RÉSOLUTION OENO 8/2005

RECOMMANDATIONS HARMONISEES POUR LA VALIDATION DES METHODES D'ANALYSE PAR UN SEUL LABORATOIRE (RAPPORT TECHNIQUE)

L'ASSEMBLEE GENERALE,

VU l'Article 2, paragraphe 2 iv de l'accord du 3 avril 2001 portant création de l'Organisation internationale de la vigne et du vin,

SUR PROPOSITION de la Sous-Commission des méthodes d'analyse et d'appréciation des vins,

DECIDE: de créer une Annexe E au Recueil des méthodes Internationales d'analyse des vins et des moûts intitulée «Assurance qualité dans les laboratoires »

DECIDE : de regrouper dans cette annexe E toutes les résolutions déjà adoptées relatives à l'assurance qualité dans les laboratoires.

DECIDE : d'introduire dans l'annexe E dudit Recueil les directives suivantes:

Recommandations harmonisées pour la validation desmethodes d'analyse par un seul laboratoire (rapport technique)

Synthèse

La validation de méthode figure parmi les mesures universellement reconnues comme faisant nécessairement partie d'un système exhaustif d'assurance qualité dans le domaine de la chimie analytique. L' ISO, l'UICPA (Union internationale de chimie pure et appliquée) et AOAC INTERNATIONAL ont coopéré afin de produire des protocoles accordés ou des recommandations sur la Conduite et interprétation des études de performance de méthodes"[1] , sur la Vérification de l'Aptitude des Laboratoires de chimie analytique[2] , sur le Contrôle interne de Qualité dans les Laboratoires de Chimie Analytique"[3] et sur l'Utilisation des données de recouvrement dans les Mesures Analytiques)".[4] Le Groupe de Travail ayant formulé ces protocoles/recommandations a été mandaté par l' UICPA pour préparer des recommandations sur la validation des méthodes d'analyse par un seul laboratoire. Ces recommandations précisent les bases des procédures qui devraient être utilisées pour assurer une validation adéquate des

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

méthodes d'analyse.

Un projet de recommandations a été discuté au cours d'un Symposium international sur l'harmonisation des systèmes d'assurance qualité dans les laboratoires d'analyse chimique dont les actes ont été publiés par la Royal Society of Chemistry du Royaume-Uni.

TABLE DES MATIERES

Synthèse

1. INTRODUCTION

1.1. Contexte

1.2. Protocoles, Normes et Guides existants

2. DEFINITIONS ET TERMINOLOGIE

2.1. Définitions Générales

2.2. Définitions utilisees uniquement dans ce document :

3. Validation des Méthodes, Incertitude et Assurance Qualité

4. Principes de base de la validation de methode

4.1. Spécification et portee de la validation

4.2. Hypothèses des essais

4.3. Sources d'Erreur en Analyse

4.4. Effets liés à la méthode et au laboratoire

5. Conduite des etudes de Validation

6. Portee des etudes de validation

6.1. Le laboratoire est tenu d'utiliser une methode "entierement" validee

6.2. Le laboratoire doit utiliser une methode entierement validee mais avec une nouvelle matrice

6.3. Le laboratoire doit utiliser une methode bien etablie mais qui n'a pas fait l'objet d'une etude collaborative

6.4. La methode a fait l'objet d'une publication dans la litterature scientifique celle-ci fournissant quelques caracteristiques analytiques

6.5. La methode a fait l'objet d'une publication dans la litterature scientifique, mais sans precision sur les caracteristiques, ou bien la methode a été developpee en interne

6.6. La methode est empirique

6.7. L'analyse est "ad hoc"

6.8. Changements de personnel et d' equipement

7. Recommandations

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

8. Références Bibliographiques

ANNEXE A: Notes sur les conditions requises pour l'etude des caracteristiques de performance d'une methode

A1 Domaine d'application

A2 Spécificité

A.3. Etalonnage et linearite

A3.1 Linéarité et ordonnée à l'origine

A3.2 Etude de l'effet matrice

A3.3 Procédure finale d'étalonnage

A4 Justesse

A4.1 Evaluation de la justesse

A4.2 Conditions à appliquer pour les expériences de justesse

A4.3 Valeurs de référence à utiliser pour les expériences de justesse

A4.3.1 Matériaux de Référence Certifiés (MRC)

A.4.3.2 Matériaux de référence :

A4.3.3 Utilisation d'une méthode de référence

A4.3.4 Utilisation des ajouts dosés / recouvrement

A5 Fidelite

A6 RECOUVREMENT

A7 Domaine de validité

A8 limite de détection

A9 Limite de determination ou limite de quantification

A10 Sensibilité

A11 Robustesse

A12 Aptitude à l'essai

A13 variation de matrice

A14 Incertitude de Mesure réf. roger wood

ANNEXE B. considerations supplémentaires pour l'evaluation de l'incertitude dans les etudes de VALIDATION

B1 Analyse de la Sensibilité

B2 Jugement

© OIV 2005

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

1. INTRODUCTION

1.1. Contexte

Des méthodes d'analyse fiables sont requises pour assurer la conformité avec les réglementations nationales et internationales dans tous les domaines d'analyse. Par conséquent, il est reconnu partout dans le monde qu'un laboratoire doit prendre les dispositions appropriées pour s'assurer qu'il est en mesure de fournir (et de fournir effectivement) des données du niveau de qualité requis. De telles dispositions comprennent:

- L'utilisation de méthodes d'analyse validées;
- L'utilisation de procédures de contrôle interne de qualité;
- La participation à des programmes d'essai d'aptitude technique; et
- L'obtention d'une accréditation selon une Norme Internationale, généralement ISO/CEI 17025.

Notons que l'accréditation ISO/CEI 17025 porte plus particulièrement sur la mise en place de la traçabilité des mesurages, tout en requérant un ensemble d'exigences techniques et de gestion, qui comprennent notamment celles figurant dans la liste ci-dessus.

La validation de méthode est ainsi une composante essentielle des mesures qu'un laboratoire devrait mettre en œuvre pour lui permettre de produire des données analytiques fiables. D'autres aspects du sujet ont été abordés par le Groupe de Travail Interdivisionnel de l' UICPA sur l'Harmonisation des Programmes d'Assurance Qualité pour les Laboratoires d'Analyse, tout particulièrement en élaborant des Protocoles/ Recommandations sur les études de performance des méthodes interlaboratoires,[1] sur les essais d'aptitude technique [2], et sur le contrôle qualité interne. [3]

Dans certains domaines, notamment en matière d'analyse alimentaire, la législation en vigueur exige l'utilisation de méthodes ayant été " complètement validées". [5,6] Il est généralement attendu que la validation " complète" d'une méthode d'analyse comprenne l'examen des caractéristiques de la méthode dans le cadre d'une étude interlaboratoire des performances de celle-ci (également appelée étude collaborative ou essai interlaboratoire). Des protocoles acceptés partout dans le monde ont été

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

établis pour la validation "complète" d'une méthode d'analyse par un essai interlaboratoires, notamment par le Protocole Harmonisé international[1] et par la procédure de l' ISO.[7] Ces protocoles / normes exigent la participation d'un nombre minimum de laboratoires et de matériaux d'essais soumis à l'étude interlaboratoires pour valider complètement la méthode d'analyse. Toutefois, il n'est pas toujours évident dans la pratique ou nécessaire de parvenir à la validation complète d'une méthode d'analyse. Dans ce cas, une "validation interne de méthode" (effectuée par un seul laboratoire) peut être appropriée.

La validation de méthode par un seul laboratoire est adéquate dans plusieurs situations, notamment dans les circonstances suivantes :

- Pour s'assurer de la viabilité de la méthode avant de s'engager dans l'exercice onéreux de l'essai interlaboratoires formel;
- Pour fournir la preuve de la fiabilité d'une méthode d'analyse lorsque des données d'essai interlaboratoires ne sont pas disponibles ou lorsque l'organisation d'un essai interlaboratoires formel n'est pas possible ;
- Pour s'assurer que des méthodes validées "prêtes à l'emploi" sont utilisées correctement.

Lorsqu'il s'agit de déterminer en interne les caractéristiques d'une méthode, il est important que le laboratoire définisse précisément, en accord avec son client, les caractéristiques à évaluer. Cependant, dans certaines circonstances, ces caractéristiques peuvent être imposées par la législation en vigueur (par exemple pour les résidus de médicaments vétérinaires dans l'alimentation ou encore pour les pesticides dans le secteur alimentaire). La portée de l'évaluation entreprise par un laboratoire doit répondre aux exigences définies par la législation.

Cependant, dans certains domaines d'analyse, la même méthode analytique est utilisée par un grand nombre de laboratoires pour déterminer les composants chimiques stables dans des matrices bien définies. Il faut souligner que lorsque l'on peut mettre à la disposition de tous ces laboratoires une méthode appropriée étudiée collectivement, le poids financier de l'essai interlaboratoires venant valider cette méthode peut tout à fait se justifier. L'utilisation d'une méthode étudiée en interlaboratoire diminue de façon importante les efforts qu'un laboratoire doit engager dans un travail de validation approfondi avant d'utiliser une méthode en routine. Un laboratoire qui utilise une méthode étudiée en interlaboratoire, et qui a été reconnue comme étant adaptée à l'utilisation prévue, n'a plus qu'à prouver qu'il

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

peut obtenir les niveaux de performance spécifiées pour la méthode. Une telle vérification du bon usage de la méthode coûte bien moins cher qu'une validation interne complète. Il revient généralement moins cher à la Communauté Analytique de valider une méthode grâce à un essai interlaboratoires dans un premier temps, puis dans une deuxième temps de vérifier les niveaux de performance dans les différents laboratoires souhaitant utiliser cette méthode, que d'assurer indépendamment dans de nombreux laboratoires une validation interne de cette même méthode.

1.2. Protocoles, Normes et Guides existants

Un certain nombre de protocoles et de recommandations[8-19] sur la validation de méthode et sur l'incertitude ont été élaborés, notamment dans AOAC INTERNATIONAL, dans la Conférence Internationale sur l' Harmonisation (CIH) et dans les documents Eurachem :

- Le Manuel de Statistiques de l' AOAC, qui comprend des recommandations sur les études internes (un seul laboratoire) précédant des études collaboratives[13]
- Le texte de la CIH[15] et sa méthodologie,[16] qui établissent des exigences minimales pour les études de validation quand il s'agit de test utilisés pour motiver une demande d'autorisation de mise sur le marché.
- The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics (1998)[12]
- Quantifying Uncertainty in Analytical Measurement (2000)[9]

La validaton de méthode a été également largement débattue lors d'une Consultation Conjointe d'Experts FAO/IAEA en Décembre 1997, sur la Validation des Méthodes Analytiques pour les Contrôles Alimentaires; le rapport de cette consultation est disponible[19].

Les présentes "recommandations" regroupent les principes scientifiques essentiels des documents mentionnés ci-dessus, ceci afin de fournir les informations ayant été soumises à l'approbation internationale, et surtout afin d'indiquer la marche à suivre pour la meilleure pratique de la validation interne d'une méthode (validation effectuée par un seul laboratoire).

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

2. DEFINITIONS ET TERMINOLOGIE

2.1. Définitions Générales

Les termes utilisés dans ce document respectent les définitions ISO et UICPA lorsqu'elles sont disponibles. Les documents suivants contiennent les définitions pertinentes:

i. IUPAC: Compendium of chemical terminology, 1987

ii. International vocabulary of basic and general terms in metrology. ISO 1993

2.2. Définitions utilisees uniquement dans ce document :

Incertitude relative : Incertitude exprimée comme un écart-type relatif.

Domaine de validité : La partie d'un domaine de concentration d'une méthode d'analyse qui a été soumise à la validation.

3. Validation des Méthodes, Incertitude et Assurance Qualité

La validation de méthode utilise un ensemble de tests qui vérifient à la fois les hypothèses de base de la méthode d'analyse et établissent et documentent les caractéristiques de performance d'une méthode; ce faisant, elle démontre si une méthode est adaptée à son emploi analytique. Les caractéristiques de performance habituelles des méthodes d' analyse sont : définition du champs d'application; la spécificité; l'étalonnage (linéarité); la justesse; la fidélité; la gamme de validité; la limite de quantification; la limite de détection; la sensibilité et la robustesse. On peut y ajouter l'incertitude de la mesure et l'aptitude-à l'essai.

Stricto sensu, la validation devrait faire référence à un 'système d'analyse' plutôt qu'à une 'méthode d'analyse', le système comprenant la définition d'un protocole de méthode, d'un domaine de concentration de l'analyte et la spécification de la nature du matériau d'essai. Pour les besoins de ce document, une référence à la 'validation de méthode' s'entendra comme faisant référence à un système d'analyse dans son ensemble. Quand on fera référence à la procédure d'analyse en tant que telle, on parlera du 'protocole'.

Dans ce document la validation de méthode est considérée comme étant distincte des activités continues telles que le contrôle interne de qualité (CIQ) ou les essais

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

d'aptitude technique. La validation de méthode est réalisé une seule fois, ou à intervalles relativement espacés tout au long de la durée d'utilisation de la méthode; elle nous renseigne sur la performance que l'on est en droit d'attendre d'une méthode à l'avenir. Le contrôle interne de qualité (CIQ) nous renseigne sur les performances de la méthode dans le passé. Le CIQ est par conséquent abordé comme étant une activité distincte dans le Programme d'Harmonisation de l' UICPA.[3]

Dans la validation de méthode les caractéristiques quantitatives présentant un intérêt sont celles concernant le degré de précision du résultat que l'on est susceptible d'obtenir. Par conséquent on peut considérer de manière générale que la validation de méthode revient à évaluer l'incertitude de la mesure. Au fil des ans et à des fins de validation est née l'habitude qui consiste à représenter divers aspects de la performance de la méthode en faisant référence aux différents éléments énoncés ci-dessus; et d'ailleurs les recommandations présentes s'inscrivent dans cette tendance dans une large mesure. Cependant avec la confiance croissante portée sur l'incertitude de la mesure comme indicateur clé à la fois de l'aptitude-à l'essai et de la fiabilité des résultats; les analystes feront de plus en plus appel à la validation des mesures pour étayer leur évaluation de l'incertitude, et certains souhaiteront recourir à cette pratique immédiatement. Par conséquent l'incertitude de la mesure est abordée brièvement dans l'Annexe A comme une caractéristique de la performance d'une méthode d'analyse, tandis que l' Annexe B fournit des recommandations supplémentaires sur certaines procédures non abordées ailleurs.

4. Principes de base de la validation de methode

4.1. Spécification et portee de la validation

La validation s'applique selon un protocole bien défini, à la quantification, dans un cadre donné, d'un analyte spécifié associé à un domaine de concentration, se trouvant dans un matériau d'essai. D'une façon générale, la validation devrait vérifier que les performances de la méthode répondent aux besoins dans tout le domaine de concentration de l'analyte et pour tous les matériaux d'essai auxquels elle s'applique. Il s'en suit que ces données devraient être parfaitement précisées avant de procéder à une quelconque validation et être accompagnées d'une énonciation de tout critère d'aptitude-à l'essai

4.2. Hypothèses des essais

Les études de validation fournissent des données de performance qui renseignent sur

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

l'aptitude–à l'essai et qui prennent le pas aujourd'hui sur l'utilisation pratique faite des données de validation. De plus, les études de validation constituent un test objectif des hypothèses de départ sur lesquelles toute méthode d'analyse est fondée. Par exemple, si un résultat est obtenu à partir d'une simple droite d'étalonnage, on suppose implicitement que l'analyse est exempte de tout biais significatif, que la réponse est proportionnelle à la concentration de l'analyte et que la dispersion des erreurs aléatoires s'effectue de manière constante sur tout la domaine de concentration qui nous intéresse. Dans la plupart des cas, de telles hypothèses sont formulées sur la base de l'expérience accumulée tout au long du développement de la méthode, ou sur le plus long terme encore; elles sont par conséquent raisonnablement fiables. Néanmoins, les bonnes pratiques en matière de science de la mesure reposent sur des hypothèses *testées*. C'est pourquoi de très nombreuses études de validation se fondent sur des tests d'hypothèses statistiques, l'objectif étant de vérifier rapidement que les hypothèses raisonnables formulées concernant les principes de la méthode ne sont pas gravement erronées.

Cette remarque apparemment abstraite a une importante conséquence pratique. Il est plus facile de vérifier que l'on s'éloigne grandement d'une hypothèse fiable que de 'prouver' qu'une hypothèse donnée est correcte. Ainsi, lorsque l'on utilise depuis longtemps et avec succès une technique d'analyse particulière (par exemple l'analyse chromatographique en phase gazeuse ou encore les méthodes de digestion acide) pour un ensemble d'analytes et de matrices, les vérifications de validation se résument à juste titre à des tests relativement simples pour apporter des garanties sur la méthode. À l'inverse, lorsque l'on dispose de peu d'expérience, l'étude de validation doit apporter des preuves solides de l'adéquation des hypothèses retenues dans le cas particulier à l'étude; il est alors généralement nécessaire d'étudier tous les cas de figure dans le détail. Il s'en suit que l'ampleur requise d'une étude de validation dans un cas donné dépendra en partie de l'expérience accumulée concernant la technique d'analyse utilisée.

Dans l'exposé suivant, on considérera comme allant de soi que le laboratoire a une longue pratique de la technique concernée, et que l'objectif de tout test statistique de signification est de s'assurer qu'il n'existe aucune preuve solide allant à l'encontre des hypothèses sur lesquelles le protocole en question repose. Le lecteur se souviendra que des vérifications plus strictes peuvent se révéler nécessaires pour des techniques de mesure mal connues ou moins bien établies.

4.3. Sources d'Erreur en Analyse

Les erreurs de mesures en analyse proviennent de différentes sources[*] et à des

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

niveaux d'organisation différents. On peut utilement représenter ces différentes sources (pour une concentration précise de l'analyte) de la manière suivante[+24]:

- Erreurs aléatoires de la mesure (répétabilité);
- Erreur systématique (biais) de série;
- Erreur systématique (biais) de laboratoire;
- Erreur systématique (biais) de méthode;
- Effet de variation de la matrice.

Bien que ces différentes sources ne soient pas nécessairement indépendantes, cette liste constitue une manière utile de vérifier les domaines dans lesquels une étude de validation donnée doit chercher les sources d'erreur.

La répétabilité intra-série comprend des contributions de toutes les étapes de la procédure qui varient au cours d'une série, ceci comprend les contributions dues aux erreurs bien connues, gravimétriques et volumétriques, à l'hétérogénéité du matériau d'essai et à une variation dans les étapes de la chimie de l'analyse; on l'observe facilement à partir de la dispersion d'analyses répétées. L'effet de série explique les variations supplémentaires quotidiennes du système analytique, comme un changement de l'analyste, des lots, des réactifs, du réétalonnage des instruments, ainsi que de l'environnement du laboratoire (par ex. des changements de température). Lors d'une validation par un seul laboratoire, l'effet de série s'évalue habituellement en procédant à des répliques d'analyse d'un matériau approprié au cours de plusieurs séries distinctes. La variation inter-laboratoires s'explique par différents facteurs, tels que des variations dans les normes d'étalonnage, des différences dans les interprétations d'un protocole d'un endroit à l'autre, des changements dans l'équipement ou dans les réactifs utilisés, ou encore des facteurs environnementaux , par exemple des conditions climatiques moyennes différentes. La variation inter-laboratoires apparaît très clairement dans les résultats d'essais interlaboratoires (études de performance de la méthode) et de tests de compétence technique; parfois, on peut discerner la variation inter-méthode dans les résultats de ces derniers.

Généralement, la répétabilité, l'effet de série et l'effet de laboratoire sont d'amplitude comparable, on ne peut par conséquent en négliger aucun sans risque lors d'une validation. Dans le passé, on a eu tendance à négliger certains aspects, surtout quand il s'agissait d'évaluer et de faire état des informations concernant l'incertitude, avec pour conséquence des intervalles d'incertitude trop étroits. Par exemple, l'essai

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

interlaboratoires tel qu' il se pratique normalement ne donne pas l'ensemble des informations car la contribution sur l'incertitude du biais de la méthode et des variations de la matrice n'y sont pas pris en compte, et doivent donc être abordées par ailleurs (habituellement lors d'études internes préalables). La validation interne présente quant à elle un danger particulier: le biais de laboratoire risque d' être négligé , or cette source d'erreur est généralement le premier pourvoyeur d'incertitude de la liste ci-dessus. Il faut donc porter une attention toute particulière au biais de laboratoire lors d'une validation interne.

En plus des problèmes évoqués ci-dessus, la validation d'une méthode se limite à la portée de son application, c'est-à-dire à la méthode appliquée à une catégorie bien particulière de matériau d'essai. S'il existe des variations importantes du type de matrice à l'intérieur d'une catégorie définie, intervient alors une source supplémentaire de variation due aux effets de matrice intra-catégorie. Bien évidemment, si la méthode est ultérieurement utilisée pour des matériaux d'essai hors de la catégorie définie (c'est-à-dire hors de la portée de la validation), le système analytique ne peut pas être considéré comme étant validé: en effet une erreur supplémentaire d'ampleur inconnue est introduite dans le processus de mesure.

Il est également important que les analystes tiennent compte des variations de la performance d'une méthode en fonction de la concentration de l'analyte. Dans la plupart des cas la dispersion des résultats augmente avec la concentration et le recouvrement peut être très différent à forte et à faible concentrations. L'incertitude de la mesure associée aux résultats dépend par conséquent souvent de ces deux effets et d'autres facteurs dépendants de la concentration. Heureusement, il est souvent raisonnable de supposer qu'il existe une relation simple entre la performance et la concentration de l'analyte; couramment les erreurs sont proportionnelles à la concentration de l'analyte.[*] Cependant, dans les cas où l'on s'intéresse à la performance de la méthode à des concentration très différentes, il est important de vérifier la relation supposée entre la performance et la concentration de l'analyte. Cela se fait habituellement en vérifiant la performance aux extrémités du domaine d'application probable ou bien à quelques niveaux sélectionnés. Des contrôles de linéarité fournissent également des informations du même ordre.

4.4. Effets liés à la méthode et au laboratoire

Il est d'une importance critique lors d'une validation interne de méthode de tenir compte du biais de la méthode et du biais du laboratoire. Il existe quelques laboratoires aux équipements particuliers dans lesquels de tels biais peuvent être considérés comme étant négligeables, mais il s'agit là d'une situation tout à fait

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

exceptionnelle.(Cependant dans le cas où un seul laboratoire pratiquerait une analyse spécifique, le biais de la méthode et le biais du laboratoire revêtent une signification particulière). Généralement, les effets liés à la méthode et au laboratoire doivent être inclus dans le budget d'incertitude, mais souvent ils sont plus difficiles à traiter que l'erreur de répétabilité et que l'effet de série. En règle générale, pour pouvoir évaluer les incertitudes respectives, il est nécessaire d'utiliser des informations indépendantes du laboratoire. Généralement parlant, les sources d'informations indépendantes les plus utiles sont (i) les statistiques provenant d'essais inter-laboratoires (données souvent non disponibles dans les cas de validation interne de la méthode), (ii) de statistiques provenant d'essais d'aptitude technique, et (iii) des résultats provenant de l'analyse des matériaux de référence certifiés.

Les essais inter-laboratoires évaluent directement la variance des biais inter-laboratoires. Bien qu'il puisse exister des lacunes théoriques dans la conception de tels essais, ces estimations de variance sont adaptées à de nombreux usages pratiques. Par conséquent il est toujours instructif de tester les validations internes en comparant les estimations d'incertitude avec les estimations de reproductibilité provenant des essais inter-laboratoires. Si le résultat de la validation interne est nettement le plus petit, il est fort probable que des sources importantes d'incertitude ont été négligées. (Il peut arriver qu'un laboratoire particulier travaille en fait à un degré moindre d'incertitude que ce que l'on trouve dans les essais inter-laboratoires : dans ce cas, ce laboratoire devrait mettre en place des mesures spécifiques pour justifier de sa situation). Si aucun essai inter-laboratoires n'a été effectué sur la combinaison méthode/matériau d'essai considérée, une estimation de l'écart type de reproductibilité σ_H pour une concentration c de l'analyte supérieure à 120 ppb environ peut généralement être obtenue en utilisant la fonction de Horwitz, $\sigma_H = 0.02c^{0.8495}$, les deux variables étant exprimées comme des fractions de la masse. (L'estimation de Horwitz varie normalement du simple au double par rapport aux résultats observés lors d'études collaboratives). On a observé que la fonction de Horwitz est incorrecte pour des concentrations inférieures à 120 ppb environ, et dans ce cas une fonction modifiée est plus appropriée.[21, 25] Toutes ces informations peuvent être intégrées au domaine de la validation interne avec des changements minimes.

Les statistiques provenant des essais d'aptitude technique sont tout particulièrement intéressantes parce qu'elles fournissent des informations d'ordre général sur l'ampleur des biais combinés de laboratoire et de méthode; dans certains cas spécifiques, le participant obtient, quant à lui, des informations sur l'erreur totale. Des statistiques telles que l'écart type consolidé des résultats des participants, pour un analyte au cours d'un service d'essai inter-laboratoire peuvent en principe être utilisées de la

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

même façon que l'écart type de reproductibilité inter-laboratoire, par exemple, pour obtenir une valeur raisonnable de l'incertitude globale à comparer avec des estimations individuelles provenant de validations internes. En pratique il peut être plus difficile d'avoir accès aux statistiques provenant des essais d'aptitude technique car elles ne sont pas systématiquement répertoriées et publiées, comme c'est le cas lors d'essais inter-laboratoires (souvent elles ne sont disponibles que pour les participants). Bien sûr, si l'on doit utiliser de telles statistiques il faut qu'elles fassent référence à la matrice et à la concentration de l'analyte appropriées. Les différents participants à des programmes d'essais d'aptitude technique peuvent également évaluer la validité de leur estimation de l'incertitude en comparant leurs résultats avec les valeurs assignées de plusieurs séries successives[26]. Cependant il s'agit là d'une activité continue, qui par conséquent ne ressort pas stricto sensu du domaine de compétence de la validation interne (qui est un événement singulier).

Si un matériau de référence certifié (MRC) est disponible, un test interne permet au laboratoire d'évaluer les biais combinés du laboratoire et de la méthode en analysant le MRC un certain nombre de fois. L'estimation du biais combiné est la différence entre la moyenne des résultats et la valeur certifiée.

Les matériaux de référence certifiés adéquats ne sont pas toujours disponibles, aussi peut-on se trouver dans l'obligation d'utiliser d'autres matériaux. Parfois des matériaux restant après des essais d'aptitude technique sont utilisés à cet effet et, même si les grandeurs assignées à ces matériaux peuvent présenter des incertitudes sur les valeurs desquelles on peut s'interroger, leur utilisation permet cependant une vérification du biais global. De façon spécifique, les valeurs assignées aux essais d'aptitude technique sont généralement choisies pour fournir une estimation avec le minimum de biais, par conséquent un test du biais à partir d'un tel matériau est une pratique tout à fait raisonnable. Il existe aussi la possibilité d'utiliser des informations à partir de tests d'ajouts dosés et de recouvrement,[4] pour fournir des estimations de ces biais, même s'il peut y avoir des sources d'incertitude non mesurées par ces techniques.

Couramment l'effet le moins reconnu en matière de validation est celui dû à la variation des matrices à l'intérieur de la catégorie de matériau d'essai défini. L'exigence théorique pour pouvoir estimer ce composant de l'incertitude serait d'analyser en une seule fois un ensemble représentatif des matériaux d'essai, en évaluant leurs biais individuels et en calculant la variance de ces biais. (Le fait d'analyser en une seule fois signifie que des biais de plus hauts niveaux n'ont pas d'incidence sur la variance. Si un large intervalle de concentration est concerné, il faut

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

prendre en considération l'évolution du biais en fonction de la concentration).

Si les matériaux représentatifs sont des matériaux de référence certifiés, les biais peuvent être directement évalués comme étant les différences entre les résultats et les valeurs de référence, et toute la procédure est directe. Dans le cas plus probable où le nombre de matériaux de référence disponibles est insuffisant, on doit parfois faire appel à des tests de recouvrement avec une gamme de matériaux d'essais types, en procédant avec tout le soin que cela mérite. Actuellement il existe fort peu de données quantitatives quant à l'ordre de grandeur des incertitudes provenant de cette source, même si dans certains cas on suspecte qu'elles sont importantes.

5. Conduite des etudes de Validation

La conception détaillée et la réalisation des études de validation de méthode sont des sujets traités de façon approfondie dans d'autres documents, et ils ne seront donc pas repris ici. Cependant, les principes les plus importants qui sont pertinents, sont pris en considération ci-dessous:

Il est essentiel que des études de validation soient représentatives. C'est-à-dire que, dans la mesure du possible, les études devraient être conduites pour fournir une représentation réaliste du nombre et et de l'étendue des effets intervenants lors de l'utilisation normale de la méthode, et conduite également pour couvrir les domaines de concentration et les types d'échantillons inclus dans le champ d'application de la méthode. Lorsqu'un facteur (ex : la température ambiante) a varié significativement et de manière aléatoire au cours d'une étude de fidélité, par exemple, les effets de ce facteur apparaissent directement dans la variance observée et ne nécessitent pas d'étude supplémentaire, sauf si une optimisation de la méthode est souhaitable par la suite.

Dans le contexte de la validation de méthode, la "variation représentative" signifie que le facteur observé doit présenter une distribution de valeurs en accord avec le domaine de variation attendu du paramètre en question. Pour des paramètres mesurables en continu, il peut s'agir d'un domaine de variation autorisé, d'une incertitude déclarée ou d'un domaine de variation attendu ; pour des facteurs discontinus, ou pour des facteurs avec des effets imprévisibles comme la matrice d'échantillon, un intervalle représentatif correspond à la diversité des types ou "niveaux de facteur" autorisés ou rencontrés lors de l'utilisation normale de la méthode. Idéalement, la représentativité ne s'applique non pas seulement à la fourchette de valeurs mais aussi à leur distribution. Malheureusement, il est souvent trop onéreux de mettre en place la variation totale de nombreux facteurs à de

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

nombreux niveaux. Cependant, dans la plupart des cas pratiques, le minimum acceptable est de mettre en place des essais basés sur les points extrêmes de l'intervalle attendu, ou sur un intervalle plus large que celui prévu dans la méthode.

Lorsque l'on choisit des facteurs de variation, il est important de s'assurer que les effets les plus importants sont le plus étudiés possible. Par exemple, quand la variation d'un jour à l'autre (pouvant être due aux effets de réétalonnage) est importante comparée à la répétabilité, deux déterminations par jour pendant cinq jours fourniront une meilleure estimation de la fidélité intermédiaire que cinq déterminations par jour pendant seulement deux jours. Dix déterminations individuelles sur plusieurs jours distincts seraient préférables encore, à condition d'exercer un contrôle suffisant, même si cela ne fournit pas d'information supplémentaire sur la répétabilité observée au cours d'une même journée.

Il est évident que lorsque l'on prévoit des tests d'hypothèse, toute étude devrait être capable de détecter de tels effets avant qu'ils ne deviennent importants pratiquement parlant (c'est-à-dire comparables à la plus grande composante de l'incertitude).

De plus, les considérations suivantes peuvent se révéler importantes:

- Lorsque l'on sait ou suspecte que des facteurs interagissent entre eux, il est important de s'assurer que l'effet de l'interaction est pris en compte. Pour ce faire, on peut soit procéder à une sélection aléatoire à différents niveaux des paramètres qui interagissent, soit concevoir avec une attention particulière un système adapté à l'obtention des effets d''interaction' ou des informations sur la covariance.
- Lorsque l'on procède à des études du biais global, il est important que les matériaux de référence et les valeurs soient en rapport avec les matériaux soumis aux essais de routine.

6. Portee des etudes de validation

La portée souhaitable de la validation par un laboratoire donné d'une méthode nouvelle, modifiée ou mal connue, dépend en partie du statut actuel de la méthode et de la compétence du laboratoire. Ci-dessous nous émettons des suggestions sur la portée de la validation et des mesures de vérification dans diverses situations. Nous supposons dans l'ensemble que la méthode en question est destinée à une utilisation de routine, sauf exception clairement indiquée.

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

6.1. Le laboratoire est tenu d'utiliser une methode "entierement" validee

La méthode a été étudiée lors d'un essai interlaboratoires et le laboratoire se doit donc de vérifier qu'il est en mesure d'obtenir les caractéristiques de performance de la méthode ayant été publiées (ou sinon qu'il est capable de répondre aux exigences de la tâche analytique). Le laboratoire devrait entreprendre des études de fidélité, des études de biais (y compris des études de variation de la matrice), et éventuellement des études de linéarité, certains essais néanmoins, et par exemple celui de la robustesse pouvant être laissés de côté.

6.2. Le laboratoire doit utiliser une methode entierement validee mais avec une nouvelle matrice

La méthode a été étudiée lors d'un essai interlaboratoires et le laboratoire se doit donc de vérifier que la nouvelle matrice n'introduit pas de nouvelles sources d'erreur dans le système. La même portée de l'étude de validation que précédemment est requise.

6.3. Le laboratoire doit utiliser une methode bien etablie mais qui n'a pas fait l'objet d'une etude collaborative

La même portée de l'étude de validation que précédemment est requise.

6.4. La methode a fait l'objet d'une publication dans la litterature scientifique celle-ci fournissant quelques caracteristiques analytiques

Le laboratoire devrait entreprendre des études de fidélité, des études de biais (incluant des études de variation de la matrice), ainsi que des études de robustesse et de linéarité.

6.5. La methode a fait l'objet d'une publication dans la litterature scientifique, mais sans precision sur les caracteristiques, ou bien la methode a été developpee en interne

Le laboratoire devrait entreprendre des études de fidélité, des études de biais (incluant des études de variation de la matrice), ainsi que des études de robustesse et de linéarité.

6.6. La methode est empirique

Dans une méthode empirique la quantité estimée est tout simplement le résultat obtenu en suivant la procédure spécifiée. Elle diffère des mesurages visant à évaluer

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

des grandeurs indépendantes de la méthode, comme la concentration d'un analyte donné dans un échantillon, dans le fait que le biais de la méthode est conventionnellement nul et que la variation de la matrice (dans la catégorie définie) n'est pas à considérer. Le biais de laboratoire quant à lui ne peut pas être ignoré, mais il a tendance à être difficilement évalué au moyen d'une étude intra-laboratoire. De plus, les matériaux de référence sont rarement disponibles. En l'absence de données provenant d'étude collaborative, on peut évaluer la fidélité inter laboratoire en utilisant une étude de robustesse spécialement conçue à cet effet ou en utilisant la fonction de Horwitz.

6.7. L'analyse est "ad hoc"

On a parfois besoin de recourir à une analyse "ad hoc" pour estimer une valeur de façon peu précise à moindre frais et sans exigence poussée. L'effort dispensé à des fins de validation est par conséquent strictement limité. Les biais devraient être étudiés par des méthodes comme l'estimation du recouvrement ou des ajouts dosés, de plus la fidélité peut être étudiée par réplique de plusieurs mesures.

6.8. Changements de personnel et d' equipement

On peut donner des exemples importants suivants: changements concernant les principaux instruments; nouveaux lots de réactifs pouvant fortement varier (par exemple les anticorps polyclonaux); changements touchant les locaux du laboratoire; méthodes utilisées pour la première fois par un nouveau personnel ; ou bien une méthode validée, employée après une période de mise en sommeil. Il s'agit alors et avant tout de prouver qu'aucune modification délétère n'est survenue. La vérification minimale requise est une simple étude de biais basée sur une étude de type "avant et après" sur des matériaux d'essai types ou sur des matériaux de contrôle. En général, les tests effectués devraient mettre en évidence l'impact possible du changement sur la procédure analytique.

7. Recommandations

Les recommandations suivantes sont formulées pour l'utilisation de la validation intra-laboratoire de la méthode:

- A chaque fois que cela est possible et pertinent, un laboratoire devrait utiliser une méthode d'analyse dont les caractéristiques de performance ont été évaluées lors d'un essai interlaboratoires conforme à un protocole international.

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

- Quand de telles méthodes ne sont pas disponibles, une méthode doit être validée en interne avant d'être utilisée pour produire des données analytiques pour un client.
- Pour une validation interne, il faut que le laboratoire choisisse des caractéristiques d'évaluation adéquates dans la liste suivante: domaine d'application, spécificité, étalonnage, justesse, fidélité, domaine de validité, limite de quantification, limite de détection, sensibilité, robustesse et adéquation à l'usage prévu. Le laboratoire doit tenir compte des besoins du client lors du choix des caractéristiques à déterminer.
- On doit fournir aux clients du laboratoire la preuve que ces caractéristiques ont bien été évaluées s'ils en font la demande.

8. Références Bibliographiques

1. "Protocol for the Design, Conduct and Interpretation of Method Performance Studies", W Horwitz, Pure Appl. Chem., 1988, **60**, 855 864, révisé W. Horwitz, Pure Appl. Chem., 1995, **67**, 331-343.
2. "The International Harmonised Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories", M Thompson et R Wood, Pure Appl. Chem., 1993, **65**, 2123-2144. (également publié dans J. AOAC International, 1993, **76**, 926-940.
3. "Harmonised Guidelines For Internal Quality Control in Analytical Chemistry Laboratories", Michael Thompson et Roger Wood, J. Pure & Applied Chemistry, 1995, 67(4), 49-56.
4. "Harmonised Guidelines for the Use of Recovery Information in Analytical Measurement", Michael Thompson, Stephen Ellison, Ales Fajgelj, Paul Willetts et Roger Wood, J. Pure & Applied Chemistry, 1999, 71(2), 337-348.
5. "Council Directive 93/99/EEC on the Subject of Additional Measures Concerning the Official Control of Foodstuffs", J. O., 1993, L290.
6. "Procedural Manual of the Codex Alimentarius Commission, 10th Edition", FAO, Rome, 1997.
7. "*Precision of Test Methods*", Genève, 1994, ISO 5725, Editions précédentes en 1981 et 1986.

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

8. *"Guide to the Expression of Uncertainty in Measurement"*, ISO, Genève, 1993.
9. *"Quantifying Uncertainty in Analytical Measurement"*, EURACHEM Secretariat, Laboratory of the Government Chemist, Teddington, Royaume-Uni, 1995, EURACHEM Guide (en cours de révision).
10. *"International vocabulary of basic and general terms in metrology"* ISO, Genève 1993
11. *"Validation of Chemical Analytical Methods"*, NMKL Secretariat, Finlande, 1996, NMKL Procedure No. 4.
12. *"EURACHEM Guide: The fitness for purpose of analytical methods. A Laboratory Guide to method validation and related topics"*, LGC, Teddington 1996. Egalement disponible auprès du Secrétariat de EURACHEM et sur leur site web.
13. *"Statistics manual of the AOAC"*, AOAC INTERNATIONAL, Gaithersburg, Maryland, USA, 1975
14. *"An Interlaboratory Analytical Method Validation Short Course developed by the AOAC INTERNATIONAL"*, AOAC INTERNATIONAL, Gaithersburg, Maryland, USA, 1996.
15. *"Text on validation of analytical procedures"* International Conference on Harmonisation. Federal Register, Vol. 60, 1 Mars 1, 1995, pages 11260
16. *"Validation of analytical procedures: Methodology"* Conférence Internationale sur l' Harmonisation. Registre Fédéral Federal Register, Vol. 62, No. 96, 19 Mai 1997, pages 27463-27467.
17. *"Validation of Methods"*, Inspectorate for Health Protection, Rijswijk, Pays-Bas, Rapport 95-001.
18. *"A Protocol for Analytical Quality Assurance in Public Analysts' Laboratories"*, Association of Public Analysts, 342 Coleford Road, Sheffield S9 5PH, Royaume-Uni, 1986.
19. *"Validation of Analytical Methods for Food Control"*, Rapport d'une Consultation Conjointe d'Experts FAO/IAEA, Décembre 1997, FAO Food and Nutrition Paper No. 68, FAO, Rome, 1998
20. *"Estimation and Expression of Measurement Uncertainty in Chemical Analysis"*, Secrétariat NMKL, Finlande, 1997, Procédure NMKL No. 5.

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

21. M Thompson, PJ Lowthian, J AOAC Int, 1997, 80, 676-679
22. Recommandation IUPAC: “Nomenclature in evaluation of analytical methods, including quantification and detection capabilities” Pure and Applied Chem.” 1995, 67 1699-1723
23. ISO 11843. “Capability of detection.” (Plusieurs parties). International Standards Organisation, Genève.
24. M. Thompson, Analyst, 2000, 125, 2020-2025
25. “Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing” M Thompson, Analyst, 2000, 125, 385-386.
26. “How to combine proficiency test results with your own uncertainty estimate - the zeta score”, Analytical Methods Committee of the Royal Society of Chemistry, AMC Technical Briefs, publié par M. Thompson, AMC Technical Brief No. 2, www.rsc.org/lap/rsccom/amc

ANNEXE A: Notes sur les conditions requises pour l'etude des caracteristiques de performance d’une methode

Les conditions générales requises pour les caractéristiques individuelles de performance d'une méthode sont les suivantes.

A1 Domaine d’application

Après validation, la documentation devrait fournir, en plus de toute spécification de la performance, les informations suivantes:

- L'identification de l'analyte, y compris la spécification si nécessaire (Exemple: ‘arsenic total’);
- Le domaine de concentration couvert par la validation (Exemple: ‘0-50 ppm’);
- La spécification des types de matrices du matériau d’essai couverte par la validation (Exemple: ‘vin rouge sec’);
- Un protocole décrivant l'équipement, les réactifs, la procédure (y compris la variation autorisée dans les instructions spécifiées, par ex. ‘chauffer à 100 □ 5□

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

pendant 30 □ 5 minutes'), l'étalonnage et les procédures qualité, ainsi que toute mesure de précaution ou de sécurité éventuellement nécessaire;

- L'application prévue et ses exigences critiques d'incertitude (Exemple: 'L'analyse alimentaire à des fins de sélection rapide. L'incertitude normale (standard) $u(c)$ du résultat c devrait être inférieure à 0.1□c.').

A2 Spécificité

La spécificité se définit comme étant le niveau de capacité d'une méthode à quantifier l'analyte exactement en présence d'interférents. Dans l'idéal, on devrait évaluer la spécificité pour chaque interférent important susceptible d'être présent. Il est particulièrement important de vérifier les interférents susceptibles, sur la base de principes chimiques, de donner une réponse à l'essai. Par exemple, on peut raisonnablement s'attendre à ce que des réactions colorimétriques pour l'ammoniac répondent aux amines aliphatiques primaires. Il peut se révéler peu ou pas pratique de tenir compte et d'évaluer chaque interférent potentiel; dans ce cas, on recommande d'envisager les situations les plus défavorables susceptibles de se produire. Le principe général est que la spécificité devrait être suffisamment bonne pour ne pas avoir à se préoccuper des interférences.

Dans de nombreux types d'analyses, la spécificité est essentiellement une évaluation qualitative fondée sur la réponse, significative ou non, des tests appropriés pour les études des interférences. Cependant, il existe des mesures quantitatives utiles. Par exemple l'indice de spécificité b_{an}/b_{int}, où b_{an} représente la sensibilité de la méthode (pente de la droite d'étalonnage) et où b_{int} représente la pente de la réponse produite indépendamment par un interférent potentiel; cet indice fournit une mesure quantitative de l'interférence. On peut déterminer b_{int} de façon approximative en exécutant la procédure sur un blanc matrice et sur le même blanc supplémenté avec l'interférent potentiel à une concentration donnée adaptée. Si aucun blanc matrice n'est disponible et qu'on utilise à la place un matériau type, on ne peut estimer b_{int} à partir d'une expérience aussi simple que si l'on supposait que les effets mutuels de matrice sont absents. Notons qu'il est plus facile de déterminer b_{int} en l'absence de l'analyte, car il existe le risque de confondre son effet avec un autre genre d'interférence lorsque la sensibilité du dosage de l'analyte est elle même affectée par l'interférent (effet matrice).

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

A.3. Etalonnage et linearite

Si l'on exclut les erreurs grossières de préparation des étalons, les erreurs d'étalonnage constituent généralement (mais pas toujours) une composante mineure du budget total de l'incertitude, et l'on peut en général raisonnablement les répartir dans diverses catégories que l'on estime par des méthodes dites "descendantes". Par exemple, les erreurs aléatoires résultant de l'étalonnage font partie du biais de série, qui est évalué de manière globale, tandis que des erreurs systématiques provenant de cette source peuvent se retrouver dans le biais du laboratoire, évalués eux aussi de manière globale. Néanmoins, il est utile de connaître certaines caractéristiques de l'étalonnage dès le départ de la validation de la méthode, car elles ont une incidence sur la stratégie à adopter pour la réalisation optimale de la procédure. Dans cette catégorie il est question de savoir s'il est plausible que la fonction d'étalonnage (a) soit linéaire, (b) passe par l'origine et (c) ne soit pas affectée par la matrice du matériau d'essai. Les procédures décrites dans le présent document s'appliquent aux études d'étalonnage dans le cadre d'une validation, celles-ci sont nécessairement plus exigeantes que celles appliquées à l'étalonnage réalisé lors d'analyses de routine. Par exemple, une fois que l'on a établi lors de la validation que la fonction d'étalonnage est linéaire et passe par l'origine, on peut instaurer une stratégie d'étalonnage beaucoup plus simple au moment de l'utilisation en routine (par exemple, un plan d'expérience avec deux points répétés). Normallement, les erreurs provenant de cette méthode simplifiée d'étalonnage se répartissent dans le cadre de la validation, dans des niveaux d'erreurs supérieurs.

A3.1 Linéarité et ordonnée à l'origine

On peut tester la linéarité de manière informelle en examinant la représentation graphique des erreurs résiduelles produites par la régression linéaire des réponses sur les concentrations dans un ensemble d'étalonnage approprié. Tout graphe en forme de courbe suggère un manque d'ajustement dû à une fonction d'étalonnage non linéaire. Un test statistique de signification peut être entrepris en comparant la variance de l'erreur d'ajustement avec celle due à l'erreur pure. Cependant il existe d'autres causes d'erreur d'ajustement que la non linéarité, et qui se produisent lors de certains types d'étalonnages analytiques ; il faut donc utiliser le test statistique de signification en conjonction avec la représentation graphique des erreurs résiduelles. Bien que largement utilisé comme indicateur de la qualité de l'ajustement, le coefficient de corrélation porte à confusion; il n'est pas approprié comme test de linéarité et ne devrait pas être utilisé.

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

Le plan d'expérience est de la plus grande importance pour les études d'erreur d'ajustement, car il est facile de confondre la non linéarité avec la dérive. Des mesures répétées sont nécessaires pour fournir une estimation de l'erreur pure s'il n'y a pas d'estimation indépendante. En l'absence de spécifications, les consignes suivantes devraient être appliquées:

- Il devrait y avoir au moins six calibrants;
- Les calibrants devraient être placés à intervalles réguliers sur toute la gamme de concentration étudiée;
- La gamme devrait englober 0-150% ou 50-150% de la concentration susceptible d'être rencontrée, selon le pourcentage le plus approprié;
- Les calibrants devraient être analysés deux fois au minimum et de préférence trois fois ou plus, dans un ordre aléatoire.

Après un premier essai d'ajustement à partir d'une régression linéaire simple, les erreurs résiduelles devraient être examinées pour trouver les modèles adaptés. L' hétéroscédasticité est assez fréquente dans l'étalonnage analytique et un modèle la suggérant signifie que les données de l'étalonnage doivent de préférence être traitées au moyen de la régression pondérée. Le non recours à la régression pondérée dans ces circonstances pourrait occasionner des erreurs exagérées dans la partie inférieure de la courbe d'étalonnage.

L'étude de l'erreur d'ajustement peut être effectuée en utilisant soit la régression simple soit la régression pondérée. Une étude de l'ordonnée à l'origine testant si elle est significativement différente de zéro peut également être effectué sur cette donnée s'il n'y a pas d'erreur d'ajustement significative.

A3.2 Etude de l'effet matrice

L'étalonage se trouve grandement simplifié si les étalons peuvent être préparés à partir d'une solution simple de l'analyte. Les effets d'une possible erreur de matrice doivent être évalués lors de la validation si cette stratégie est adoptée. On peut tester les effets matrice en appliquant la méthode des ajouts d'analyte (appelée aussi "ajouts de standard") à une solution d'essai dérivée d'un matériau d'essai type. Il faudrait effectuer l'essai de façon à obtenir la même dilution finale que la procédure normale, et la gamme des ajouts devrait recouvrir la même fourchette que celle de la validation de l'étalonnage définie par la procédure. Si l'étalonnage est linéaire, la pente de la fonction habituelle d'étalonnage peut être comparée à la pente de la représentation

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

graphique des ajouts dosés, pour rechercher une différence significative. L'absence de signification indique qu'il n'existe pas d'effet matrice détectable. Si l'étalonnage n'est pas linéaire on a besoin d'une méthode plus complexe avec un test statistique de signification, mais une comparaison visuelle à des concentrations égales pourra généralement suffire. L'absence de signification dans cet essai indique souvent que l'effet de variation de la matrice [Section A13] est aussi absent.

A3.3 Procédure finale d'étalonnage

On peut être amené à valider séparément la stratégie d'étalonnage telle qu'elle est définie dans la procédure, bien que les erreurs impliquées contribuent à des incertitudes estimées conjointement. Il est important de noter ici que les incertitudes estimées à partir des plans d'expérience spécifiques pour la linéarité etc,.. seront plus petites que celles issues de l'étalonnage simple, défini dans le protocole de la procédure

A4 Justesse

A4.1 Evaluation de la justesse

La justesse est le degré de concordance entre le résultat d'un essai et la valeur de référence acceptée du paramètre mesuré. La justesse s'exprime quantitativement en termes de "biais"; plus le biais est petit, plus la justesse est grande. Le biais est typiquement déterminé en comparant la réponse de la méthode à un matériau de référence présentant une valeur connue affectée. Le test statistique de signification est recommandé. Quand l'incertitude de la valeur de référence n'est pas négligeable, l'évaluation des résultats devrait tenir compte de l'incertitude liée au matériau de référence ainsi que de la variabilité statistique.

A4.2 Conditions à appliquer pour les expériences de justesse

Des biais peuvent apparaître à différents niveaux d'organisation dans un système analytique, par exemple, le biais de série, le biais de laboratoire et le biais de la méthode. Il est important de savoir quel biais est traité par les différentes méthodes d'étude des biais. En particulier:

- La moyenne d'une série de mesurages d'un matériau de référence, réalisés ensemble au cours de la même série, donne des informations sur la somme des effets de méthode, de laboratoire et de série pour cette série en particulier. Puisque l'on suppose que l'effet de série est aléatoire d'une série à l'autre, la

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

variation du résultat d'une série à l'autre sera plus importante que celle que l'on pourrait attendre à partir de la dispersion observable des résultats; il faut en tenir compte dans l'évaluation des résultats (par exemple, en comparant le biais mesuré avec l'écart type entre séries déterminé par ailleurs).

- La moyenne des analyses d'un matériau de référence, répétées sur plusieurs séries permet d'estimer les effets combinés des biais produits par la méthode et par le laboratoire pour un laboratoire donné (sauf quand la valeur est attribuée en utilisant la méthode en question).

A4.3 Valeurs de référence à utiliser pour les expériences de justesse

A4.3.1 Matériaux de Référence Certifiés (MRC)

Les MRC sont raccordés aux étalons internationaux, avec une incertitude connue, et peuvent, par conséquent, être utilisés pour aborder simultanément tous les aspects des biais (méthode, laboratoire, et intra-laboratoire), en supposant qu'il n'y a pas d'erreur lié à la matrice. En conséquence, on devrait faire usage des MRC pour la validation de la justesse, lorsque cela est possible en pratique. Il faut s'assurer que les incertitudes des valeurs certifiées sont suffisamment petites pour permettre la détection d'un biais de grande ampleur. Quand elles ne le sont pas, on recommande malgré tout l'utilisation des MRC, mais des vérifications supplémentaires doivent être effectuées.

Une expérience de justesse type engendre une réponse moyenne pour un matériau de référence. Au moment de l'interprétation du résultat , il faut prendre en considération l'incertitude associée à la valeur certifiée ainsi que l'incertitude due à la variation statistique dans le laboratoire. Ce dernier élément peut être fondé sur l'écart type intra-série, inter-séries, ou sur une estimation de l'écart type inter-laboratoires; tout dépendra alors du but de l'expérience (statistique ou matériaux). Quand l'incertitude de valeur certifiée est petite, un test de Student- t est généralement pratiqué, en utilisant le niveau de confiance approprié.

Lorsque cela se révèle nécessaire et possible en pratique, on devrait examiner un certain nombre de MRC appropriés, avec les matrices et les concentrations d'analyte adéquates. Quand c'est le cas, et que les incertitudes des valeurs certifiées sont inférieures à celles des résultats d'analyse, il serait raisonnablement bon d'utiliser la régression simple pour évaluer les résultats. Ainsi on pourrait exprimer le biais en fonction de la concentration et il pourrait se révéler comme étant une ordonnée à l'origine différente de zéro ("biais transitionnel ou constant") ou comme étant une

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

pente différente de 1 (biais “de rotation” or proportionnel). Il faut faire preuve de prudence lors de l'interprétation des résultats lorsque la gamme de matrices est importante.

A.4.3.2 Matériaux de référence :

Quand des *Matériaux de Référence Certifiés* (MRC) ne sont pas disponibles, ou en complément des MCR, on peut faire usage de tout matériau suffisamment caractérisé pour le besoin présent (un matériau de référence[10]), en ne perdant jamais de vue que, si un biais insignifiant n'est pas une preuve de l'absence de biais, en revanche un biais significatif sur quelque matériau que ce soit doit être étudié de façon approfondie. Les matériaux de référence comprennent par exemple: des matériaux caractérisés par un fournisseur de matériaux de référence, mais dont les valeurs ne sont pas accompagnées par des données d'incertitude ou sont qualifiées autrement; des matériaux caractérisés par un fabricant du matériau; des matériaux caractérisés dans le laboratoire pour être utilisés comme matériaux de référence; des matériaux soumis à un essai comparatif restreint, ou distribués lors d'un test d’aptitude technique. Même si la traçabilité de ces matériaux n’est pas parfaite, il est nettement préférable de les utiliser plutôt que de ne procéder à aucune évaluation des biais. Ces matériaux seront utilisés autant que possible de la même manière que des MRC, cependant, en l'absence de données d'incertitude, tout test statistique de signification reposera entièrement sur la fidélité observable des résultats.

A4.3.3 Utilisation d'une méthode de référence

On peut en principe utiliser une méthode de référence pour tester les biais d’une autre méthode en cours de validation. C'est une possibilité très utile quand on veut vérifier une alternative à /ou une modification d'une méthode courante établie déjà validée et utilisée par le laboratoire. Les deux méthodes sont utilisées pour analyser un certain nombre de matériaux d’essais types, recouvrant de façon bien répartie de préférence, un domaine de concentration pertinent. Une comparaison des résultats sur le domaine à l'aide d'une méthode statistique appropriée (par exemple un test-*t bilatéral*, avec les vérifications nécessaires d'homogénéité de variance et de normalité) mettrait en évidence tout biais entre les méthodes.

A4.3.4 Utilisation des ajouts dosés / recouvrement

En l'absence de matériaux de référence, ou pour étayer les études à partir de matériaux de référence, le biais peut être étudié grâce à des ajouts dosés et à l’étude de recouvrement. Un matériau d’essai type est analysé par la méthode à valider à la

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

fois dans son état d'origine et après ajout dosé d'une masse connue de l'analyte. La différence entre les deux résultats, exprimée en proportion de la masse ajoutée est appelée recouvrement de substitution ou parfois recouvrement marginal. Des recouvrements sensiblement différents de 1 indiquent qu'un biais affecte la méthode. Au sens strict, les études de recouvrement ici décrites n'évaluent que le biais dû aux effets s'exerçant sur l'analyte ajouté; les mêmes effets ne s'appliquent pas nécessairement dans la même mesure à l'analyte natif, et des effets supplémentaires peuvent s'appliquer sur lui. Les études d'ajouts dosés et de recouvrements sont par conséquent fortement sujettes à l'observation suivante: même si un bon recouvrement n'est pas une garantie de justesse, un mauvais recouvrement est certainement une indication d'un manque de justesse. Des méthodes concernant la manipulation des données de l'ajout dosé/recouvrement ont été traitées en détail ailleurs.[4]

A5 Fidelite

La fidélité est l'étroitesse d'accord entre des résultats d'essais indépendants obtenus dans des conditions bien stipulées. Elle est généralement définie en termes d'écart type ou de coefficient de variation. La distinction entre la fidélité et le biais est fondamentale, mais dépend du niveau auquel le système analytique est considéré. Ainsi, du point de vue d'une détermination donnée, toute déviation affectant l'étalonnage de la série serait perçue comme un biais. Du point de vue de l'analyste passant en revue une année de travail, le biais de série sera différent chaque jour et agira comme une variable aléatoire avec une fidélité associée. Les conditions stipulées pour l'évaluation de la fidélité tiennent compte de cette différence de point de vue.

Pour la validation interne, deux types de conditions sont pertinents: (a) la fidélité dans des conditions de répétabilité, décrivant les variations observées pendant une seule série d'analyse présentant une espérance statistique de 0 et un écart type σ_r, et (b) la fidélité inter-séries décrivant les variations du biais de série $\square_{run}$ avec une espérance de 0 et un écart type σ_{run}. Généralement ces deux sources d'erreurs affectent les résultats d'analyse individuels, qui par conséquent ont une fidélité combinée $\sigma_{tot} = (\sigma_r^2/n + \sigma_{run}^2)^{1/2}$, avec n nombre de résultats répétés obtenus à l'intérieur d'une même série exprimés sous la forme de leur moyenne dans les rapports d'étude. Les deux estimations de la fidélité peuvent être obtenues plus simplement en analysant le matériau d'essai sélectionné en double pendant un certain nombre de séries successives. Les divers composants de la variance peuvent alors être calculés en appliquant une analyse unidirectionnelle de la variance. Chaque analyse dupliquée doit être issue d'une mise en oeuvre de la procédure de façon indépendante, que l'on

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

applique à un échantillon indépendant du matériau d'essai. Par une autre approche, la fidélité combinée σ_{tot} peut être directement évaluée en procédant à l'analyse du matériau d'essai une seule fois au cours de séries successives, et en estimant l'écart type à partir de l'équation classique. (Notons que les écart types observés portent généralement le symbole *s*, pour les distinguer des écarts types vrais □).

Il est important que les valeurs de fidélité soient représentatives des conditions d'essai probables. Premièrement, la variation dans les conditions au cours des séries doit refléter ce qui se passerait normalement dans le laboratoire dans les conditions de l'utilisation en routine de la méthode. Par exemple, des variations dans les lots de réactifs, dans les analystes et les instruments devraient être représentatives. Deuxièmement, le matériau d'essai employé devrait être typique, en termes de matrice et (idéalement) d'état de fractionnement, des matériaux susceptibles d'être rencontrés dans l'application en routine. Il en résulte que des matériaux d'essai vrais ou, à un degré moindre, des matériaux de référence correspondants à la matrice seraient adéquats, mais les solutions synthétiques ne le seraient pas. Notons également que les MRC et les matériaux de référence préparés sont fréquemment beaucoup plus homogènes que des matériaux d'essai typiques, et la fidélité calculée à partir de leur analyse peut par conséquent sous-estimer la variation que l'on observera pour les matériaux d'essai.

La fidélité varie souvent en fonction de la concentration de l'analyte. On suppose généralement que i) il n'y a pas de changement de la fidélité en fonction du niveau de l'analyte, ou ii) l'écart type est proportionnel à, ou dépend linéairement du niveau de l'analyte. Dans les deux cas, il faut vérifier la supposition de départ si l'on s'attend à ce que le niveau de l'analyte varie de façon substantielle (c'est-à-dire de plus de 30% environ par rapport à sa valeur centrale). L'expérience la plus économique est probablement de procéder à une simple évaluation de la fidélité au niveau des, ou proche des, points extrêmes du domaine d'application, en l'accompagnant d'un test statistique approprié appliqué sur la différence de variance. Le test F est adapté lorsque l'erreur suit une distribution normale.

Des données sur la fidélité peuvent être obtenues à partir d'un grand nombre d'ensembles de conditions différents, en plus des conditions les plus simples qui sont celles de la répétabilité et celles inter-séries, stipulées ici, il peut être pertinent d'acquérir des informations supplémentaires. Par exemple, il peut être utile, pour l'évaluation des résultats, ou pour améliorer le mesurage, d'avoir des indications séparées sur les effets des séries et des opérateurs, d'un jour à l'autre ou au cours d'une même journée ou sur la fidélité à laquelle on peut prétendre en utilisant un ou plusieurs instruments. Une gamme de différents plans d'expérience et de différentes

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

techniques d'analyse statistique est disponible, et l'on recommande fortement de toujours bien choisir son plan d'expérience pour de telles études.

A6 RECOUVREMENT

Des méthodes d'évaluation du recouvrement sont traitées dans le chapitre consacré aux méthodes d'évaluation de la justesse (ci-dessus).

A7 Domaine de validité

Le domaine de validité est l'intervalle de concentration de l'analyte à l'intérieur duquel la méthode peut être considérée comme étant valide. Il est important de bien comprendre que ce domaine n'est pas nécessairement identique au domaine utile de l'étalonnage. Tandis que l'étude d'étalonnage peut recouvrir un large domaine de concentration, le restant de la validation (qui représente généralement une partie beaucoup plus importante en termes d'incertitude) recouvrira un domaine plus restreint. Dans la pratique, la plupart des méthodes ne seront validées qu'à un ou deux niveaux de concentration. On peut considérer le domaine de validité comme une extrapolation raisonnable de ces points sur l'échelle de concentration.

Quand l'utilisation de la méthode se focalise sur une concentration donnée bien supérieure à la limite de détection, une validation proche de ce niveau critique précis serait appropriée. Il est impossible de définir une extrapolation générale sûre de ce résultat à d'autres concentrations de l'analyte, car beaucoup d'éléments dépendent du système analytique individuel. Par conséquent le rapport de l'étude de validation devrait préciser l'intervalle autour de la valeur critique, dans lequel, la personne qui procède à la validation, en recourant à un jugement professionnel, considère que l'incertitude estimée est vraissemblable.

Quand le domaine de concentration étudié s'approche de zéro, ou de la limite de détection, il est incorrect de faire l'hypothèse soit d'une incertitude absolue constante soit d'une incertitude relative constante. On peut procéder à une approximation utile de cette situation courante en posant l'hypothèse d'une relation fonctionnelle linéaire, avec une ordonnée à l'origine positive, entre l'incertitude u et la concentration c, de la forme suivante

- $u(c) = u_0 + \theta c$

où θ représente l'incertitude relative estimée pour une concentration bien supérieure à la limite de détection. u_0 représente l'incertitude type estimée pour une

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

concentration nulle, ce que dans certains cas pourrait être estimée comme étant $c_L/3$. Il serait alors raisonnable de considérer que le domaine de validité s'étend de zéro à un petit multiple entier du point de validation supérieur. Encore une fois il s'agira de recourir au jugement professionnel.

A8 limite de détection

En termes généraux la limite de détection est la plus petite quantité ou concentration de l'analyte dans l'échantillon d'essai que l'on peut distinguer de zéro de façon fiable.[22,1] La limite de détection ne doit pas nécessairement faire partie d'une validation pour les systèmes analytiques dans lesquels le domaine de validation ne l'inclut pas, ni ne s'en approche.

L'idée peut paraître simple, mais toute la question de la limite de détection est en proie aux problèmes énoncés ci-dessous:

- Il existe plusieurs approches conceptuelles sur le sujet, chacune donne une définition quelque peu différente de la limite. Les essais de clarification du sujet semble encore plus confus.
- Bien que chacune de ces approches dépende d'une estimation de la fidélité, à ou près d'une concentration nulle, on ne sait pas clairement si l'on doit appliquer des conditions de répétabilité ou quelque autre condition pour l'estimation.
- Sauf dans les cas où l'on a accès à une quantité exceptionnelle de données, les estimations de la limite de détection seront sujettes à une variation aléatoire assez importante.
- Les estimations de la limite de détection tendent souvent à être sous évaluées à cause de facteurs opérationnels.
- Des interférences statistiques liées à la limite de détection dépendent de l'hypothèse de répartition normale des résultats, ce qui est pour le moins douteux aux faibles concentrations.

A des fins pratiques, en matière de validation de méthode, il semble préférable de choisir une définition simple, conduisant à une estimation rapidement mise en œuvre que l'on utilisera seulement à caractère indicatif à des fins utiles pour la méthode. Il faut cependant reconnaître que la limite de détection telle qu'on l'estime lors de la mise au point de méthode peut ne pas être identique sur le plan du concept ou

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

numériquement à la limite utilisée pour caractériser une méthode analytique complète. Par exemple la “limite de détection instrumentale”, telle qu'on la trouve dans la littérature ou dans les brochures accompagnant les instruments et qui est ensuite ajustée pour la dilution, se trouve être souvent très inférieure à une limite “pratique” de détection; elle se révèle donc inadaptée pour la validation de méthode.

Par conséquent, pour la validation de méthode, on recommande de fonder l'estimation de fidélité utilisée ($\hat{\sigma}_0$) sur au moins 6 déterminations complètes et indépendantes de la concentration de l'analyte dans un blanc matrice type ou avec un matériau de faible concentration, sans censurer les résultats zéro ou négatifs; on calcule alors la limite de détection approximative comme $\widehat{3\sigma}_0$. Notons qu'avec le nombre minimal recommandé de degrés de liberté, cette valeur est assez incertaine, et peut facilement être erronée d’un facteur de deux. Quand on requiert des estimations plus rigoureuses (par exemple pour étayer des décisions prises sur la base de la détection ou non d'un matériau), certaines références bibiolgrapliques devraient apporter des recommandations adaptées (voir par exemple les références 22-1).

A9 Limite de determination ou limite de quantification

Il est parfois utile de déterminer une concentration en dessous de laquelle la méthode analytique ne peut pas fonctionner avec un degré acceptable de fidélité. On définit parfois arbitrairement cette fidélité comme étant 10% de l’ecart-type de l’erreur de la méthode, parfois la limite est tout aussi arbitrairement définie comme un multiple fixe (généralement 2) de la limite de détection. Bien qu’il soit dans une certaine mesure rassurant d'opérer au delà d'une telle limite, nous devons reconnaître qu’il s'agit là d'une délimitation assez artificielle de l'échelle de concentration: des mesurages inférieurs à une telle limite ne sont pas dépourvues d'informations et peuvent répondre à l’attente que l’on fait d’eux. C'est pourquoi l'utilisation de ce type de limite pour la validation n’est pas recommandée ici. Il est préférable d’essayer d'exprimer l'incertitude de la mesure en tant que fonction de la concentration, et de comparer cette fonction avec un critère d'aptitude à l’essai issu d’un accord entre le laboratoire et le client ou l'utilisateur final des données.

A10 Sensibilité

La sensibilité d'une méthode est la pente de la fonction d’étalonnage. Comme elle est souvent arbitraire, dépendant du paramétrage de l’instrument, la sensibilité n'est pas utile en validation, (elle peut néanmoins être utile dans les procédures d'assurance qualité, pour vérifier si la performance d'un instrument se situe à un niveau constant

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

et satisfaisant.).

A11 Robustesse

La robustesse d'une méthode d'analyse est la résistance au changement des résultats qu'elle produit quand des modifications mineures ont lieu par rapport aux conditions expérimentales décrites dans la procédure. Les limites pour des paramètres expérimentaux devraient être stipulées dans le protocole de la méthode (même si cela n'a pas toujours été fait dans le passé); et de telles modifications acceptées, séparément ou sous quelque combinaison possible, ne devraient pas produire de changement significatif dans les résultats produits. (Un "changement significatif" se traduit par le fait que la la méthode ne pouvait pas fonctionner sans franchir des limites d'incertitude adoptées qui définissent l'aptitude à l'essai.). Les aspects de la méthode susceptibles d'avoir une incidence sur les résultats devraient être identifiés, et leur influence sur la performance de la méthode devrait être évaluée en ayant recours à des études de robustesse.

La robustesse d'une méthode est testée en introduisant délibérément de légères variations dans la procédure et en observant les effets sur les résultats. Il peut être nécessaire de tenir compte d'un certain nombre d'aspects de la méthode, mais comme la plupart d'entre eux ont un effet négligeable, on peut généralement en faire varier plusieurs à la fois. Youden[13] a décrit une expérience économique basée sur des modèles factoriels fractionnés. Par exemple, il est possible de formuler une approche en utilisant 8 combinaisons de 7 facteurs variables, c'est-à-dire d'observer les résultats de sept paramètres à l'aide de seulement huit résultats analytiques. Les approches univariées sont également possibles quand une seule variable à la fois est modifiée.

Les exemples de paramètres que pourrait traiter un test de robustesse sont les suivants : changements d'instrument, d'opérateur ou de marque de réactif; concentration d'un réactif; pH d'une solution; température d'une réaction; temps imparti à l'achèvement du processus etc...

A12 Aptitude à l'essai

L'aptitude à l'essai, c'est le domaine à l'intérieur duquel la performance d'une méthode répond aux critères qui, d'accord entre l'analyste et l'utilisateur final des données, décrivent les besoins de l'utilisateur final. Par exemple, les erreurs dans les données ne devraient pas être d'une ampleur telle qu'elles suscitent des décisions incorrectes plus souvent qu'une petite probabilité définie; mais ces erreurs ne devraient pas être

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

rendues faibles au point que des dépenses inutiles soient demandées à l'utilisateur final. On peut fonder les critères d'aptitude à l'essai sur quelques unes des caractéristiques décrites dans cette annexe, mais en fin de compte ces critères s'expriment en terme d'incertitude totale acceptable.

A13 variation de matrice

Dans de nombreux domaines, la variation de matrice est l'une des sources d'erreurs les plus importantes mais les moins reconnues en matière de mesures analytiques. Lorsque nous définissons le système analytique à valider en précisant, entre autres choses, la matrice du matériau d'essai, des variations considérables peuvent exister à l'intérieur de la catégorie définie. Pour donner un exemple extrême, un échantillon de la catégorie "sol" peut être composé d'argile, de sable, de craie, de latérite (principalement Fe_2O_3 et Al_2O_3), de tourbe, etc.

Ou d'un mélange de ceux ci. On peut aisément imaginer que chacun de ces types produira un effet matrice singulier (unique) sur une méthode d'analyse comme que la spectrophotométrie d'absorption atomique. Si nous ne disposons pas d'informations au sujet du type de sols que nous analysons, il se produira une incertitude supplémentaire dans les résultats, du fait de cet effet matrice variable.

Il convient de quantifier séparément les incertitudes liées aux variations de matrice car elles ne sont prises en compte nulle part ailleurs dans le processus de validation. On acquiert cette information en recueillant un ensemble représentatif des matrices susceptibles d'être rencontrées à l'intérieur de la catégorie définie, toutes ayant des concentrations d'analytes dans le domaine approprié. Les matériaux sont analysés conformément au protocole, et l'on estime les biais dans les résultats. Sauf quand les matériaux de test sont des MRC, il faut généralement procéder à l'estimation des biais en utilisant les ajouts dosés et l'estimation du recouvrement. L'incertitude est estimée par l'écart type des biais. (Note: cette estimation renfermera également une contribution à la variance venant de l'analyse répétée. L'ampleur de celle-ci sera de

si l'ajout dosé a été employé. Si l'on exige un budget d'incertitude strict, ce terme devrait être déduit de la variance de la variation de matrice pour éviter de le comptabiliser deux fois.)

A14 Incertitude de Mesure réf. roger wood

L'approche formelle de l'estimation de l'incertitude de mesure calcule une estimation de l'incertitude de la mesure à partir d'une équation, ou modèle mathématique. Les procédures décrites pour la validation de méthode sont conçues pour assurer que

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

l'équation utilisée pour évaluer le résultat, prenant en compte les erreurs aléatoires de tous genres, soit une expression valable qui englobe tous les effets reconnus et significatifs sur le résultat. Il s'en suit, à une réserve près détaillée ci-après, que l'équation ou 'modèle' soumis à une validation peut être utilisé(e) directement pour estimer l'incertitude de mesure. Pour ce faire on suit des principes bien établis, fondés sur la 'loi de propagation de l'incertitude' qui, pour des effets d'entrée indépendants s'écrit:

- $u(y(x_1,x_2,...)) = \sqrt{\sum_{i=1,n} c_i^2 u(x_i)^2}$

où $y(x_1,x_2,....x_n)$ est une fonction de plusieurs variables indépendantes $x_1,x_2...$, et c_i un coefficient de sensibilité évalué par l'expression $c_i = \partial y / \partial x_i$, la différentielle partielle de y par rapport à x_i. $u(x_i)$ et $u(y)$ sont des *incertitudes type*, c'est-à-dire des incertitudes de mesure exprimées sous la forme d'écart-type. Puisque $u(y(x_1,x_2,...))$ est fonction de plusieurs estimations d' incertitudes distinctes, on y réfère comme étant *l' incertitude type combinée*.

Ainsi, pour estimer l'incertitude de mesure à partir de l'équation $y=f(x_1,x_2...)$ utilisée pour calculer le résultat, il faut, premièrement établir les incertitudes $u(x_i)$ pour chacun des termes x_1, x_2 etc..., et deuxièmement combiner ceux-ci avec les termes supplémentaires nécessaires pour représenter les effets aléatoires tels qu'ils sont trouvés en validation; et enfin il faut tenir compte de tous les effets supplémentaires. Dans la discussion de la fidélité figurant ci-dessus, le modèle statistique est:

- $y=f(x_1,x_2...) + \delta_{run} + e$

où e représente l'erreur aléatoire pour un résultat donné. Puisque δ_{run} et e ont respectivement des écarts types σ_{run} et σ_r qui sont connus à partir des expériences de fidélité, ces écarts type, (ou, strictement parlant, leurs estimations s_{run} et s_r) constituent les incertitudes associées à ces termes supplémentaires. Quand on fait la moyenne des résultats individuels intra-série, l'incertitude combinée associée à ces deux termes est (comme donné précédemment) $s_{tot} = (s_r^2/n + s_{run}^2)^{1/2}$. Notons que lorsqu'il est démontré que les termes de fidélité varient avec le niveau d'analyte, l'estimation d'incertitude pour un résultat donné doit utiliser le terme de fidélité approprié pour le niveau donné. Par conséquent, la base servant à estimer l'incertitude découle directement du modèle statistique posé comme hypothèse et

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

testé en validation. Il faut ajouter à cette estimation tous les termes supplémentaires nécessaires pour tenir compte (notamment) du manque d'homogénéité et de l'effet matrice (voir section A13). Enfin, l'incertitude type calculée est multipliée par un 'facteur de couverture', *k*, pour fournir une incertitude élargie, c'est-à-dire “un intervalle attendu couvrant une large fraction de la dispertion des valeurs que l'on peut attribuer au mesurande”[8]. Quand le modèle statistique est bien établi, la distribution reconnue comme étant normale et le nombre de degrés de liberté associés avec l'estimation est élevé, alors on choisit généralement *k* comme étant égal à 2. L'incertitude élargie correspond alors approximativement à un intervalle de confiance de 95%.

Il faut ici ajouter une réserve importante. Quand on teste le modèle statistique posé comme hypothèse, des essais imparfaits sont utilisés par la force des choses. Il a déjà été signalé que ces essais ne prouvent pas que quelque effet que ce soit fût identique à zéro ; ils ne peuvent montrer seulement qu'un effet est trop petit pour être détecté à l’intérieur de l'incertitude associée au test statistique de signification en question. Un exemple particulièrement important est le test statistique de signification du biais de laboratoire. D'évidence, s'il s'agit là du seul test effectué pour confirmer la justesse, il doit rester une incertitude résiduelle sur l'absence véritable et effective de biais dans la méthode. Il s'en suit que quand de telles incertitudes sont significatives par rapport à l'incertitude calculée jusqu'alors, il conviendrait de les prendre en compte.

Dans le cas d'une valeur de référence présentant une incertitude, le plus simple est de prendre en compte l'incertitude stipulée pour le matériau, combinée avec l'incertitude statistique de la méthode utilisée. Une discussion exhaustive de ce point dépasse la portée du texte ici présent ; la référence biobliographique 9 fournit des détails supplémentaires. Cependant, il est important de noter que, si l'incertitude estimée directement à partir du modèle statistique posé comme hypothèse constitue l'incertitude *minimale* qui peut être associée au résultat analytique, elle sera presque sûrement toujours sous estimée; de même, une incertitude élargie fondée sur les mêmes considérations et utilisant k=2 ne constituera pas un intervalle de confiance suffisant.

Le Guide ISO[8] recommande que, pour augmenter l’intervalle de confiance, plutôt que d'additionner des termes de façon arbitraire, la valeur de *k* devrait être augmentée si besoin est. L'expérience pratique suggère que, pour des estimations de l'incertitude fondées sur un modèle statistique validé, et alors qu'il n'existe aucune preuve au delà des études de validation qui renforce la confiance qu'on a dans ce modèle, *k* ne devrait pas être inférieur à 3. Quand on a de fortes raisons de soupçonner que l'étude de

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

validation n'est pas exhaustive, k devrait être encore augmenté si nécessaire.

ANNEXE B. considerations supplémentaires pour l'evaluation de l'incertitude dans les etudes de VALIDATION

B1 Analyse de la Sensibilité

L'expression de base utilisée pour l'évaluation de l'incertitude

- $u(y(x_1,x_2,\ldots)) = \sqrt{\sum_{i=1,n} c_i^2 u(x_i)^2}$

requiert les 'coefficients de sensibilité' c_i. Il est courant au cours de l'évaluation de l'incertitude d'observer que, alors qu'un facteur d'influence donné x_i a une incertitude connue $u(x_i)$, le coefficient c_i n'est pas suffisamment caractérisé ou pas immédiatement disponible à partir de l' équation d'expression du résultat. Ceci est particulièrement courant dans les cas où un effet n'est pas inclus dans l'équation d'expression de mesure parce qu'il n'est pas normalement significatif, ou parce que la relation n'est pas suffisamment expliquée pour justifier une correction. Par exemple, l'effet de la température d'une solution T_{sol} sur une procédure d'extraction à température ambiante est rarement établi dans le détail.

Quand on désire évaluer l'incertitude relative à un résultat soumi à un tel effet, on peut déterminer le coefficient de façon expérimentale. Cela peut être fait très simplement en changeant x_i et en observant l'effet sur le résultat, d'une façon très semblable aux essais de robustesse de base. Dans la plupart des cas, il suffit en premier lieu de choisir tout au plus deux valeurs pour x_i en dehors de la valeur nominale, et de calculer une pente approximative à partir des résultats observés. Cette pente nous fournit une valeur approximative pour c_i. Le terme $c_i.u(x_i)$ peut alors être déterminé. (Notons qu'il s'agit là d'une méthode pratique pour démontrer la signification ou l'absence de signification d'un effet possible sur les résultats).

Lors d'une telle expérience, il est important que le changement de résultat observé soit suffisant pour permettre un calcul fiable de c_i. Cela est difficile à prévoir à l'avance. Cependant, pour un intervalle donné où varie la grandeur d'influence x_i, ou alors pour une incertitude élargie relative à cette grandeur, dans lesquels on s'attend à voir produire un changement insignifiant sur le résultat, il est évidemment important

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005

d'évaluer c_i à partir d'un plus grand intervalle. Par conséquent on recommande, pour une grandeur d'influence avec un intervalle attendu de ± *a*, (où ± *a* peut être, par exemple : l'intervalle autorisé, l'intervalle d'incertitude élargi ou un intervalle de confiance à 95%), que l'expérience de sensibilité utilise dans la mesure du possible un changement d'au moins 4*a* pour garantir des résultats fiables.

B2 Jugement

Il n'est pas rare d'observer que même si un effet est reconnu et peut être significatif, il n'est pas toujours possible d'obtenir une estimation fiable de l'incertitude. Dans de telles circonstances, le Guide ISO dit assez clairement qu'une estimation de l'incertitude effectuée par le professionnel est préférable au fait de négliger l'incertitude. Ainsi, quand on ne dispose pas d'une estimation de l'incertitude pour un effet potentiellement important, l'analyste devrait se fier à son meilleur jugement pour évaluer l'incertitude probable et l'appliquer à l'estimation de l'incertitude combinée. La référence 8 donne des indications supplémentaires sur l'utilisation du jugement en matière d'évaluation de l'incertitude.

---

[*]On ne tient pas compte dans le document présent de l'incertitude d'échantillonnage, au sens strict d'incertitude due à la préparation de l'échantillon de laboratoire à partir du volume d'origine. L'incertitude liée au prélèvement sur l'échantillon de laboratoire d'une prise d'essai fait intégralement partie de l'incertitude de la mesure et donc est automatiquement incluse aux différents niveaux de l'analyse suivante.

[+24]+ Plusieurs groupes alternatifs ou 'partitions d'erreur' sont possibles et peuvent être utiles lors de l'étude plus approfondie de sources d'erreur bien spécifiques ou pour un ensemble différent de situations. Par exemple, le modèle statistique ISO 5725 combine généralement les effets de l'expérience et du laboratoire, tandis que la procédure d'évaluation de l'incertitude de ISO GUM se prête bien à l'estimation de l'influence sur le résultat de chaque élément distinct et mesurable.

[*] Cela peut ne pas s'appliquer à des concentrations en deçà de dix fois la limite de détection.

Exemplaire certifié conforme Paris, 20 juillet 2005
Le Directeur Général de l'OIV
Secrétaire de l'Assemblée Générale
Frederico CASTELLUCCI

© OIV 2005