la méthode du profil d'exactitude exige que le plan d'expérience de validation soit équilibré, c'est-à-dire que le nombre de répétitions I par série soit constant. C'est pourquoi, elle ne prévoit pas de méthodes de détection et/ou de rejet de données atypiques (aberrantes).

Toutefois, si le responsable de la validation juge que le rejet de données atypiques est nécessaire après un examen visuel d'une représentation graphique, il a toute latitude pour mettre en œuvre la procédure qui lui semblera la mieux adaptée, par exemple, en s'appuyant sur les recommandations de la norme ISO 5725-2. De plus, il est toujours possible de recommencer les mesures pour un niveau de concentration et une série donnés.

Un trop grand nombre de points aberrants souligne la mise en œuvre médiocre de la méthode d'analyse et met en doute le bien-fondé de sa validation.

2. Transférer la méthode en routine

Finalement, lors de l'application en routine de la méthode ainsi validée, il convient de vérifier de façon régulière si les résultats obtenus restent acceptables, par exemple au moyen d'une carte de contrôle, en particulier, si on utilise un facteur de correction (Feinberg M., 2010c).

Des actions correctives inefficaces peuvent ainsi conduire à la revalidation de la méthode. La revalidation peut alors être complète ou partielle. Une revalidation partielle consistant à faire un plan d'expérience à un seul niveau de concentration pour vérifier si l'intervalle de tolérance obtenu reste bien entre les limites d'acceptabilité.

Bibliographie

AFNOR (1994) Norme NF ISO 5725-2, *Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 2 : Méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée.*

AFNOR (1996) Norme NF ISO 11095 *Étalonnage linéaire utilisant des matériaux de référence*, Afnor, Paris

Azaïs J.M., Bardet J.M.; (2006) *Le modèle linéaire par l'exemple, Régression, analyse de la variance et plans d'expérience illustrés avec R, SAS et Splus*, Dunod, Paris

Commission européenne (2002) Décision de la commission 2002/657 du 12 août 2002 portant modalités d'application de la directive 96/23/CE du Conseil en ce qui concerne les performances des méthodes d'analyse et l'interprétation des résultats

Commission SFSTP, P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat ; Validation des procédures analytiques quantitatives : harmonisation des démarches.

Partie I. : *STP Pharma Pratiques* (2003) 13:3, 101-138

Partie II. : *Statistiques, STP Pharma Pratiques* (2006) 16:1, 30-60

Partie III : *Exemples d'application, STP Pharma Pratiques* (2006) 16:2, 87-121

FAO (1997) Manuel sur le contrôle de la qualité des produits alimentaires. 14 : Assurance de la qualité dans le laboratoire d'analyse chimique des aliments, Rome, <http://www.fao.org/docrep/T0845F/t0845f00.HTM>

- Feinberg M., (2009) *Labo-Stat : Guide de validation des méthodes d'analyse* Lavoisier tec&Doc, Paris
- Feinberg M. (2010c) Interprétation du profil d'exactitude in Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 45-59
- Huet S., Bouvier A., Poursat M.A., Jolivet E. (2004) *Statistical Tools for Nonlinear Regression* Springer Verlag, New-York, ISBN 0-387-40081-8
- Mee R.W. (1984) β -Expectation and β -Content Tolerance Limits for Balanced One-Way ANOVA Random Model *Technometrics* 26:3, 251-254
- Commission du Codex Alimentarius (2005) Manuel de procédure de la 15^{ème} édition, http://www.codexalimentarius.net/web/procedural_manual.jsp)
- CX/MAS 09/30/7, Directives pour établir des critères pour l'identification des méthodes d'analyse adéquates (Conversion des méthodes pour les éléments traces en critères) <ftp://ftp.fao.org/codex/>)