

CHAPITRE 2.2.6.

SELECTION ET UTILISATION DES ECHANTILLONS ET PANELS DE REFERENCE

INTRODUCTION

Les Recommandations de l'OMSA pour la validation fournissent des informations détaillées et des exemples à l'appui de la Norme de validation publiée au Chapitre 1.1.6 Validation des épreuves diagnostiques pour les maladies infectieuses des animaux terrestres du présent Manuel terrestre. L'expression « Norme de validation de l'OMSA » dans le présent chapitre doit être comprise comme renvoyant à ce chapitre.

Les échantillons et panels de référence sont essentiels, depuis l'étude de faisabilité initiale dans le laboratoire de développement jusqu'à l'entretien et au contrôle de la performance de l'essai dans le laboratoire diagnostique, en passant par toutes les étapes intermédiaires. On ne saurait trop insister sur l'importance fondamentale des échantillons et panels de référence. Le choix d'un matériel de référence inadéquat peut mener à des biais et fausser les conclusions dès la mise au point de l'essai et jusqu'à sa validation et à son utilisation. C'est pourquoi il faut apporter beaucoup de soin au choix des échantillons de référence et à la conception des panels.

Fig. 1. Échantillons et panels de référence groupés sur la base de leurs caractéristiques communes et de leur composition.
Les rubriques et les sous-titres alphanumériques (p. ex. Faisabilité, A.2.1) renvoient à la section correspondante de la Norme de validation de l'OMSA.

Groupe A		Groupe B		Groupe D
Faisabilité A.2.1		SpA B.1.2.		Comparaison des méthodes normalisées B.2.6.
Plage de fonctionnement A.2.2.		Précision analytique B.1.4.		Reconnaissance provisoire B.2.6.
Optimisation A.2.3.		Échantillons et panels de référence		Modifications biologiques B.5.2.2.
Robustesse A.2.5.				
Étalonnage A.2.6.		Groupe C		Groupe E
Contrôle du procédé A.2.6.		Répétabilité B.1.1.		SpD et SeD Gold standard B.2.1.
SeA B.1.3.		Reproductibilité initiale B.2.6.		
Modifications techniques B.5.2.1.		Reproductibilité B.3.		Groupe F
Remplacement des réactifs B.5.2.3.		Contrôle des compétences B.5.1.		SpD et SeD No gold standard B.2.2.

SpA = spécificité analytique ; SeA = sensibilité analytique ; SpD = spécificité diagnostique ;
SeD = sensibilité diagnostique

Comme le montre la Figure 1, les échantillons et/ou panels de référence sont mentionnés tout au long de la Norme de validation de l'OMSA. Selon la définition du glossaire des Normes de qualité et Lignes directrices de l'OMSA pour les laboratoires vétérinaires : maladies infectieuses, « les matériels de référence sont des substances dont les propriétés sont suffisamment homogènes et bien établies pour être utilisées pour l'étalonnage d'un appareil, l'évaluation d'une méthode de mesure ou pour assigner des valeurs aux matériels¹ ». Dans le cadre de la validation des méthodes de test, les matériels ou échantillons de référence contiennent l'analyte faisant l'objet du test à des concentrations ou à des activités variables et ils sont utilisés pour mettre au point et évaluer les caractéristiques de performance analytique et diagnostique de l'essai candidat. Dans le cas présent, le terme « analyte » désigne le composant spécifique d'un échantillon de test, qui est détecté ou mesuré par la méthode de test, à savoir l'anticorps, l'antigène ou l'acide nucléique. Ces échantillons de référence peuvent être des sérums, des liquides, des tissus, des excréments, des aliments ou des produits environnementaux contenant l'analyte faisant l'objet du test et sont généralement prélevés sur des animaux infectés ou dans leur environnement. Cependant, dans certains cas, ils peuvent être préparés au laboratoire à partir d'un matériel de base original (p. ex. une dilution de sérum fortement positif dans un sérum négatif) ou éventuellement produits en ajoutant l'analyte extrait à la matrice choisie (p. ex. culture bactérienne ou virale, protéine recombinante/exprimée ou construction génomique). Qu'ils soient naturels ou fabriqués, ils sont utilisés tout au long des processus de mise au point et de validation et sur toute la durée de vie d'un essai afin d'en contrôler la performance.

Dans la Figure 1, les échantillons et panels de référence sont groupés selon leurs caractéristiques communes et leur composition, ces regroupements servant de base aux descriptions qui suivent. Une indication renvoie à la Section correspondante de la Norme de validation de l'OMSA pour chaque utilisation des échantillons et panels de référence.

Les échantillons de référence peuvent être utilisés à différentes fins dès les premières étapes du développement et de l'optimisation, au cours de l'étape 1 de la validation ainsi que pour le contrôle et l'entretien continu de l'essai. Lorsque cela est possible, de grandes quantités de ces échantillons de référence doivent être collectées ou préparées, puis conservées pour un usage à long terme. Changer d'échantillons de référence durant le processus de validation introduit une variable irréductible qui peut fortement discréditer l'interprétation des données expérimentales et donc l'intégrité du processus de développement et de validation. Pour les essais susceptibles de cibler plusieurs espèces, les échantillons doivent être représentatifs de l'espèce primaire concernée. Il est fondamental qu'ils reflètent à la fois l'analyte cible et la matrice où l'analyte se trouve dans la population à laquelle l'essai est destiné. Le matériel de référence doit représenter de manière appropriée la plage de concentration de l'analyte que l'essai est supposé détecter.

Il est important de souligner que tous les critères de sélection, toutes les procédures de préparation et toutes les exigences d'analyse doivent faire l'objet d'une description intégrale et qu'il faut les réunir dans un document de contrôle, que les échantillons de référence soient d'une provenance naturelle ou qu'ils aient été préparés en laboratoire. Il ne s'agit pas uniquement de bonnes pratiques de gestion de la qualité, cela apporte également un degré accru de continuité et de confiance pour toute la durée de vie de l'essai.

A. GROUPE A

La question de savoir s'il est possible de regrouper des échantillons pour créer un échantillon de référence revient souvent. Lorsque du matériel de référence est prélevé sur un seul animal, il est important de vérifier si celui-ci est ou non représentatif de l'évolution typique d'une phase de l'infection dans le contexte de la population à tester. Sinon, cela peut induire un biais et fausser les conclusions relatives à la validation. Le regroupement est une bonne alternative, mais il est impératif de regrouper des prélèvements effectués sur des animaux se trouvant dans la même phase infectieuse. C'est particulièrement important pour les systèmes de détection des anticorps. Le regroupement permet aussi de résoudre le problème des quantités de matériel à stocker pour un usage à long

1 https://www.techlab.fr/Commun/UK_Def_MRC.asp

terme, notamment lorsque l'on travaille avec des espèces d'hôtes de taille réduite. Avant de regrouper des échantillons, il est préférable de les tester indépendamment pour démontrer qu'ils sont similaires du point de vue de leur concentration en analyte et/ou de leur réactivité. Une évaluation doit se faire après le regroupement pour s'assurer qu'aucune interférence inattendue n'ait surgi lors du regroupement de plusieurs échantillons. Par exemple, il arrive que certains types de sang ou certaines compositions d'anticorps au sein des différents échantillons indépendants présentent une réactivité croisée lors du regroupement et que l'échantillon groupé adopte alors un comportement différent de celui des échantillons individuels testés indépendamment avec l'essai en question.

Il est souvent difficile d'obtenir des échantillons individuels véritablement représentatifs des concentrations ou des réactivités des analytes sur tout le spectre attendu. Compte tenu de la dynamique des infections et des réponses aux agents pathogènes, les plages intermédiaires sont souvent très passagères. Dans le cas des réponses en anticorps, les phases d'infection précoces chez les animaux pris individuellement résultent souvent d'anticorps d'isotypes et d'avidités variables et hétérogènes. En général, ces échantillons ne font pas de bons échantillons de référence pour évaluer les caractéristiques analytiques d'un essai. Ils sont néanmoins importants pour différents types de panels de référence comme nous le verrons plus loin. Pour la plupart des rubriques du groupe A, il est acceptable d'utiliser des échantillons préparés qui seront ajoutés, selon des concentrations en analyte connues ou en séries de dilution fortement positives, à une matrice négative pour créer une gamme de concentrations.

Qu'ils soient naturels ou préparés, les échantillons de référence doivent représenter la gamme attendue des concentrations d'analyte, de faiblement à fortement positif, telle qu'elle pourrait se présenter durant le cours typique de l'infection. Un échantillon de référence négatif doit être inclus comme contrôle de fond. Si une matrice négative est utilisée comme diluant pour la préparation d'un échantillon de référence positif (p. ex. un sérum négatif utilisé pour diluer un sérum fortement positif ou un tissu enrichi d'une construction), cette matrice doit impérativement être incluse comme échantillon de référence négatif.

Comme mentionné plus haut, tous les échantillons de référence doivent être bien caractérisés. Cela inclut la documentation concernant l'agent pathogène et l'hôte donneur. Pour les agents pathogènes, cela peut comprendre des détails relatifs à la souche, au sérotype, au génotype, à la lignée, etc. La provenance du matériel hôte doit être bien décrite quant à l'espèce, la race, l'âge, le sexe, le statut reproducteur, les commémoratifs de vaccination, les antécédents du troupeau, etc. Lorsque c'est possible, la phase d'infection doit être notée. Cela peut également inclure des détails relatifs aux signes cliniques, aux profils d'anticorps, à la charge pathogène ou à l'excrétion de l'agent, etc. Tout aussi importants, les tests utilisés pour déterminer le statut sanitaire/infectieux doivent être bien documentés (voir Section E de ce chapitre pour de plus amples explications). Parfois, une infection/exposition expérimentale peut constituer la seule option viable pour produire le matériel de référence. Dans ce cas, toutes les considérations ci-dessus doivent être détaillées en sus du protocole expérimental.

Point fondamental : qu'il soit naturel ou préparé, le statut du matériel de référence doit être sans ambiguïté, l'échantillon représentant soit un vrai positif, soit un vrai négatif. Cela peut nécessiter que ce statut soit confirmé au moyen d'un autre test ou d'une batterie de tests. Ainsi, de nombreux sérums de référence pour les anticorps sont caractérisés à l'aide de plusieurs tests sérologiques. Cela renforce la confiance et apporte des caractéristiques documentées supplémentaires, qui s'avéreront éventuellement nécessaires lorsqu'il s'agira de remplacer ou de dupliquer le matériel de référence.

Les recommandations concernant la stabilité et le stockage du matériel de référence figurent sous : <http://www.woah.org/fr/expertise-scientifique/produits-veterinaires/reactifs-de-reference/>.

1. Faisabilité (Norme de validations de l'OMSA, Section A.2.1)

La Norme de validation de l'OMSA préconise que les méthodes de test et les procédures associées soient appropriées à leur utilisation diagnostique spécifique afin que les résultats de test soient pertinents. En d'autres termes, l'essai doit être « adapté à l'objectif prévu ». De nombreux essais sont développés dans de bonnes intentions, mais sans avoir à l'esprit leur utilisation spécifique. Au tout début, il est fondamental que le ou les objectifs diagnostiques soient définis par rapport à la ou aux populations qui seront testées. Les objectifs les plus courants sont énumérés en termes génériques dans la Section A de la Norme de validation de l'OMSA. En tant que tels, ils incluent des utilisations plus étroites et spécifiques. Ces objectifs spécifiques doivent cependant être clairement définis dès le début et sont extrêmement importants dans le cadre d'essais entièrement validés. Comme le montrent les descriptions qui suivent, une définition claire de l'utilisation aura des conséquences à la

fois sur le choix des échantillons et des panels de référence et sur la conception des évaluations analytiques et diagnostiques.

2. Plage de fonctionnement (Norme de validation de l'OMSA, Section A.2.2) et sensibilité analytique (Norme de validation de l'OMSA, Section B.1.3)

2.1. Méthodes analytiques

La plage de fonctionnement d'un essai est l'intervalle des concentrations (quantités) d'analyte sur lequel la méthode fournit une exactitude et une précision adéquates. Elle définit également les limites de détection inférieure et supérieure de l'essai. Pour établir cette plage, un échantillon de référence fortement positif sera choisi. Cet échantillon fortement positif, qu'il soit naturel ou préparé, sera dilué en série jusqu'à extinction dans une matrice négative représentative de la matrice des échantillons provenant d'animaux de la population ciblée par l'essai. Cela inclut les titrages d'anticorps où un sérum de référence fortement positif doit être dilué dans un sérum de référence négatif pour créer une série de dilutions. La sensibilité analytique (SeA) est indiquée par la limite inférieure de détection (LD) d'un analyte dans un essai. Le même échantillon de référence fortement positif pourra être utilisé pour déterminer à la fois la plage de fonctionnement et la limite de détection analytique.

2.2. Méthodes comparatives

Si l'objectif prévu est de détecter de faibles taux d'analyte ou des infections subcliniques, il peut s'avérer difficile d'obtenir le matériel de référence approprié dès les stades précoces du processus infectieux. Dans certains cas, il peut être utile de déterminer une SeA comparative en analysant un panel d'échantillons avec l'essai candidat ainsi qu'avec un autre essai indépendant. Idéalement, ce panel d'échantillons doit être prélevé périodiquement sur des animaux infectés naturellement ou expérimentalement et doit, si possible, représenter les animaux infectés peu après l'infection, une fois les signes cliniques devenus manifestes ainsi que lors d'évolutions fulminantes. Cela permet une comparaison de la SeA entre les essais ainsi qu'une comparaison chronologique du point de détection le plus précoce par rapport à la pathogénèse de la maladie.

Une expérience comme celle décrite ci-dessus fournit une occasion unique de prélever des échantillons de référence représentatifs d'un éventail naturel de concentrations et susceptibles d'être utiles pour d'autres objectifs de validation. On prendra soin d'éviter d'utiliser de tels échantillons dans les cas qui ne s'y prêtent pas (voir Groupe D ci-dessous). Lorsque c'est possible, des séries d'échantillons doivent être prélevées sur cinq animaux au moins au cours de l'infection. Dans les cas où l'échantillonnage est létal (c'est-à-dire qu'il requiert le prélèvement des tissus d'un organe interne), le nombre d'animaux requis sera de cinq au minimum par échantillonnage. Pour les espèces d'hôtes de petite taille, il peut s'avérer nécessaire d'augmenter ce nombre afin de collecter suffisamment de matériel de référence. Les expériences de ce type nécessitant des ressources importantes, il est judicieux de maximiser la collecte non seulement d'échantillons de référence ciblés pour le cas à l'étude, mais aussi de matériel complémentaire (autres tissus, liquides, etc.) susceptibles de servir ultérieurement de matériel de référence.

3. Optimisation (Norme de validation de l'OMSA, Section A.2.3) et répétabilité initiale (Norme de validation de l'OMSA, Section A.2.6)

L'optimisation est le processus par lequel les principaux paramètres physiques, chimiques et biologiques d'un essai sont évalués et ajustés pour garantir que les caractéristiques de performance de l'essai sont les mieux adaptées à l'utilisation prévue. Trois échantillons de référence au moins (négatif, faiblement positif et fortement positif) peuvent être choisis à partir d'échantillons de référence naturels ou préparés. Les expériences d'optimisation sont relativement exhaustives, notamment lorsque les essais comportent plusieurs étapes de préparation et d'analyse. Il est très important qu'une quantité suffisante de chaque échantillon de référence soit disponible pour effectuer toutes les expériences d'optimisation. Il n'est pas recommandé de changer d'échantillons de référence en cours d'optimisation, car cela ajoute une variable non contrôlée et interrompt la continuité des preuves d'optimisation.

L'évaluation de la répétabilité doit commencer lors de la mise au point de l'essai et des étapes d'optimisation. La répétabilité sera également vérifiée au cours de l'étape 1 de la validation de l'essai (Section B.1.1). Les mêmes

échantillons de référence doivent être utilisés pour les deux processus, une fois de plus pour attester de la continuité des preuves.

4. Étalonnage et contrôles des procédés (Norme de validation de l'OMSA, Section A.2.6)

4.1. Étalons analytes de référence internationaux, nationaux ou internes

Les étalons de référence internationaux sont hautement caractérisés, contiennent des concentrations définies d'analyte et sont généralement préparés et conservés par des laboratoires de référence internationaux. Ils sont les réactifs par rapport auxquels tous les essais et/ou d'autres matériels de référence doivent être normalisés. Les étalons de référence nationaux sont étalonnés par comparaison avec un étalon de référence international lorsque c'est possible. En l'absence d'étalon international, un étalon national peut être choisi ou préparé pour devenir ensuite l'étalon de comparaison pour l'essai candidat. En l'absence des deux étalons ci-dessus, un étalon interne doit être sélectionné et préparé par le laboratoire de développement au sein de l'organisation responsable. Tous les étalons de référence, qu'ils soient naturels ou préparés, doivent être hautement caractérisés au moyen d'une analyse approfondie et les méthodes utilisées pour leur caractérisation, leur préparation et leur stockage doivent de préférence avoir fait l'objet de publications dans des revues à comité de lecture. Ces étalons de référence doivent également être stables et sans danger.

Les étalons de référence, en particulier les anticorps, sont généralement fournis dans l'un ou l'autre des deux formats. Ils peuvent être fournis comme réactif positif unique d'un titre défini, dans la perspective que l'essai candidat soit normalisé pour livrer un titre équivalent. Il s'agit d'une méthode analytique très simple mais beaucoup de ces étalons « uniques » sont préparés à partir d'échantillons fortement positifs sous forme de pré-dilution dans une matrice négative afin de maximiser le nombre d'aliquotes existants. L'inconvénient ici est qu'un éventuel effet de la matrice dans l'essai candidat n'est pas pris en compte, puisqu'aucun témoin n'est fourni pour la matrice. L'autre méthode consiste à fournir un ensemble d'étalons de référence négatifs, faiblement positifs et fortement positifs aux concentrations ou aux réactivités connues et se trouvant dans la plage de fonctionnement de la méthode normalisée utilisée pour les préparer. Cela compense l'effet de matrice éventuellement caché. Par ailleurs, cet ensemble de trois étalons sert de modèle pour la sélection et/ou la préparation des contrôles des procédés (voir discussion ci-dessous).

Traditionnellement, les étalons ci-dessus sont des étalons d'anticorps polyclonaux et, dans une moindre mesure, des étalons d'antigènes conventionnels utilisés pour l'étalonnage des essais sérologiques. Aujourd'hui toutefois, les étalons de référence peuvent aussi être des anticorps monoclonaux, des protéines recombinantes/exprimées ou des constructions génomiques lorsqu'il s'agit d'étalonner des épreuves par rapport à un étalon de performance spécifique.

4.2. Étalons de travail ou contrôles des procédés

Les étalons de travail, couramment appelés contrôles de la qualité ou du procédé, sont étalonnés par rapport à des réactifs étalons internationaux, nationaux ou internes. Ils sont sélectionnés et préparés dans la matrice locale que l'on trouve dans la population à laquelle l'essai est destiné. Idéalement, des étalons de travail négatifs ainsi que faiblement et fortement positifs doivent être sélectionnés ou préparés. Les concentrations et/ou les réactivités doivent se situer dans la plage normale de fonctionnement de l'essai. De grandes quantités doivent être préparées, réparties en aliquotes et stockées pour être utilisées en routine dans chaque cycle diagnostique de l'essai. Le but de ces contrôles est d'imiter, autant que possible, les échantillons de terrain, en les manipulant et en les testant comme les échantillons de routine. Ils sont utilisés pour établir des limites de contrôle supérieures et inférieures de la performance de l'essai et pour contrôler de manière aléatoire et/ou systématique la variabilité au moyen de diverses méthodes et chartes de contrôle. Leur performance quotidienne détermine si, oui ou non, un essai est sous contrôle et si ses cycles individuels peuvent être acceptés. À ce titre, les étalons de travail sont fondamentaux du point de vue de la gestion de la qualité.

5. Modifications techniques (Norme de validation de l'OMSA, Section B.5.2.1)

Les modifications techniques apportées à un essai validé, comme les changements d'instrumentation, de protocoles d'extraction ou la conversion d'un essai à un système semi-, voire entièrement, automatisé et robotisé, ne nécessiteront généralement pas une revalidation intégrale de l'essai. Une étude comparative des méthodes

sera plutôt faite afin de déterminer si les modifications mineures apportées au protocole d'essai sont susceptibles d'influencer les résultats du test. Voir le Chapitre 2.2.8 *Comparabilité des épreuves suite à des changements introduits dans une méthode de test validée* pour les méthodes statistiques servant à évaluer la précision de l'essai après des modifications techniques.

En général, ces méthodes requièrent l'utilisation de trois échantillons de référence : un négatif, un faiblement positif et un fortement positif. Une fois de plus, ces échantillons peuvent être naturels ou préparés. Ce qui compte, c'est que les échantillons de référence qui ont été utilisés dans les phases de mise au point de l'essai peuvent être réutilisés pour évaluer les modifications une fois la méthode appliquée à un usage diagnostique de routine. Cela apporte un niveau plus élevé de confiance pour évaluer les répercussions potentielles des modifications, car les caractéristiques de performance de ces échantillons de référence ont été bien définies. Lorsqu'il est nécessaire d'utiliser de nouveaux échantillons de référence, ceux-ci doivent être sélectionnés ou préparés en utilisant au moins les critères ou les procédures de préparation établies pour les échantillons précédents. Cela renforce à nouveau la continuité de la preuve.

6. Remplacement des réactifs (Norme de validation de l'OMSA, Section B.5.2.3)

Lorsqu'un réactif tel qu'un échantillon de contrôle du procédé arrive à épuisement, il est essentiel de préparer et de tester plusieurs fois son remplaçant. Le remplaçant envisagé doit être inclus dans plusieurs cycles de l'essai, parallèlement au contrôle d'origine, afin d'établir leur rapport de proportionnalité. Il est important de ne pas changer plus d'un réactif de contrôle à la fois pour éviter d'aggraver le problème en devant évaluer simultanément plus d'une variable.

À nouveau, on ne saurait trop insister sur le fait que tout remplaçant d'un réactif de référence doit être sélectionné ou préparé en utilisant les critères et les procédures de préparation établies pour le matériel précédent. Cela renforce la continuité de la preuve et la confiance dans l'essai.

B. GROUPE B

1. Spécificité analytique (Norme de validation de l'OMSA, Section B.1.2)

La spécificité analytique (SpA) est le degré selon lequel l'essai différencie l'analyte cible et les autres composants qui peuvent être détectés dans l'essai. Il s'agit d'une définition relativement large qui, souvent, n'est pas bien comprise. La SpA peut être décomposée en différents éléments, comme décrit ci-après.

Le choix des échantillons de référence requis pour évaluer la SpA dépend largement de l'objectif ou de l'utilisation spécifique initialement prévu(e) lors de la mise au point de l'essai. L'évaluation de la SpA est un élément crucial de l'étude de faisabilité et de la vérification de l'adéquation à l'objectif.

Un élément important est la mesure dans laquelle une méthode peut détecter et/ou quantifier avec exactitude l'analyte cible lorsqu'il s'agit d'acides nucléiques, de protéines et/ou d'anticorps dans la matrice de test. On parle parfois de « sélectivité ». Un exemple est l'utilisation d'échantillons de référence pour des tests qui sont conçus pour différencier des animaux infectés et des animaux vaccinés (tests DIVA).

Les échantillons de référence doivent être sélectionnés et testés à partir d'animaux i) non infectés/non vaccinés, ii) non infectés/vaccinés, iii) infectés/non vaccinés et iv) infectés/vaccinés. Ces échantillons peuvent être collectés en conditions réelles, mais il est important que les antécédents précis soient recueillis, idéalement en relation avec les animaux, mais au minimum avec les troupeaux concernés, ce qui inclut les habitudes de vaccination et les épisodes de maladies. Sinon, il peut s'avérer nécessaire de produire ce matériel dans des expériences comme celles décrites à la Section A.2.2 du présent chapitre, mais comprenant une combinaison d'animaux vaccinés et d'animaux inoculés. Il est important d'éviter d'utiliser le vaccin comme antigène de capture dans l'essai (p. ex. essai immuno-enzymatique indirect [I-ELISA]) car les protéines porteuses dans le vaccin pourraient stimuler des réponses humorales non spécifiques chez les animaux vaccinés et être détectées dans l'ELISA, aboutissant à des faux positifs dans l'essai. De manière similaire à l'approche comparative décrite ci-dessus par rapport à la SeA, il faut envisager de sélectionner cinq animaux par groupe au moins. Pour les espèces hôtes de petite taille, il peut s'avérer nécessaire d'augmenter ce nombre afin de récolter suffisamment de matériel de référence. Selon le test DIVA concerné, une expérience unique peut être conçue pour évaluer à la fois la SeA et la SpA.

Un deuxième élément, parfois appelé « exclusivité », est la capacité de l'essai à détecter un analyte ou une séquence génomique unique dans l'organisme cible et à exclure tous les autres organismes connus susceptibles d'avoir une activité croisée. Cela est particulièrement vrai pour les épreuves sérologiques où plusieurs exemples d'antigènes sont exprimés par d'autres organismes capables d'induire des anticorps à la réactivité croisée. Une tentative doit être faite pour obtenir des échantillons de référence à partir de cas d'infection documentés et/ou d'organismes susceptibles d'avoir une activité croisée. Selon le type d'essai, ce matériel de référence peut représenter l'organisme lui-même, des échantillons extraits de l'hôte ou des séquences génomiques. Un profil de l'exclusivité de l'essai doit être établi et élargi sur une base permanente au fur et à mesure que des microorganismes susceptibles d'avoir une activité croisée apparaissent.

Un troisième élément essentiel de la conception d'un essai se rapporte à sa capacité à détecter une ou plusieurs souches ou sérovars d'une espèce, plusieurs espèces d'un genre, ou un regroupement similaire d'organismes ou d'anticorps étroitement apparentés. Cela définit le champ de détection et, de ce fait, l'adéquation à l'objectif. Des échantillons de référence sont nécessaires pour définir la portée de l'essai. Si, par exemple, un essai est mis au point comme test de dépistage pour détecter tous les génotypes ou sérotypes connus d'un virus, les échantillons de référence de chaque type représentatif doivent alors être testés. Lorsque de nouvelles lignées ou de nouveaux variants d'un sérotype apparaissent, ils doivent également être testés et faire partie intégrante du profil de tests qui à actualiser de manière continue.

2. Exactitude analytique des tests complémentaires (Norme de validation de l'OMSA, Section B.1.4)

Certaines méthodes ou procédures de test sont uniquement des outils analytiques, généralement utilisés pour mieux caractériser un analyte qui a été détecté dans un essai primaire, par exemple les tests de neutralisation virale utilisés pour typiser un virus préalablement isolé ou pour caractériser une réponse en anticorps. Ce genre de tests complémentaires doit être validé pour ses caractéristiques de performance analytique, mais diffère des tests diagnostiques de routine puisqu'il ne nécessite pas de validation pour ses caractéristiques de performance diagnostique. La précision analytique de ces tests est souvent dépendante de l'emploi de réactifs de référence. Ces réactifs, qu'il s'agisse d'anticorps pour typiser les souches d'organismes ou de souches de référence de l'organisme, etc., doivent être soigneusement documentés, comme tout autre matériel de référence, quant à leur provenance, leur identité et leurs caractéristiques de performance.

C. GROUPE C

Les échantillons de référence du Groupe C peuvent être utilisés pour de nombreux objectifs. Dans les phases de développement initiales, ils peuvent être utilisés pour évaluer la répétabilité de l'essai ainsi que la reproductibilité initiale de l'étape 1 et l'évaluation plus approfondie de la reproductibilité de l'étape 3 du processus de validation. Une fois transférés au laboratoire de diagnostic, ces échantillons ont toutefois plusieurs autres usages potentiels. Ils peuvent être utilisés comme panels pour la formation et la qualification des analystes ou pour évaluer les compétences du laboratoire lors de programmes externes d'essais circulaires. Idéalement, 20 échantillons individuels ou plus doivent être préparés en grandes quantités. Un quart environ (25 %) doit être constitué d'échantillons négatifs et le reste (75 %) doit représenter un ensemble d'échantillons positifs couvrant toute la plage de fonctionnement de l'essai. Ils doivent être répartis en aliquotes dans des tubes individuels en quantités suffisantes pour un seul usage et stockés pour un usage à long terme. Le nombre d'aliquotes nécessaires pour chaque échantillon dépend du nombre de laboratoires utilisant l'essai dans leur diagnostic de routine et de la manière dont le contrôle des compétences est prévu. Idéalement, ils seront préparés en quantités inépuisables, ce qui est rarement faisable. Au minimum, plusieurs centaines d'aliquotes (ou plus) de chaque échantillon seront préparées en une seule fois en testant le même échantillon sur plusieurs intervalles de tests, un moyen utile de détecter une erreur systématique (biais) qui pourrait s'insinuer dans l'utilisation à long terme de l'essai.

Ces échantillons peuvent être naturels ou préparés à partir de matériel de base individuel ou regroupé. Le but est qu'ils imitent autant que possible un véritable échantillon de test. Comme le stockage en masse est toujours un problème, il peut s'avérer nécessaire de stocker ce matériel en vrac et de préparer des aliquotes de travail de temps à autre. Si l'on dispose d'espaces de stockage, il est toutefois préférable de préparer et de stocker de grands nombres d'aliquotes en une fois, car les analytes en vrac, soumis aux cycles congélation-décongélation pour préparer quelques aliquotes à la fois, risquent de se détériorer. Comme ce type de matériel de référence est consommé à un taux relativement élevé, il devra être remplacé ou réapprovisionné sur une base continue. Puisque le matériel de remplacement potentiel est identifié lors de tests de routine ou lors de foyers, il est conseillé d'avoir des interlocuteurs sur le terrain à même d'obtenir du matériel de référence en vrac et de le

stocker pour un usage ultérieur. Sinon, il peut s'avérer nécessaire de produire ce matériel dans des expériences comme celles décrites à la Section A.2.2 de ce chapitre. Comme pour la méthode comparative décrite plus haut à propos de la SeA, il faut envisager de sélectionner cinq animaux par groupe au moins. Pour les espèces hôtes de petite taille, il peut être nécessaire d'augmenter ce nombre afin de collecter suffisamment de matériel de référence.

1. Répétabilité (Norme de validation de l'OMSA, Section B.1.1) et reproductibilité initiale (Norme de validation de l'OMSA, Section B.2.6)

La répétabilité est le degré de concordance entre les résultats des copies d'un échantillon, intra-cycle et inter-cycles, selon la même méthode de test dans un laboratoire donné. La répétabilité est estimée en évaluant la variation des résultats des copies à partir d'un minimum de trois échantillons (ou cinq de préférence) représentant l'activité de l'analyte à l'intérieur de la plage de fonctionnement de l'essai. Pour évaluer la répétabilité, voir le Chapitre 2.2.4 *Incertitude des mesures*, qui décrit les méthodes statistiques utilisées pour mesurer l'incertitude.

La reproductibilité est l'aptitude d'une méthode de test à fournir des résultats constants, comme défini par les valeurs estimatives de précision, lorsque la méthode est appliquée aux aliquotes des mêmes échantillons testés dans des laboratoires différents. Les estimations de la reproductibilité initiale de l'essai candidat doivent cependant être déterminées durant les étapes de son développement. Un petit panel de trois échantillons (ou cinq de préférence) représentant les négatifs ainsi que les faiblement positifs et les fortement positifs, comme ceux décrits plus haut, suffit. Ce type de panel peut également être utilisé pour une évaluation limitée de la reproductibilité afin d'accélérer la reconnaissance provisoire de l'essai. La méthode de test est généralement évaluée dans un ou plusieurs laboratoires disposant d'un niveau d'expérience et de compétence élevé dans des tests similaires à l'essai candidat. Le panel d'échantillons « aveugles » est évalué en utilisant l'essai candidat dans chacun de ces laboratoires, selon un protocole identique, avec les mêmes réactifs et un équipement comparable. Il s'agit d'une version réduite de l'étape 3 de la validation de l'essai. Voir le Chapitre 2.2.4 pour plus d'explications à ce sujet et sur son utilisation.

2. Reproductibilité (Norme de validation de l'OMSA, Section B.3)

La reproductibilité est un indicateur important de la précision d'un essai lorsque celui-ci est utilisé dans différents laboratoires situés dans des régions ou des pays distincts et utilisant le même essai (protocole, réactifs et contrôles). Plus le nombre de laboratoires augmente, plus le nombre de variables en relation avec l'environnement de laboratoire, les différences d'équipement et l'expertise technique augmentent également. Ces études sont une mesure de la capacité d'un essai à ne pas être altéré par des changements substantiels ou des substitutions dans les conditions de test, éléments attendus lors d'un emploi par de multiples laboratoires (p. ex. conditions d'expédition, transfert de technologies, lots de réactifs, équipement, plateformes et/ou environnement de test). Un laboratoire sur trois au moins doit tester le même panel d'échantillons « aveugles », contenant un minimum de 20 échantillons représentant les échantillons négatifs ainsi qu'un éventail d'échantillons positifs. Si des échantillons négatifs et/ou positifs sélectionnés dans le panel sont dupliqués, les estimations de la répétabilité intralaboratoire peuvent augmenter avec l'analyse répétée des échantillons utilisés dans les études de reproductibilité.

3. Essais d'aptitude (Norme de validation de l'OMSA, Section B.5.1)

Un essai validé et utilisé en routine dans plusieurs laboratoires nécessite d'être continuellement contrôlé pour garantir que sa performance est uniforme et qu'il inspire une confiance générale dans ses résultats. Cela est évalué à l'aide de programmes externes d'assurance qualité. Les essais d'aptitude constituent une mesure des compétences du laboratoire au moyen d'une comparaison interlaboratoires ; cela implique que les laboratoires participant à ces essais utilisent des méthodes de test, des réactifs et des contrôles identiques (ou similaires). Les résultats sont généralement exprimés de manière qualitative, c'est-à-dire comme négatifs ou positifs, pour déterminer la réussite ou l'échec. Pour les essais de type dilution, des résultats semi-quantitatifs fournissent toutefois des données supplémentaires pour l'évaluation d'erreurs non aléatoires dans les laboratoires concernés.

Les programmes d'essais d'aptitude varient selon le type d'essai utilisé. Pour les essais de type dilution, la taille des panels varie également, mais un minimum de cinq échantillons représentant les négatifs ainsi que les faiblement et les fortement positifs, comme ceux décrits plus haut, suffit. Les essais d'aptitude ne sont pas très différents d'une forme continue d'évaluation de la reproductibilité. Toutefois, par définition, la reproductibilité est une mesure de la performance de l'essai dans plusieurs laboratoires, tandis que les essais d'aptitude sont une

évaluation des compétences du laboratoire dans la réalisation d'un essai établi et validé. Les mesures de précision peuvent être estimées pour les données de reproductibilité et de répétabilité si des copies du même échantillon de référence sont incluses dans ce panel « aveugle ». Voir Chapitre 2.2.4 pour de plus amples explications à ce sujet et sur l'utilisation.

D. GROUPE D

Les échantillons de référence du groupe D diffèrent de ceux des autres groupes, car chaque échantillon du panel doit provenir d'un animal différent. Comme expliqué au Chapitre 2.2.8, les études avec inoculation expérimentale comprennent souvent des prélèvements répétés d'animaux à titre individuel pour déterminer la progression de la maladie, mais dans un objectif autre que celui de comparer les caractéristiques de performance associées à la sensibilité diagnostique (SeD) et à la spécificité diagnostique (SpD) de la méthode de test. Les échantillons prélevés à plusieurs reprises, à des jours différents, sur un même animal ne peuvent pas être considérés comme représentatifs d'animaux individuels dans les populations ciblées par l'essai, car ils violent la règle d'indépendance des échantillons requise pour ce genre d'études.

Un soin particulier doit être apporté au choix des échantillons de référence et à la méthode (indépendante) de référence utilisée pour ce type de comparaison afin de garantir que les analytes détectés (s'ils sont différents) présentent le même type de profil pathogénique en ce qui concerne l'apparition après exposition à l'agent infectieux et l'abondance relative dans les échantillons de test choisis.

1. Comparaison des méthodes normalisées et reconnaissance provisoire (Norme de validation de l'OMSA, Section B.2.6)

Dans certaines situations, il n'est pas possible ou pas souhaitable d'effectuer l'étape 2 du processus de validation, les échantillons appropriés provenant de la population cible étant rares ou les animaux difficiles à atteindre (comme c'est le cas pour les maladies exotiques). Un panel restreint, mais choisi, d'échantillons de test bien caractérisés représentant la plage de concentration de l'analyte doit toutefois être analysé avec la méthode de l'essai candidat, parallèlement à une méthode normalisée de l'OMSA, telle que publiée dans les *Manuels* de l'OMSA. Si les méthodes sont jugées comparables (Chapitres 2.2.8) et selon l'utilisation prévue pour l'essai, il est légitime de considérer qu'une validation diagnostique complémentaire n'est pas nécessaire. Si, par exemple, la méthode est destinée au dépistage d'agents pathogènes exotiques chez des animaux importés ou dans des produits d'origine animale, ou à la confirmation de signes cliniques, une validation complète en plus d'une comparaison des méthodes de test peut être irréalisable ou injustifiée.

L'expérience a prouvé que le nombre défini d'échantillons requis pour estimer les paramètres de performance avec un degré élevé de certitude (Norme de validation de l'OMSA, Section B.2) constituait le principal obstacle à l'étape 2 du processus de validation. Dans certains cas, une reconnaissance provisoire par les autorités internationales, nationales ou locales peut être accordée pour un essai qui n'a pas été entièrement évalué au-delà des étapes analytiques. Les différentes raisons présidant à une reconnaissance provisoire sont bien expliquées dans la Norme de validation de l'OMSA. Dans tous les cas, des preuves solides de la comparabilité des estimations de la SpD et de la SeD sur la base d'un petit panel choisi d'échantillons bien caractérisés contenant l'analyte ciblé doivent toutefois exister.

Idéalement, qu'il s'agisse de la comparaison avec une méthode normalisée ou d'une reconnaissance provisoire, un panel de 60 échantillons, par exemple, peut être réuni pour garantir une taille suffisante de l'échantillon aux fins de l'analyse statistique des données générées. Ce panel inclurait 30 « vrais » négatifs et 30 « vrais » positifs. Les positifs reflèteraient, si possible, toute la gamme de concentrations ou d'activités de l'analyte attendue dans la population cible. Comme mentionné plus haut, chaque échantillon de ce panel représenterait un seul animal. Voir le Chapitre 2.2.5 pour les méthodes statistiques servant à déterminer la comparabilité des méthodes qui utilisent des échantillons diagnostiques.

2. Modifications biologiques (Norme de validation de l'OMSA, Section B.5.2.2)

Il peut arriver que des modifications de certains produits biologiques utilisés dans l'essai soient nécessaires et/ou justifiées. Il peut s'agir de modifications des réactifs eux-mêmes ou du remplacement par un type différent d'échantillon contenant le même analyte que celui ciblé par l'essai initialement validé (p. ex. passer du sérum à la salive). Tous les critères analytiques du processus de validation devront au minimum être réévalués avant d'aller

plus loin. Si les exigences analytiques sont remplies, l'ultime question est de savoir si une validation diagnostique complète est nécessaire. Une approche similaire à celle citée plus haut, utilisant un panel de 60 échantillons de référence individuels, peut être envisagée. Dans ce cas, la méthode de test d'origine sera considérée comme le test normalisé (indépendant) et la méthode modifiée comme le test candidat. Voir le Chapitre 2.2.5 pour les approches statistiques servant à déterminer la comparabilité des méthodes qui utilisent des échantillons diagnostiques.

E. GROUPE E

Les animaux de référence et les échantillons de référence du groupe E sont bien décrits dans la Norme de validation de l'OMSA, Section B.2.1. Certains points méritent cependant d'être répétés ici.

1. « Gold standard »² – spécificité diagnostique et sensibilité diagnostique (Norme de validation de l'OMSA, Section B.2.1)

Pour les estimations conventionnelles de la SpD, les échantillons de référence négatifs désignent des échantillons vrais négatifs provenant d'animaux n'ayant eu aucune possibilité d'infection par ou d'exposition à l'agent pathogène. Dans les situations où la maladie n'a jamais été observée dans un pays ou qu'elle s'est limitée à certaines régions du pays, l'identification d'échantillons de référence vrais négatifs n'est généralement pas un problème. Lorsqu'en revanche la maladie est endémique, ce type d'échantillons peut être difficile à localiser. Il est souvent possible d'en obtenir dans certaines régions de grands pays, voire dans d'autres pays où la maladie concernée a soit été éradiquée, soit n'a jamais été observée.

De même, pour les estimations conventionnelles de la SeD, les échantillons de référence positifs désignent des échantillons vrais positifs. Il faudra s'assurer que l'échantillon de population soit bien représentatif de la population qui sera la cible de l'essai validé. Il est généralement problématique de trouver un nombre suffisant d'animaux de référence vrais positifs sur la base de l'isolement du microorganisme. Il peut être nécessaire de recourir à des échantillons d'animaux testés selon une combinaison de méthodes les classant de manière certaine comme infectés/exposés, tel que décrit dans la Norme de validation de l'OMSA.

Le point important ici est que tous les échantillons, quelle que soit leur origine, doivent être documentés comme le serait tout autre échantillon de référence, de manière à classer sans doute possible les animaux comme infectés ou exposés, selon l'adéquation à l'objectif et l'utilisation proposée du test. Comme mentionné à la Section A de ce chapitre, tous les échantillons de référence doivent être bien caractérisés. Cela comprend la documentation sur l'agent pathogène et sur l'hôte donneur. Pour les agents pathogènes, cela peut comprendre des détails relatifs à leur provenance, leur sérotype, leur génotype, leur lignée, etc. La provenance du matériel hôte doit être bien décrite quant à l'espèce, la race, l'âge, le sexe, le statut reproducteur, les commémoratifs de vaccination, les antécédents du troupeau, etc. Si possible, la phase d'infection doit être notée et inclure les détails relatifs aux signes cliniques, aux profils d'anticorps, à la charge pathogène, à l'excrétion de l'agent, etc. Dans certains cas, une infection/exposition expérimentale peut être la seule option viable pour produire le matériel de référence (voir Norme de validation de l'OMSA, Section B.2.3.). Tout ce qui précède ainsi que le protocole expérimental doivent alors être décrits en détail.

Particulièrement importants par rapport à ces échantillons de référence, les tests utilisés pour déterminer leur « véritable » statut sanitaire/infectieux doivent être bien documentés afin d'évaluer les erreurs d'estimation potentielles, susceptibles d'être reportées sur les estimations faites pour l'essai candidat. De fait, lorsque l'on utilise des essais normalisés imparfaits pour définir l'animal de référence ou le statut de l'échantillon, les estimations de la performance diagnostique (SeD et SpD) de l'essai candidat peuvent être faussées et souvent surestimées. Voir le Chapitre 2.2.5 pour les considérations statistiques.

2 Le terme « gold standard » se limite au modèle de référence absolu tel que décrit dans la Norme de validation de l'OMSA, Section B.2.1.2 et au Chapitre 2.2.5, Introduction et Figure 1.

F. GROUPE F

1. Animaux au statut inconnu – spécificité diagnostique et sensibilité diagnostique (Norme de validation de l'OMSA, Section B.2.2)

La Norme de validation de l'OMSA propose une introduction aux modèles de structure latente. Ceux-ci ne reposent pas sur l'hypothèse d'un test de référence (normalisé ou indépendant) parfait, mais évaluent plutôt l'exactitude du test candidat et de l'étalon de référence en combinant les résultats des tests. Comme ces modèles statistiques sont complexes et nécessitent des hypothèses fondamentales, une aide statistique doit être sollicitée pour orienter l'analyse et définir l'échantillonnage de la ou des populations cibles, les caractéristiques des autres tests inclus dans l'analyse, le choix de modèles appropriés et les méthodes d'estimation basées sur des publications à comité de lecture. Voir le Chapitre 2.2.5 pour les considérations statistiques.

Les populations de référence, et non les échantillons de référence individuels, utilisées dans les études de structure latente doivent être bien décrites. Cela comprend la documentation concernant l'agent pathogène ainsi que l'hôte donneur. Pour les agents pathogènes susceptibles de circuler dans la population, cela peut inclure des détails sur leur souche, leur sérotype, leur génotype, leur lignée, etc. La provenance du matériel de l'hôte doit être bien décrite quant à l'espèce, la race, l'âge, le sexe, le statut reproducteur, les commémoratifs de vaccination, les antécédents du troupeau, etc. Lorsque c'est possible, la phase d'infection dans les populations doit être notée et faire état des morbidités, mortalités, guérisons, etc.

Relevons que, lorsque les modèles de structure latente doivent être utilisés pour attribuer des valeurs estimatives à la SeD et à la SpD et que plusieurs laboratoires participent à la conception de l'essai, il est possible d'incorporer une estimation de la reproductibilité dans l'évaluation. Tel que recommandé plus haut, des conseils statistiques doivent être sollicités en la matière.

*
* *

N. B. : ADOPTE POUR LA PREMIERE FOIS EN 2014.