

MASTER 2^{ème} ANNÉE

Santé publique Paris Sud-Saclay et Santé UPEC

Dominante

**SURVEILLANCE ÉPIDÉMIOLOGIQUE DES MALADIES HUMAINES ET
ANIMALES**

RAPPORT DE STAGE

Refonte de l'outil d'évaluation des systèmes de surveillance OASIS

Présenté par

Fiona Porsan

Réalisé sous la direction de : Julie Rivière et Jean-Philippe Amat

Organisme et pays : École Vétérinaire de Maisons-Alfort

Période du stage : 8 janvier au 30 juin 2024

Date de soutenance : 24 juin 2024

Année universitaire 2023-2024

Remerciements

Je tiens à remercier chaleureusement mes co-encadrants de stage Julie Rivière et Jean-Philippe Amat pour leur encadrement, suivi et orientation tout au long du stage. Je vous remercie de m'avoir laissé une chance d'intégrer cette équipe si accueillante et motivée, c'est ce qui m'a donné la détermination de donner le meilleur de moi-même tout au long du stage. Merci d'avoir encouragé mes idées et valorisé mon travail.

Je remercie également mes encadrants secondaires, Pascal Hendriks et Renaud Laillier pour les mêmes raisons, pour leur partage de leurs connaissances sur la méthode OASIS et pour ces réunions toujours animées dans la bonne humeur.

À toute l'équipe de refonte d'OASIS, merci pour votre bonne humeur tout au long de ce stage.

Je tiens à remercier également tous les experts ayant participé et accompagné cette refonte de la méthode OASIS :

- Margot Bärenstrauch
- Marion Bordier
- Lucie Collineau
- Barbara Dufour
- Emilie Gay
- Vincent Leclerc
- Loïc Palumbo
- Anne Quillévé

Merci pour votre investissement, gentillesse, bienveillance et pour la richesse de vos idées et expertises partagées lors de ces deux tours.

Mes remerciements vont également à Barbara Dufour et Pascal Hendriks pour avoir cocréé la méthode OASIS, sans laquelle mon stage n'aurait pas eu lieu.

À mes camarades de Master ESMIHA, pour tous ces moments et rires partagés à l'ENVA ainsi qu'au CIRAD. Je suis heureuse d'avoir croisé votre route.

À ma chère amie Maelys, qui a illuminé ma dernière année de Master et dont la présence a rendu cette période inoubliable. Merci d'avoir été à mes côtés pendant ces dix derniers mois.

Un grand merci à ma famille, pour leur soutien, y compris mon petit compagnon à fourrure partie trop tôt mais dont la présence apaisante a été d'un grand soutien tout au long de mes études.

Enfin, un petit clin d'œil à toute l'équipe d'EPIMIM, de l'École vétérinaire de Maisons-Alfort/Anses, pour leur gentillesse et leur accueil chaleureux.

Résumé

Face aux défis liés à l'évolution constante des dispositifs de surveillance sanitaire, il est essentiel de disposer d'outils et de méthodes d'évaluation performants, génériques et adaptables à divers contextes de surveillance. La méthode OASIS, développée en 2011, a été largement utilisée pour évaluer la qualité du fonctionnement des dispositifs de surveillance dans les domaines de la santé animale, de la santé environnementale et de la sécurité sanitaire des aliments. Après plus de 10 ans d'utilisation, il est nécessaire de s'assurer de l'efficacité, la praticité et de l'adaptabilité de cette méthode aux besoins actuels d'évaluation. Dans ce contexte, un groupe de travail inter-plateformes de surveillance sanitaire a été créé dans le but d'entreprendre une démarche d'élicitation d'experts avec une approche de Delphi à deux tours, avec pour mission de réviser la méthode OASIS. Différents axes ont été visés : la pertinence, la clarté, le niveau de détail des éléments qui composent la méthode, les pondérations appliquées dans la notation des indicateurs et l'introduction de nouveaux éléments qui apporteraient précision à la méthode. Cette approche a permis d'identifier les points forts à conserver ainsi que les aspects à améliorer, en sollicitant l'avis des utilisateurs de la méthode par le biais d'un questionnaire et de matrices de pondérations. A l'issue de ces deux tours de Delphi, les principales évolutions adoptées incluent la suppression des pondérations entre les points critiques, la suppression des sous-catégories de points critiques et l'ajout de nouveaux indicateurs comme « Moyen », « Collaboration » ou encore « Outils Laboratoire » et « Outils Terrain ». De nouveaux critères relatifs à la surveillance intégrée et à l'impact de la surveillance ont également été pris en compte dans les pondérations. Plusieurs indicateurs ont été redéfinis pour plus de clarté, et certains ont été supprimés. Néanmoins, des aspects nécessitent encore des clarifications, tels que le changement d'échelle de pondération, la redéfinition de la pondération de certains indicateurs, ou encore l'évolution des sorties graphiques à l'image des préférences exprimées lors des deux premiers tours de Delphi. Finalement, bien que des progrès significatifs aient été réalisés, la refonte de la méthode OASIS nécessite encore des efforts pour aboutir à une version finalisée et pleinement opérationnelle répondant aux enjeux actuels d'évaluation des systèmes de surveillance. La poursuite des discussions entre experts sera cruciale pour affiner et valider les ajustements nécessaires.

Mots clés : Surveillance, évaluation, OASIS, élicitation, experts, Delphi, santé animale, santé végétale, sécurité alimentaire.

Table des matières

Introduction	- 1 -
Première partie : synthèse bibliographique	- 3 -
I. Généralités sur la surveillance épidémiologique et les dispositifs de surveillance	- 3 -
1. La surveillance épidémiologique en santé.....	- 3 -
II. Les dispositifs de surveillance	- 3 -
1. Définition et intérêt des dispositifs de surveillance.....	- 3 -
2. Fonctionnement des dispositifs de surveillance	- 4 -
3. Domaines d'utilisation des dispositifs de surveillance	- 6 -
III. L'évaluation des dispositifs de surveillance	- 7 -
1. Enjeu de l'évaluation des dispositifs de surveillance	- 7 -
2. Méthodologie de l'évaluation des dispositifs de surveillance : qualitative, quantitative, semi-quantitative et les différents éléments pouvant être évalués.....	- 8 -
3. Exemple de méthodes d'évaluation des dispositifs de surveillance	- 8 -
IV. Focus sur la méthode OASIS	- 9 -
1. La conception d'OASIS	- 9 -
2. Principe généraux de la méthode OASIS	- 10 -
3. Réflexion autour de l'évolution de la méthode Oasis.....	- 12 -
Deuxième partie : Travail personnel	- 13 -
I. Contexte problématique et objectifs	- 13 -
II. Matériels et Méthodes	- 14 -
1. Méthode d'élicitations d'experts	- 14 -
2. Questionnaire en ligne sur l'évolution globale de la méthode Oasis.....	- 15 -
3. Matrices Excel de pondération des critères participant aux indicateurs.....	- 16 -
4. Organisation des tours d'élicitation d'experts	- 18 -
III. Résultats	- 19 -
1. Points critiques	- 19 -
2. Attributs.....	- 25 -
3. Sorties graphiques	- 29 -
IV. Discussion	- 32 -
1. Discussion de la méthode	- 32 -
2. Discussion des résultats.....	- 36 -
3. Perspectives et Recommandations	- 39 -
Conclusion	- 41 -
Annexes	- 45 -

Index des tableaux

Tableau 1 : Synthèse des résultats du questionnaire en ligne sur la pertinence, la clarté et de la redondance des points critiques	- 20 -
Tableau 2 : Questionnements et propositions discutées et validées lors du 2 nd tour d'élicitation d'experts concernant les points critiques de la méthode Oasis	- 23 -
Tableau 3 : Résultats des pondérations des points critiques initiaux de la méthode Oasis par les experts sollicités, lors du premier tour d'élicitation.....	- 24 -
Tableau 4 : Nombre de critères participant aux points critiques dont la pondération fait l'objet d'un consensus, d'une faible ou d'une forte divergence entre les experts, et nombre de critères non retenus à l'issue du premier tour d'élicitation.....	- 24 -
Tableau 5 : Evolutions des pondérations proposées par les experts pour le CCP « objectifs » et décision finale validée lors du 2 ^{ème} tour d'élicitation	- 25 -
Tableau 6 : Synthèse des résultats du questionnaire en ligne sur la pertinence, la clarté et la redondance des attributs	- 26 -
Tableau 7 : Répartition des attributs proposée par cinq experts au sein des catégories d'attributs proposées dans le questionnaire en ligne	- 28 -
Tableau 8 : Synthèse des résultats sur la pertinence de la modification des sorties graphiques et propositions d'évolutions associées.....	- 31 -
Tableau 9 : Répartition de la fréquence de pondération des cinq experts ayant pondérés les indicateurs ...	- 38 -
Tableau 10 : Répartition des critères pour les trois points critiques discutés durant le second tour de Delphi	- 39 -

Index des figures

Figure 1 : Répartition du niveau de pertinence de la pondération des points critiques estimés par les experts interrogés par la méthode Delphi, à l'issu du premier tour d'élicitation (résultats du questionnaire en ligne)- 22 -

Figure 2 : Distribution de la fréquence des notes de pertinence des sorties graphiques estimées par les 12 experts répondants, sur une échelle de 0 à 10 (10 étant la note maximale)..... - 30 -

Figure 3 : Distribution de la fréquence des notes de clarté des sorties graphiques estimées par 12 experts répondants, sur une échelle de 0 à 10 (10 étant la note maximale) - 30 -

Liste des abréviations

Att : Attributs

ANSES : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail

AST : Appui Scientifique et Technique

CCP : Points critiques (*Critical Control Points*)

CDC : Centres de Contrôle et de Prévention des maladies (Centers for Disease Control and Prevention)

HACCP : Hazard Analysis Control Critical Points (Analyse des Risques et Maîtrise des Points Critiques)

OASIS : Outil d'Analyse des Systèmes de Surveillance Sanitaire

OMC : Organisation Mondiale du Commerce

QDD : Qualité des données

SA : Santé Animale

SP : Santé Publique

SV : Santé Végétale

SNAT : Surveillance Network Analysis Tool (outil d'analyse des dispositifs de surveillance)

SSA : Sécurité Sanitaire des Aliments

Introduction

La surveillance des systèmes de santé revêt une importance cruciale dans la gestion efficace des risques sanitaires et dans la prise de décisions stratégiques en matière de santé publique¹. Dans ce contexte, les outils et méthodes d'évaluation des systèmes de surveillance jouent un rôle essentiel en permettant d'évaluer la performance, l'efficacité, et la cohérence des objectifs de ces systèmes. Parmi ces méthodes, OASIS (Outil d'Analyse des Systèmes de surveillance Sanitaire), élaboré en 2011, a été largement utilisé dans les domaines de la santé animale², végétale et de la sécurité des aliments pour évaluer la qualité et le fonctionnement des systèmes de surveillance dans différents contextes et pays.

Dès lors, avec l'évolution constante des besoins en matière de surveillance et l'avancée des technologies et de certains concepts (telle que « Une seule santé »), il est devenu impératif de réviser et de mettre à jour cet outil afin de garantir sa pertinence, son utilité dans les pratiques actuelles, en tenant compte des particularités et des problématiques propres à chaque domaine de santé et en s'appuyant sur l'expérience des évaluations déjà réalisées avec OASIS. C'est dans ce cadre que s'inscrit le présent mémoire de stage de Master 2, qui porte sur la refonte de la méthode OASIS.

L'objectif principal de ce travail était de solliciter l'expertise des utilisateurs et des spécialistes de la méthode OASIS afin d'identifier les points forts de cet outil à conserver ainsi que les aspects à améliorer. Pour ce faire, une démarche d'élicitation d'expert a été entreprise, en utilisant la méthode Delphi en deux tours. Cette approche avait pour but de recueillir des avis et des recommandations variés, en mettant l'accent sur la pertinence, l'efficacité et la praticité d'OASIS dans le contexte actuel de la surveillance sanitaire.

Ce mémoire commencera par un état des lieux de la méthode OASIS, en mettant en lumière ses principes, ses objectifs et ses domaines d'application. Ensuite, la démarche méthodologique sera décrite en détail, en expliquant les différentes étapes de l'élicitation d'expert par la méthode Delphi. Enfin, la méthode et les résultats obtenus lors de cette élicitation seront discutés, en mettant en évidence les aspects consensuels et discordants mis en évidence au cours de la refonte de la méthode OASIS.

¹ **OMS, 1952** : « la santé publique est la science et l'art de prévenir les maladies, de prolonger la vie et d'améliorer la santé et la vitalité mentales et physiques des individus par le moyen d'une action collective [...] ».

² **Académie vétérinaire de France** : « La santé publique vétérinaire est l'ensemble des actions collectives, principalement régaliennes, en rapport avec les animaux sauvages ou domestiques, leurs services et leurs productions entrant notamment dans la chaîne alimentaire, qui visent à préserver les santé humaine et animale - y compris l'état de bien-être - et la santé des écosystèmes. Elle contribue ainsi au développement durable et à la mise en œuvre du concept "Une seule santé" ».

Première partie : synthèse bibliographique

I. Généralités sur la surveillance épidémiologique et les dispositifs de surveillance

1. La surveillance épidémiologique en santé

La surveillance épidémiologique, également connue sous le nom d'épidémiosurveillance, se définit comme une « méthode d'observation fondée sur des enregistrements en continu, permettant de suivre l'état de santé ou les facteurs de risque d'une population définie, en particulier de déceler l'apparition de processus pathologiques et d'en étudier le développement dans le temps et dans l'espace, en vue de l'adoption de mesures appropriées de lutte » (Dufour, Hendrikx and Toma, 2005). De cette définition ressortent trois notions principales : l'épidémiologie descriptive¹, le processus continu et la perspective d'action (Dufour, Hendrikx and Thonnat, 2011).

La surveillance épidémiologique est déterminante dans la mesure de l'état de santé des populations dans le temps et l'espace afin d'en cerner les besoins, priorités et d'identifier les groupes les plus vulnérables, et dans la mesure des résultats des actions de santé publique mises en place, pour évaluer leur efficacité et juger de leur pertinence (*Mesure de l'état de santé des populations et notification des données*, 2000). Elle permet également de cerner les facteurs de risques des maladies ou dangers étudiés, et les déterminants de santé des populations, champ qui relève davantage de l'épidémiologie analytique.

Les mouvements généralisés des animaux et de leurs produits à travers le monde, notamment avec la mise en place de l'Organisation mondiale du commerce (OMC) en 1995, ont entraîné un risque croissant de propagation des maladies (Dufour, Hendrikx and Thonnat, 2011). La surveillance épidémiologique s'est donc considérablement développée au cours de ces dernières années pour mieux suivre ces risques. Elle a contribué à de nombreuses évolutions, notamment dans la connaissance de l'état sanitaire et de la situation épidémiologique des pays, et dans l'élaboration de politiques de santé. Il est devenu nécessaire de favoriser l'échange entre pays d'informations comparables sur les maladies présentes ou potentiellement émergentes sur ces territoires (Hoinville *et al.*, 2013). Cette transparence repose sur une compréhension commune des approches de surveillance et de mise en œuvre de systèmes de surveillance.

Finalement une surveillance efficace permet de réduire l'impact des maladies sur la production agricole et la santé, de faciliter la confiance entre pays et la circulation des animaux, des plantes et des denrées et d'assurer une protection efficace des population animales et humaines contre les maladies exotiques ou émergentes.

II. Les dispositifs de surveillance

1. Définition et intérêt des dispositifs de surveillance

En parallèle de l'avancée de la surveillance épidémiologique, la création de systèmes de surveillance s'est intensifiée et se poursuit encore aujourd'hui. Le terme dispositif de surveillance peut être confondu avec le terme « système » ou encore celui de « réseau » de surveillance, soulignant ainsi

¹ **Larousse** : « L'épidémiologie descriptive a pour objectifs d'identifier et de décrire les problèmes de santé et d'évaluer leur importance dans des populations. Pour étudier la fréquence et la répartition de ces problèmes, elle établit des taux, c'est-à-dire qu'elle rapporte un nombre de malades à l'effectif d'une population. ».

l'importance de bien comprendre les définitions et leurs distinctions (Dufour, Hendrikx and Thonnat, 2011).

Un réseau de surveillance peut être défini par l'ensemble des personnes ou organismes structurés assurant la surveillance sur un territoire donné d'une ou de plusieurs entité(s) pathologique(s) (Dufour, Hendrikx and Thonnat, 2011). En revanche, le terme « système de surveillance » désigne plutôt un ensemble de réseaux de surveillance de la même maladie ou du même danger (par exemple le système de surveillance des salmonelles en France) (Dufour, Hendrikx and Thonnat, 2011).

Le terme plus général de « dispositif de surveillance », englobe finalement toutes activités de surveillance épidémiologique, qu'elles reposent sur des systèmes, plans ou toute autre initiative de surveillance sans nécessairement être intégrées à des réseaux structurés (Dufour, Hendrikx and Thonnat, 2011). Le terme « dispositif » sera utilisé dans la suite du manuscrit.

Face à une société en constante évolution, où les risques de développement ou d'émergence de nouvelles maladies sont croissants, il est impératif de mettre en place des dispositifs de surveillance adaptés permettant de gérer efficacement ces risques. Chaque activité de surveillance doit être planifiée et mise en œuvre selon des méthodes appropriées à ses objectifs de surveillance et à la maladie en question. Les dispositifs de surveillance doivent donc fonctionner de manière à permettre, dans l'idéal, la production de données de qualité pour suivre de manière fiable l'état sanitaire d'une zone et prendre les décisions de lutte appropriées. Ainsi, il est essentiel de définir clairement l'objet de la surveillance, les données à collecter, la population concernée, ainsi que les modalités de surveillance (active, passive ou mixte) les mieux adaptées aux objectifs fixés.

La surveillance événementielle aussi appelée « surveillance passive », se définit comme toute activité de surveillance reposant sur une déclaration spontanée et volontaire des cas ou suspicions de maladies surveillées (Dufour, Hendrikx and Thonnat, 2011). Elle ne nécessite pas de sollicitation spécifique ou d'acteurs de terrain pour assurer l'obtention des données et vise à permettre une alerte précoce en cas d'apparition ou de réapparition de dangers sanitaires. En revanche, la surveillance programmée, également appelée « active », implique des actions programmées à l'avance, qui peuvent être élaborées par l'animateur du réseau, cellule d'animation ou encore par un comité scientifique et technique (Dufour, Hendrikx and Thonnat, 2011). Contrairement à la surveillance événementielle, il faut définir à l'avance le nombre, la nature et la localisation des données qui seront collectées par le réseau de surveillance.

La surveillance syndromique constitue quant à elle une approche innovante fondée sur des indicateurs de santé non spécifiques y compris des signes cliniques, des symptômes ainsi que des mesures indirectes utilisant des informations pré-diagnostiques qui proviennent de différentes sources (telles que les bases de données électroniques de santé) (Rodríguez-Prieto *et al.*, 2015). Elle permet de fournir des indications plus précoces sur une épidémie plutôt que d'attendre un diagnostic confirmé, et de minimiser les principales limites des méthodes de surveillance événementielles sur la confirmation en laboratoire et par les cliniciens.

Ainsi, ces différentes approches de surveillance, qu'elles soient événementielles, programmée ou syndromique, contribuent toutes à la détection précoce, à la gestion efficace et à la prévention des maladies, assurant ainsi la protection des populations humaines et animales, mais il reste nécessaire de faire en sorte d'évaluer la capacité d'un dispositif de surveillance à atteindre ces objectifs.

2. Fonctionnement des dispositifs de surveillance

La mise en place d'un dispositif peut se découper en plusieurs étapes : établir des objectifs ; élaborer des définitions de cas ; déterminer les sources de données et les mécanismes de collecte de données (type de dispositif) ; déterminer les instruments de collecte de données ; tester les méthodes sur le

terrain, développer les méthodes d'analyse des données ; développer les méthodes de diffusion de l'informations (Dufour, Hendrikx and Thonnat, 2011).

Les dispositifs de surveillance fonctionnent tous de manière différente et spécifique aux objectifs fixés, et à la(aux) maladie(s) surveillée(s). C'est pourquoi ces objectifs doivent être définis précisément avant leur lancement afin qu'ils soient en mesure de mettre en place des modalités de surveillance adaptées. Ils peuvent autant consister en la détection de l'apparition d'une maladie exotique ou d'une nouvelle sur un territoire afin de mettre en place le plus rapidement possible des mesures de lutte que de démontrer le statut exempt d'infection des animaux et des produits dérivés des animaux dans une zone donnée. On peut également chercher à déterminer l'importance réelle d'une maladie (incidence, prévalence, perte économiques, etc...) et suivre son évolution (évaluation des résultats d'un plan de lutte mise en place, évaluation de "l'importance" d'une maladie), afin de choisir d'entreprendre ou non une lutte appropriée ou de modifier les modalités de la lutte. De même, il peut s'agir de l'établissement de priorités d'action, impliquant une hiérarchisation des maladies en fonction de leur importance économique ou sanitaire au sein d'une population donnée.

Une fois l'objectif et la nécessité d'un dispositif de surveillance déterminé, les méthodes pour obtenir, analyser, diffuser et utiliser les informations doivent être définies et mises en œuvre. Un dispositif de surveillance repose sur plusieurs choses : Tout d'abord une collecte de données de santé sur les maladies, les blessures et les facteurs de risque, qui peuvent être obtenues de diverses manières notamment sur terrain (Dufour, Hendrikx and Thonnat, 2011). Ces données sont par la suite centralisées (au niveau local, régional et/ou national selon les dispositifs), analysées, puis servent à mieux connaître la situation épidémiologique du pays/région/population et à guider les décideurs afin d'entreprendre les mesures de luttés appropriées. Ces étapes reposent sur des outils permettant de collecter les données, mesurer la fréquence des maladies, d'analyser mais également de communiquer entre les différents acteurs de terrain. Mais également sur des acteurs de terrain, locaux, ou encore centraux/nationaux, pouvant avoir différents rôles spécifiques à la collecte de données ex : vétérinaires, médecins), à l'analyse des données (ex : laboratoires, épidémiologistes, etc...), ou encore à l'organisation du réseau (ex : animateur de réseau) (Dufour, Hendrikx and Thonnat, 2011).

Les dispositifs de surveillance sont continus et nécessitent donc la coopération et l'échange des acteurs entre eux. Ils doivent être acceptés par ceux qui y participent, suffisamment flexibles pour répondre aux besoins évolutifs de la communauté ainsi qu'aux changements dans les tendances des maladies et des dangers, et doivent fournir des informations en temps opportun pour permettre une action rapide. Il est essentiel de trouver un équilibre entre toutes ces considérations afin de concevoir un dispositif qui répond aux besoins identifiés sans devenir trop coûteux ou contraignant.

Il est aussi important de souligner le rôle significatif de l'organisation, de l'engagement et de la formalisation des relations entre les acteurs à différents niveaux du dispositif. Cela repose sur des activités de communication (circulation multidirectionnelles des informations, ascendantes, descendantes et transversales), d'information et de formation.

Comme mentionné précédemment, clarifier les définitions des termes utilisés en surveillance épidémiologique est essentiel pour éviter toute confusion et promouvoir une communication efficace au sein du réseau mais aussi l'échange d'information et l'évaluation des activités de surveillance à travers différents secteurs, industries et pays (Hoinville *et al.*, 2013). L'établissement de normes et de standards pour ces termes renforce ainsi la confiance dans la validité des données de surveillance et encourage l'engagement des parties prenantes en garantissant une cohérence et une uniformité dans leur utilisation.

3. Domaines d'utilisation des dispositifs de surveillance

L'expansion au fil des années des dispositifs de surveillance sous-entend leurs développement dans divers domaines, qu'ils s'agissent de la santé publique (SP), animale (SA), végétale¹ (SV), ou de la sécurité sanitaire des aliments² (SSA) (Calba *et al.*, 2015). Ils occupent une place centrale des préoccupations actuelles One Health³.

Parmi ces domaines d'application, les dispositifs de surveillance sont essentiels pour suivre l'état sanitaire des pays pour des maladies infectieuses bactérienne (salmonellose, tuberculose, ...), virales (grippe aviaire, VIH, rage) ou encore non infectieuses (telles que le cancer, les maladies cardiovasculaires...). Certains dispositifs sont exclusifs à un domaine particulier, tandis que d'autres peuvent concerner différents domaines de manière décloisonnée, c'est-à-dire en favorisant la collaboration et l'échange d'informations entre eux. Par exemple, un dispositif intégré peut permettre aux experts en santé publique, en santé animale et végétale de travailler ensemble pour détecter et répondre rapidement à une épidémie de zoonose. Cette approche interdisciplinaire et collaborative permet de mieux comprendre et gérer les risques pour la santé dans une perspective globale pour protéger la santé publique, animale et végétale.

Les interactions étroites entre les humains et les animaux, ainsi que les risques sanitaires associés aux maladies d'origine animale (zoonoses, maladies vectorielles, etc.), mettent en évidence la frontière étroite qui subsiste entre ces deux monde. Ces enjeux sont particulièrement pertinents dans le contexte contemporain actuel, marqué par l'urbanisation et des pandémies telles que la COVID-19.

La santé végétale est également un domaine clé où les dispositifs de surveillance sont utilisés pour évaluer les risques liés à l'environnement, non seulement pour la santé humaine, mais aussi pour la santé des écosystèmes. Cela inclut la surveillance de paramètres tels que la qualité de l'air, de l'eau et du sol, ainsi que l'exposition aux produits chimiques et toxiques (Smith, 2019).

Enfin la surveillance dans le domaine de la sécurité sanitaire des aliments pour surveiller la présence d'agents pathogènes, de contaminants, de toxines microbiennes, de résidus de pesticides tout au long de la chaîne alimentaire, garantissant ainsi la protection des populations contre les épidémies et les infections d'origine alimentaire ('Systèmes de contrôle des aliments', no date). Les salmonelles constituent la deuxième cause d'infection d'origine alimentaire en Europe ('Qu'est-ce que la salmonellose et comment s'en prémunir?', 2022), soulignant ainsi l'importance cruciale de la surveillance en sécurité alimentaire dans la prévention et le contrôle de telles maladies.

La nécessité de diversifier les dispositifs de surveillance est en constante évolution, car elle vise à étendre sa couverture à un large éventail de maladies, de populations et de régions, y compris dans les pays en développement. Cette diversification est cruciale pour répondre aux défis spécifiques rencontrés dans ces contextes, tels que le coût des outils de collecte de données ou encore les contraintes climatiques et territoriales. De même, il est essentiel d'évaluer la qualité de fonctionnement et la fiabilité des informations produites d'un dispositif, pour assurer qu'il produise des données utiles pour éclairer les mesures de gestion et de lutte. Il est également impératif de réfléchir à des améliorations pour garantir leur efficacité à long terme.

¹ **CORDIS** : « La santé des végétaux est un terme englobant les risques émergents, dont les organismes nuisibles, les maladies et les mauvaises herbes, la lutte intégrée contre les organismes nuisibles et l'innovation dans la protection des végétaux ».

² **Ministère de l'Agriculture et de la Souveraineté animales** : « La sécurité alimentaire nécessite le contrôle des denrées alimentaires d'origine animale et végétale, sur toute la chaîne commerciale, depuis leur production jusqu'à la remise au consommateur en passant par toutes les étapes de l'élaboration des denrées alimentaires ».

³ **OMS définition** : « approche qui vise la conception et la mise en œuvre de programmes, de politiques, de législations et de recherche impliquant une communication et une collaboration entre de multiples secteurs en vue d'obtenir de meilleurs résultats en matière de santé publique ».

III. L'évaluation des dispositifs de surveillance

1. Enjeu de l'évaluation des dispositifs de surveillance

La planification de la surveillance est un processus itératif nécessitant une réévaluation régulière de ses objectifs et de ses méthodes. La question fondamentale à répondre lors de l'évaluation est de savoir si les objectifs du dispositif de surveillance ont été atteints, et que pourrait-on faire pour améliorer son fonctionnement (organisation, traitement des données ou encore communication) et ses attributs (simplicité, flexibilité ou encore acceptabilité).

La qualité des données est essentielle dans le cadre de la surveillance, car elle garantit la précision et la fiabilité des informations utilisées pour prendre des décisions éclairées. Des données de mauvaise qualité peuvent mener à des diagnostics erronés, des réponses inappropriées et une gestion inefficace des risques sanitaires. Plusieurs facteurs influencent la qualité des données, notamment la précision des outils de collecte, la pertinence des informations collectées, ou encore la perception des utilisateurs quant à l'utilité du dispositif (ou encore des outils ou des données). Par conséquent des évaluations régulières sont cruciales dans le maintien d'un dispositif de surveillance en marche.

L'évaluation de la surveillance peut être définie comme une « évaluation systématique et objective de la pertinence, de l'adéquation, des progrès, de l'efficacité, de l'efficience et de l'impact d'un ensemble d'actions d'un réseau de surveillance, par rapport à des objectifs prédéfinis et en tenant compte des ressources et des installations déployées par le système » (Dufour, Hendrikx and Thonnat, 2011). Elle joue également un rôle essentiel dans l'établissement et le maintien de la confiance et de la crédibilité des dispositifs de surveillance auprès des décideurs, professionnels de santé, du public, mais également des autres dispositifs de surveillance (collaboratifs ainsi qu'au niveau international).

Ces dernières années, plusieurs méthodes ont été développées pour évaluer les dispositifs de surveillance, permettant d'analyser le dispositif dans son ensemble (méthode générique comme la méthode OASIS) ou certains aspects spécifiques (tels que l'organisation, la collaboration ou les coûts). Ces évaluations peuvent s'appuyer sur des approches variées, qu'elles soient qualitatives, quantitatives ou encore semi-quantitatives. Parmi les méthodes et outils d'évaluation, une méthode développée il y a plusieurs années et reconnue internationalement, celle des CDC guidelines (Centers For Disease Control and Prevention, aux Etats Unis)), développe des recommandations et des directives utilisées pour standardiser et améliorer divers aspects de la santé publique. Dans le contexte de la surveillance épidémiologique, les CDC guidelines fournissent des critères et des méthodes pour évaluer l'efficacité des dispositifs de surveillance (incluent des aspects tels que la simplicité, la flexibilité ou encore l'acceptabilité). Elle est souvent utilisée et intégrées dans d'autres méthode et outils de surveillance tel que la méthode OASIS.

Le choix d'une méthode d'évaluation est une étape préliminaire visant à définir et à choisir quel aspect du dispositif de surveillance on souhaite évaluer. Deux approches complémentaires peuvent être envisagées pour évaluer la qualité de fonctionnement des dispositifs surveillance épidémiologique. D'une part l'évaluation interne des dispositifs, qui repose sur l'élaboration et le calcul d'indicateurs de performance et de diagnostic (*Principes d'évaluation des réseaux de santé*, 1999). Pilotée par l'animateur du dispositif en collaboration avec les acteurs, cette approche permet un suivi continu de son fonctionnement. D'autre part, une évaluation externe peut être réalisée par des membres extérieurs (donc neutres vis-à-vis du dispositif), et permet d'identifier les points forts et faibles du dispositif et de proposer des recommandations pour son amélioration. Contrairement à l'évaluation interne, l'évaluation externe est effectuée ponctuellement et ne repose pas sur des mesures d'indicateurs en continu mais ponctuels (*Principes d'évaluation des réseaux de santé*, 1999).

La collecte d'informations est une phase descriptive de l'évaluation, elle se doit d'être le plus organisée et standardisée possible. Elle englobe à la fois la collecte de données sur l'objet de l'évaluation (le dispositif de surveillance en lui-même) mais également sur les performances et les résultats qui lui sont associées (organisation, informations produites et diffusion de l'information).

Enfin, l'étape d'évaluation ou d'analyse des informations collectées permet de comparer ce que l'on souhaite évaluer avec un élément de référence : un gold standard, (dans l'idéal, si notre dispositif de surveillance était le plus optimal possible). Puis s'ensuit une étape de recommandations permettant de formuler des propositions d'amélioration dans le but de faire un lien avec l'objectif initial de l'évaluation du dispositif.

2. Méthodologie de l'évaluation des dispositifs de surveillance : qualitative, quantitative, semi-quantitative et les différents éléments pouvant être évalués

L'évaluation des réseaux de surveillance épidémiologique peut reposer sur différentes approches : qualitative, quantitative ou semi-quantitative.

L'évaluation qualitative, repose sur la collecte des données permettant de décrire et de comprendre le fonctionnement du dispositif de surveillance en profondeur, sous forme d'entretiens, d'observations sur le terrain ou d'analyses documentaires. Elle permet d'obtenir des données sur les processus et sur les pratiques des acteurs du dispositifs (Bronner *et al.*, 2015). L'évaluation quantitative repose quant à elle sur le recueil de données numériques pour évaluer la performance et la qualité du dispositif de surveillance (Bronner *et al.*, 2015). Cette méthode utilise ainsi des indicateurs quantitatifs pour mesurer divers aspects du dispositif (comme la sensibilité, la spécificité ou encore la rapidité), ce qui permet une évaluation plus objective et mesurable de la performance du dispositif. Entre ces deux types d'évaluation se trouve une approche intermédiaire, l'évaluation semi-quantitative, qui combine des aspects qualitatifs et quantitatifs. Cette méthode repose généralement sur des échelles de notation ou des scores pour évaluer différents aspects d'un dispositif de surveillance. Elle implique souvent une collecte d'information fondée sur des entretiens semi-directif et des questionnaire avec les coordinateurs et acteurs du dispositif. On retrouve dans cette catégorie la méthode OASIS (Partie IV) ou encore l'outil EcoSur (qui évalue l'organisation et le fonctionnement des collaborations multisectorielles dans un dispositif de surveillance).

Plusieurs aspects d'un dispositif de surveillance peuvent être évalués : des aspects techniques, tels que l'exécution des tâches au sein du réseau, la formalisation des rôles de chaque acteur, ainsi que tout ce qui touche à la collecte de données. Des aspects organisationnels comme la coordination, la répartition et la gestion des tâches au sein du dispositif. Des aspects sociaux pouvant inclure la communication entre les membres au sein du dispositif et avec des acteurs externes, ainsi que la diffusion et la transmission des informations. Des aspects liés aux résultats, se concentrant sur l'efficacité de la surveillance et la qualité des données produites, ainsi que l'impact global des actions entreprises par le réseau. En dernier lieu, des aspects plus précis peuvent être examinés, notamment la rapidité, la collaboration, la représentativité ou encore l'acceptabilité, qui constituent des éléments plus ciblés du dispositif.

3. Exemple de méthodes d'évaluation des dispositifs de surveillance

Plusieurs méthodes d'évaluation des dispositifs de surveillance ont vu le jour ces dernières années. Elles recouvrent toutes différents aspects de la surveillance et possèdent leurs spécificités permettant de choisir la plus appropriées à l'évaluation d'un dispositif de surveillance en question.

Certaines méthodes peuvent être semi-quantitatives, basées sur une série de « critères¹ » ou « attributs² » selon l'appellation choisie, noté (sur une échelle définie) lors d'entretiens par exemple, permettant le calcul de sorties graphiques ou de résultats quantitatifs. Leur objectif est d'identifier les points forts et les aspects à améliorer du dispositif en question. On compte parmi ces méthodes, EcoSur, un outil permettant d'évaluer la collaboration multisectorielle des dispositifs de surveillance (Bordier *et al.*, 2019) (à l'aide de 23 attributs organisationnels), la méthode OASIS qui permet d'évaluer le fonctionnement global d'un dispositif de surveillance (Hendrikx *et al.*, 2011), (à l'aide de 78 critères, répartis en 10 sections) ou encore l'outil OH-EpiCap, qui permet d'évaluer les capacités et mécanismes de coordination multisectorielle (Tegegne *et al.*, 2023) (à l'aide de 30 indicateurs). Il existe également des méthodes qualitatives, telles que la méthode SERVAL (Drewe *et al.*, 2015), une méthode générique permettant d'évaluer les dispositifs de surveillance en santé animale (à l'aide de 22 attributs) ayant la particularité d'inclure des critères socio-économique ou encore la méthode Surf Surveillance Evaluation Framework (Muellner *et al.*, 2018), qui permet de fournir un cadre commun pour la surveillance dans les secteurs sanitaire animale et environnemental.

IV. Focus sur la méthode OASIS

1. La conception d'OASIS

L'évaluation des dispositifs de surveillance nécessite une approche standardisée et flexible, permettant son application dans la plupart des situations de surveillance épidémiologique (réseau local, régional, national), dans divers domaines sanitaire, pour et pour une importante variété de dispositifs. C'est dans cette optique que l'outil OASIS (Outil d'Analyse des Systèmes de Surveillance) a été développé par une équipe multidisciplinaire de l'ANSES, (Agence Nationale de la Sécurité Sanitaire) regroupant des épidémiologistes, développeurs et utilisateurs de méthodes d'évaluation des dispositifs de surveillance (Hendrikx *et al.*, 2011).

La méthode OASIS créée en 2010 (parfois appelée « outil OASIS »), résulte de la combinaison de plusieurs méthodes d'évaluation existantes : un outil fondé sur l'évaluation qualitative de la réalisation des activités clés de réseaux de surveillance notamment utilisé dans les Caraïbes, une méthode semi-quantitative fondée sur des points de contrôle critiques de réseaux de surveillance, et une méthode qualitative standardisée développées par les Centre de contrôle des maladies (CDC - Center for Disease Control) et l'OMS (Organisation Mondiale de la Santé) fondé sur des attributs tels que la sensibilité, la rapidité ou encore l'acceptabilité (Hendrikx *et al.*, 2011).

Si l'on revient plus en détail sur ces différentes méthodes de surveillance qui composent OASIS :

- La première méthode appelée SNAT (Surveillance Network Assessment Tool), est une méthode d'audit semi-quantitative, développée dans les Caraïbes par 4 épidémiologistes puis mise à jour en 2007 par un groupe du réseau de surveillance régional CaribVET. La méthode SNAT se divise en deux phases distinctes : la première phase consiste en un inventaire détaillé des structures et des procédures du réseau de surveillance épidémiologique, tandis que la deuxième phase présente un résumé de la situation du réseau dans ses principaux domaines d'activité. Un questionnaire en 10 sections est utilisé pour décrire le dispositif de surveillance, chaque section étant résumée par 4 critères et représentée graphiquement par des camemberts, au total 10 (Hendrikx *et al.*, 2011).

¹ Larousse : « Principe, élément de référence qui permet de juger, d'estimer, de définir quelque chose ».

² Larousse : « Terme indiquant une manière d'être ou une qualité attribuée au sujet ou à l'objet [...] ».

- La deuxième méthode est celle des “Critical Control Point”, développée par Barbara Dufour en 1997. Elle repose sur les principes de l’HACCP¹ (Hazard Analysis Control Critical Points) et son objectif est d’identifier les points critiques du fonctionnement d’un réseau de surveillance en santé animale associé à des critères, permettant ainsi d’évaluer sa qualité. Elle recense huit points critiques (objectifs, échantillonnage, programme de sensibilisation, outils, collecte et circulation des données, environnement, interprétation et traitement des données, diffusion des informations), dont 1 (environnement) qui n’est intégré que pour les maladies exotiques. Leur degré de satisfaction est apprécié à l’aide d’un score attribué à chaque critère grâce à des questionnaires qui permettent de récolter les informations nécessaires.
- Enfin la dernière méthode de surveillance est fondée sur des lignes directrices publiées en 1988 par le CDC aux Etats-Unis. Cette méthode recense 10 attributs (sensibilité, spécificité, représentativité, rapidité, flexibilité, fiabilité, stabilité, acceptabilité, simplicité, utilité) qui permettent d’évaluer l’efficacité d’un dispositif dans l’accomplissement de ses objectifs. D’autres organisations comme l’OMS et Canadian Public Health Organization se fondent sur des listes de critères assez similaires.

2. Principe généraux de la méthode OASIS

2.1 Contenu de la méthode OASIS

Cette approche d’évaluation semi-quantitative repose tout d’abord sur un questionnaire, rempli soit par le biais d’entretiens avec les acteurs du réseau (dans le cadre d’OASIS approfondi), soit par une analyse documentaire (dans le cadre d’OASIS flash, une application web d’OASIS permettant une évaluation dans un laps de temps plus court). Ce questionnaire permet de collecter diverses informations pour décrire et comprendre en profondeur le fonctionnement du dispositif de surveillance évalué. Ces informations servent ensuite à remplir une grille de notation, divisée en dix sections composées de 78 critères. Chaque section est synthétisée par un nombre variable de critères (de 4 à 14 critères) qui font chacun l’objet d’une notation, de 0 à 3, en fonction du degré de satisfaction/conformité estimé pour le dispositif étudié (3 correspondant à une situation “idéale” et 0 à une situation non-idéale, soulignant ainsi certaines faiblesses du dispositif). Cette notation est réalisée en se référant à un guide détaillé qui explicite les conditions d’attribution de chaque note, assurant ainsi un certain niveau de standardisation. Lorsqu’un critère considéré n’est pas pertinent pour le dispositif, il est déclaré et codifié “SO” (sans objet), ce qui garantit une certaine flexibilité dans l’utilisation d’OASIS et son adaptabilité à des dispositifs de surveillance pouvant être très divers et concerner différents secteurs de la santé.

La notation des critères peut être accompagnée de commentaires détaillant et justifiant les choix de la notation effectués. De plus, à la fin de chaque notation de section, un commentaire général reprend les principaux commentaires des sections et permet de mettre en lumière des recommandations visant à améliorer le fonctionnement du dispositif. Ainsi, lorsque l’ensemble des notes ont été attribuées à chacun des critères, leur intégration dans un tableur de type Excel (qui se fait de manière automatique dans une feuille de calcul), permet de représenter et d’analyser les résultats de l’évaluation à travers trois types de représentations graphiques (cf. annexe 1). Ces graphiques mettent en lumière les points forts et les points à améliorer du dispositif de surveillance. Chaque sortie graphique est représentée par une combinaison spécifique de critères notés, faisant l’objet d’une synthèse, et chaque résultat

¹ *FDA, Food and Drug Administration (« Agence américaine des produits alimentaires et médicamenteux ») ; HACCP est une approche systématique pour l’identification, l’évaluation et le contrôle des dangers liés à la sécurité alimentaire, basée sur les sept principes suivants : Réaliser une analyse des dangers, Déterminer les points de contrôle critiques (PCC), Établir des limites critiques, Établir des procédures de surveillance, Établir des actions correctives, Établir des procédures de vérification, Établir des procédures d’enregistrement et de documentation.*

chiffré est mesuré par une proportion (exprimée en pourcentages), calculée en faisant le rapport entre le score obtenu et le score maximum possible.

La première sortie graphique (cf. annexe 1) est composée de dix camemberts (ou graphiques en secteur), un par section fonctionnelle, (Objectifs et champ de la surveillance, Organisation institutionnelle centrale, Organisation institutionnelle de terrain, Laboratoire, Outils de surveillance, Modalités de surveillance, Gestion des données, Formation, Communication, Évaluation), ce qui permet de représenter le niveau de satisfaction du fonctionnement du dispositif dans son ensemble ainsi que d'identifier les points d'amélioration du dispositif de surveillance. Cette sortie est basée sur la méthode SNAT. Chaque camembert représente, pour chaque section, la note obtenue par rapport à la note maximale attribuable pour la section.

La seconde sortie (cf. annexe 1) se présente sous forme d'histogrammes, composé de 8 barres (une barre par point critique) et permet d'identifier le niveau de contrôle de 8 points critiques¹ d'un dispositif (objectifs, l'échantillonnage, environnement, animation, outils, recueil et la circulation des données, traitement et l'interprétation et diffusion de l'information) avec une pondération² spécifique répartie sur 100. Certains points critiques ont un poids plus important que d'autres : les outils et l'échantillonnage ont chacun une pondération de 20, tandis que les objectifs et l'animation ont chacun une pondération de 15 et le reste possède une pondération de 10. Les points critiques sont particulièrement utiles pour faire des recommandations par la suite, puisqu'ils permettent de savoir directement sur quels aspects critiques il faut agir pour améliorer des aspects essentiels du dispositif de surveillance (les priorités d'amélioration du dispositif).

La troisième sortie (cf. annexe 1), sous forme de « radar chart » ou graphique en toile d'araignée, est constituée de 10 attributs (sensibilité, spécificité, représentativité, rapidité, flexibilité, fiabilité, stabilité, acceptabilité, simplicité et utilité), (une branche/axe par attributs), représentant le niveau de qualité du dispositif de surveillance. Cette sortie est fondée sur les lignes directrices élaborées par le CDC et l'OMS permettant de visualiser plus facilement les points forts et les points d'amélioration d'un dispositif. Un résultat numérique est également associé à chaque attribut, correspondant à la combinaison des notes de chaque critère d'intérêt, associées à une pondération indiquant l'importance relative de chaque critère dans la constitution de l'attribut.

2.2 Méthode de calculs (pondérations, pourquoi pondérer ?)

La méthode OASIS utilise un processus de pondération des points critiques entre eux (participant à la largeur des barres d'histogramme de la sortie 2) et des critères qui participent à la notation des attributs. La pondération des critères permet de souligner la différence de poids relative des critères dans la définition de l'attribut, car chaque critère ne possède pas la même importance et n'influence pas de la même manière les différents attributs. Dans le cadre de la méthode OASIS elle permet de préciser et d'équilibrer l'évaluation.

2.3 Equipe d'évaluation

Afin de mener à bien l'évaluation, il est essentiel de créer une équipe mixte de notation, composée d'environ 4 personnes : au moins deux évaluateurs externes au dispositif évalué (par exemple un épidémiologiste avec de l'expérience dans la surveillance épidémiologique, et une personne qui maîtrise la maladie ou le danger sanitaire surveillé), et au moins un auditeur interne³ au dispositifs (si possible le coordinateur du dispositif de surveillance). Cette complémentarité dans cette équipe mixte est utile à la fois pour avoir un point de vue interne et une bonne connaissance du dispositif évalué, mais également pour pouvoir prendre du recul sur celui-ci. Il faut noter qu'en cas de désaccord sur

¹ **Reverso** : provient de la méthode HACCP, et constitue des principes permettant d'évaluer les points de fonctionnement cruciaux d'un dispositif de surveillance.

² **Larousse** : « Procéder à la pondération des variables entrant dans le calcul d'un indice, d'une note ».

³ **Pages Jaunes** : « L'audit externe est une activité de contrôle et de conseil menée par une personne extérieure à une organisation ».

une note, un commentaire ou une recommandation, c'est l'avis des évaluateurs externes qui l'emporte.

3. Réflexion autour de l'évolution de la méthode Oasis

Des difficultés dans la compréhension de termes, dans la formulation et définition de certains critères ainsi que dans l'adaptation de la méthode à certains dispositifs spécifique ont été soulignées. De même, un manque de transparence sur le calcul des points critiques et attributs a été mentionné par certains évaluateurs. Cette prise de conscience, ainsi que le souhait d'élargir l'utilisation de la méthode OASIS à d'autres secteurs que celui de la santé animale, a conduit à la création d'un groupe de suivi inter plateformes en santé animale, végétale, et sécurité sanitaire des aliments, portant sur l'évaluation de dispositifs de surveillance (GS EDS). Au sein de ce groupe de suivi, plusieurs groupes de travail ont vu le jour, notamment : un groupe de travail visant à réaliser des retours d'expérience sur la méthode OASIS, un groupe de travail visant à réviser le vocabulaire de la méthode pour l'adapter à différents types de dispositifs, ou encore un groupe de travail visant à revoir en profondeur la méthode OASIS, (ses points critiques, attributs, méthodes de calcul et sorties graphiques). D'autres groupes de travail plus spécifiques, portant sur des aspects économiques ou collaboratifs entre dispositifs de surveillance, ont par ailleurs identifié depuis 2020 des critères à ajouter à la méthode OASIS (section 11 « surveillance intégrée » et critères « impacts »), qui ont par ailleurs été intégrer aux réflexions autour de la révision en profondeur d'OASIS. Toutefois, il est essentiel de constamment évaluer toute méthode afin d'en améliorer les résultats, la qualité des données produites et par conséquent le bon fonctionnement des dispositifs de surveillance. Après une dizaine d'années d'utilisation, il semblait nécessaire d'initier une réflexion sur l'évolution de la méthode OASIS et d'apporter de manière exhaustive les mises à jour les plus avancées possibles à cette méthode.

Deuxième partie : Travail personnel

I. Contexte problématique et objectifs

Depuis sa création en 2011, la méthode OASIS (Outil d'Analyse des Systèmes de surveillance Sanitaire) a été un outil précieux pour l'évaluation des dispositifs de surveillance dans divers domaines, tels que la santé animale, la santé environnementale et la sécurité alimentaire : en permettant d'identifier les points forts et les faiblesses des dispositifs de surveillance évalués, facilitant leur amélioration continue, en essayant ainsi de garantir une gestion efficace des risques sanitaires. Toutefois, après plus d'une décennie d'utilisation, plusieurs défis et points faibles de la méthode ont été identifiés, rendant nécessaire une révision approfondie de l'outil pour maintenir sa pertinence et son utilité dans un contexte de surveillance en constante évolution.

En effet, les utilisateurs de la méthode OASIS ont signalé plusieurs difficultés. Parmi celles-ci figurent la complexité de certains termes et éléments de la méthode, (tel que les sous-catégories de points critiques) l'adaptation limitée de la méthode à certains dispositifs spécifiques, et le manque de transparence dans le calcul des points critiques et des attributs.

Face à ces enjeux, un groupe de suivi inter-plateformes a été créé, réunissant des experts des secteurs de la santé animale, végétale, et de la chaîne alimentaire. Ce groupe a entrepris plusieurs initiatives, notamment la collecte de retours d'expérience sur l'utilisation d'OASIS, la révision du vocabulaire de la méthode pour l'adapter à divers secteurs d'application, et une refonte en profondeur des points critiques, des attributs et des sorties graphiques associées. Ce stage de Master 2 s'inscrit dans le cadre de ce dernier groupe de travail, avec pour mission principale la refonte de la méthode OASIS.

Le présent mémoire s'intéresse ainsi à la refonte de la méthode OASIS en s'appuyant sur l'expertise des utilisateurs et des spécialistes, afin d'améliorer son utilité et sa pertinence dans les pratiques actuelles de surveillance sanitaire. La problématique à laquelle cette étude répond est ainsi la suivante : Quelles seraient les améliorations possibles permettant de favoriser le maintien de la pertinence et de l'utilité de la méthode OASIS dans divers domaines de santé et pour divers types de dispositifs de surveillance ? Quelles approches seraient à adopter pour en améliorer les résultats et la qualité des données produites ? Pour atteindre cet objectif global, plusieurs sous-objectifs spécifiques ont été définis :

- **Evaluer la pertinence des indicateurs actuels (points critiques, attributs), leur clarté et les sorties graphiques associées ;**
- **Réviser les pondérations actuellement attribuées aux points critiques et attributs ;**
- **Identifier des pistes d'amélioration et d'évolutions de la méthode.**

II. Matériels et Méthodes

1. Méthode d'élicitations d'experts

1.1 Méthode Delphi

Dans le cadre de ce travail de refonte, une méthode par éléicitati6n d'experts a été retenue. Cette approche a été choisie pour plusieurs raisons. Tout d'abord, la méthode OASIS se veut générique et adaptable à divers dispositifs et secteurs de la santé. En intégrant des experts issus de divers domaines tels que la santé animale, la santé humaine et la chaîne alimentaire, il a été possible de recueillir des avis et des points de vue spécifiques à chaque secteur. Cette diversité d'expertise avait pour objectif de prendre en compte les particularités et les problématiques propres à chaque domaine, ce qui était essentiel pour ajuster les besoins et réviser les pondérations de manière adéquate. Ainsi, l'objectif était de développer une version 2.0 de la méthode OASIS qui soit la plus polyvalente et applicable possible, indépendamment de la nature du dispositif étudié. Le mot « indicateurs » pourra être utilisé pour désigner les points critiques ET attributs dans la suite de ce mémoire.

La méthode Delphi a été jugée comme la méthode d'éléicitati6n d'experts la plus adaptée. Elle correspond à une méthode de facilitation de groupe visant à obtenir un consensus¹, en recueillant les opinions d'experts sur une question de recherche ou un sujet spécifique (Boulkedid *et al.*, 2017). Elle se déroule généralement à travers une série de tours de consultation (sous forme de questionnaires anonymes ou de réunions) permettant, dans l'idéal, d'arriver à une convergence des opinions des experts sollicités. Le processus itératif sur plusieurs tours, avec un retour d'information aux participants fondé sur l'agrégation de leurs réponses, facilite le développement d'un consensus et permet d'affiner les idées progressivement tout au long de l'éléicitati6n. Elle peut être considérée comme une méthode mixte car elle permet d'intégrer à la fois des données qualitatives et quantitatives, pouvant être analysées en parallèle. Elle se distingue par sa flexibilité, permettant d'ajuster les méthodes utilisées à chaque tour en fonction des contraintes et des objectifs de leur mise en œuvre (par exemple, en fonction des disponibilités des experts, si les rencontres en présentiel ne sont pas envisageables, il est possible de privilégier des questionnaires en ligne). Cette adaptabilité permet de moduler l'approche méthodologique selon les besoins et les contextes particuliers de sa mise en œuvre.

Au premier tour, l'expert est sollicité pour répondre à divers sujets/questions ; puis, au(x) tour(s) suivant(s), chacun reçoit les résultats agrégés du groupe d'experts (résultats quantitatifs complétés éventuellement par les commentaires anonymisés des participants) et un rappel de sa propre réponse. Cela permet à chaque expert de réévaluer son degré d'accord avec chaque assertion, en confrontant ses réponses avec celles de l'ensemble du groupe, également de préparer au mieux les sujets et problématiques qui doivent être discutées lors des tours suivants. L'expert a alors la possibilité de conserver l'avis émis au premier tour, ou de le modifier en justifiant l'évolution de sa réflexion.

Dans le cadre de la révision de la méthode OASIS, nous avons souhaité appliquer la méthode Delphi avec deux tours d'éléicitati6n d'experts :

- Le premier tour avait pour objectif de solliciter les experts sur leur avis général sur la méthode OASIS (clarté, facilité d'utilisation, articulation entre points critiques et attributs, sorties graphiques...) à travers un questionnaire en ligne, ainsi que sur les critères contribuant aux 8 points critiques et 10 attributs évalués par la méthode et leurs pondérations associées, à travers des matrices Excel à compléter. Ce premier tour était réalisé individuellement par chaque

¹ Larousse : « Accord et consentement du plus grand nombre [...] ».

expert, sans possibilité de communication entre eux, et s'est étalé 2 semaines, du 8 au 22 mars 2024 ;

- Le second tour avait pour objectif de réunir les experts au travers de deux réunions qui ont eu lieu en distanciel sur la plateforme Teams, les 29 avril et 2 mai 2024, d'une durée de 3h chacune. Planifiée à l'aide de l'outil Doodle pour convenir des dates et horaires qui conviendraient au plus grand nombre d'experts, il n'a toutefois pas été possible de réunir l'ensemble des experts lors de ces deux réunions. Ainsi un total de 10 personnes ont participé à la première réunion, et 9 à la seconde. Ces réunions avaient pour but de valider les quelques décisions pouvant être prises à l'issues du premier tour, et surtout, discuter de manière collégiale des résultats du premier tour faisant l'objet de fortes divergences entre les experts, l'objectif étant d'atteindre un consensus, à la fin du second tour.

1.2 Critères de sélection des experts

Le choix des experts à solliciter a été guidé par deux critères principaux : (i) leur connaissance approfondie de la méthode OASIS, traduisant une certaine expertise dans l'évaluation de dispositifs de surveillance ou dans l'application de cette méthode ; et (ii) leurs domaines d'activité (santé animale, santé végétale et sécurité de la chaîne alimentaire), afin d'assurer une bonne représentativité des principaux secteurs d'application de la méthode. Ainsi, une liste de 15 experts a été établie, et leur sollicitation s'est faite au travers de courriers électroniques (mails).

Cette approche a été adoptée dans le but de recueillir une quantité suffisante d'informations, pour une analyse approfondie et pour pouvoir recueillir l'avis d'experts de profils variés (Booto Ekionea, Bernard and Plaisent, 2011), pouvant représenter différentes situations, et ainsi identifier les besoins majeurs d'évolution de la méthode et les points de consensus et de divergence concernant les modifications pouvant être envisagées. L'implication des experts était jugée comme cruciale pour favoriser l'acceptation et l'intégration de la méthode dans leurs futures activités d'épidémiosurveillance. Il convient de souligner l'importance d'intégrer les créateurs de la méthode dans ce panel, afin de bénéficier de leurs opinions et explications concernant les choix ayant conduit à l'élaboration de la version initiale de la méthode (appelée version 1.0). De plus, cela permet d'assurer que toute évolution soit réalisée avec leur approbation.

2. Questionnaire en ligne sur l'évolution globale de la méthode Oasis

2.1 Objectifs

L'objectif de ce questionnaire était de recueillir l'avis général des experts sur la méthode OASIS, en évaluant sa clarté, sa facilité d'utilisation, et en identifiant des points d'évolution souhaités par les experts. Un focus particulier a été mis sur les trois grands types de résultats produits par la méthode : les points critiques, les attributs et les sorties graphiques. Des questions spécifiques pour chaque type de résultat visaient à obtenir les avis des experts sur leur clarté, leur pertinence, et la possibilité de les conserver tels quels ou d'envisager des évolutions, telles que la reformulation, la clarification des définitions, ou leur suppression.

2.2 Elaboration et administration du questionnaire

Afin d'obtenir les avis des experts sur la méthode OASIS dans son ensemble et sur les évolutions à envisager, un questionnaire en ligne a été créé et paramétré avec le logiciel Kobotoolbox. Il comprenait à la fois des questions ouvertes et fermées, à choix simple et à choix multiples, afin de collecter des données directement triées et facilement exploitables dans différents formats.

Le questionnaire comprenait un total de 27 questions, et était structuré en trois parties (cf. annexe 2) : « **Points critiques** », « **Attributs** » et « **Sortie graphiques** ».

- Dans les deux premières parties, les experts étaient invités à se prononcer sur la clarté, la compréhensibilité, la pondération, ou encore la redondance des points critiques et des attributs. Ils devaient également évaluer la pertinence de conserver des points critiques et attributs dans la méthode, en précisant pour chacun d'entre eux s'il était correctement défini ou nécessitait d'évoluer (par exemple, suppression) en justifiant leurs choix lorsque possible. Des questions conditionnelles permettaient d'explorer la pertinence de maintenir une pondération entre les points critiques, ou d'introduire une pondération entre les attributs entre eux (actuellement absente dans la version 1.0), avec des tableaux leur permettant de répartir, dans ce cas, des pondérations sur 100 entre les CCP et attributs. Des questions spécifiques permettaient d'aborder la pertinence de distinguer les CCP et attributs dans les sorties graphiques, et de manière générale dans la méthode, de l'utilité de la sous-catégorisation des CCP, de la possibilité d'introduire une catégorisation pour les Att, de donner leurs avis sur des suggestions d'amélioration.
- Dans la partie sortie graphiques, les experts devaient se positionner sur leur niveau de compréhensibilité, leur pertinence, redondance, les évolutions possibles, ainsi que sur la suffisance des détails qui leur sont associés. Les questions sur le niveau de détail et de compréhension des sorties graphiques étaient présentées sous forme d'échelles de 0 à 10.

Pour chaque question fermée, une zone de commentaires libres permettait aux experts de justifier leur choix, d'apporter des précisions et de proposer des pistes d'évolution de la méthode.

2.3 Analyses

2.3.1 1^{er} tour d'élicitation d'experts

Un document synthétique comprenant les résultats du questionnaire en ligne a été élaboré (Booto Ekionea, Bernard and Plaisent, 2011), les réponses ayant été regroupées par thématique d'analyse (sous-parties « Pertinence et clarté », « Redondance », « Suggestions/Nouveautés/Propositions », « Pondérations » au sein de parties « Point critiques », « Attributs », « Sorties graphiques »). A la fin de chaque partie, un bilan des questionnements était proposé sous forme d'encadré, ce qui a permis de hiérarchiser et de sélectionner les problématiques par ordre de priorité à discuter lors du 2nd tour de l'élicitation d'experts (notamment les points ne faisant pas consensus).

2.3.2 2^{ème} tour d'élicitation d'experts

La discussion de la partie questionnaire a été abordée de la manière suivante : une présentation synthétique des résultats consensuels a été faite en début de réunion, afin de permettre une exploration plus approfondie des points de divergence et problématiques classées par ordre d'importances pour chaque partie. Durant la réunion, les experts étaient encouragés à indiquer leur niveau d'accord avec les résultats et à proposer de nouvelles idées.

3. Matrices Excel de pondération des critères participant aux indicateurs

3.1 Objectifs

L'objectif des matrices Excel de pondération était de recueillir toutes les informations relatives à la notation des indicateurs, notamment les critères considérés ainsi que leurs pondérations. De plus, elle permettait de recueillir les avis et opinions des experts sur les indicateurs eux-mêmes, ainsi que sur la pertinence des critères. Cette phase était essentielle, car la notation des critères et leur compilation pour estimer des indicateurs (CCP ou attributs) constitue le cœur même de la méthode OASIS. En

effet, ces données sont cruciales pour le calcul des résultats et sorties graphiques, qui servent de fondement aux recommandations de l'évaluation des dispositifs de surveillance.

3.2 Elaboration des matrices de pondération

Lors du premier tour, des matrices Excel (cf. annexe 3) de pondération ont été envoyées aux experts, accompagnées d'une notice explicative.

Elles avaient pour objectifs de recueillir plusieurs informations :

- La pertinence des critères, notamment s'ils étaient suffisamment pertinents (à conserver tel quel), s'ils devaient être modifiés (renommés/reformulés) ou supprimés ;
- Les critères devant entrer dans la pondération de chaque CCP et attribut ;
- La pondération associée à ces critères ;
- La suggestion éventuelle de nouveaux critères ;
- Les commentaires globaux liés aux CCP, attributs ou critères, avec des colonnes et une lignes dédiées permettant aux experts de fournir librement leurs observations.

Les matrices ont été construites de la manière suivante : une ligne pour chaque critère de choix pondéré, une colonne pour chaque éléments étudiés (CCP et Att). La grille "type" utilisée dans la version 1.0 de la méthode, a été reprise dans les matrices, en y intégrant de nouveaux critères et sections développées récemment dans les travaux du groupe de suivi EDS (notamment critères "surveillance intégrée", et section "impacts").

3.3 Administration et modalités de pondération

Pour chacun des 8 CCP et 10 attributs existants, les experts devaient décider, pour chaque critère sélectionné, s'il devait être pris en compte dans le calcul de l'indicateur. Dans ce cas, ils devaient proposer un poids à attribuer à ce critère, parmi les poids suivants : 1, 2, 5 ou 10 (10 étant le poids le plus important pouvant être attribué et une case vide correspondant au poids nul), en fonction de l'importance qu'ils attribuaient à chaque critère. Les critères participant à chaque indicateur de la version 1.0 de la méthode étaient repérables grâce à un fond de case coloré. Les experts avaient donc la possibilité, dans cette révision de la méthode, de les choisir à nouveau (dans ce cas en y ajoutant une pondération) ou de ne pas les inclure.

De même, il leur été possible de proposer de nouveaux points critiques ou attributs (dans ce cas en indiquant les critères qui les composaient) ou nouveaux critères, ainsi qu'une pondération associée.

Chaque expert devait réaliser cet exercice d'identification des critères et de leur pondération pour six indicateurs (points critiques ou attributs), permettant ainsi une répartition équilibrée du travail entre les experts. Les indicateurs ont été répartis en amont entre les experts afin d'assurer la participation de 4 à 7 experts par indicateur, garantissant ainsi une représentation suffisante des opinions consensuelles ou divergentes. De plus, les experts avaient la possibilité de travailler sur un nombre supérieur d'indicateurs s'ils le souhaitaient.

3.4 Analyses

3.4.1 1^{er} tour d'élicitation d'experts

Suite de la collecte des réponses des experts dans le cadre du premier tour d'élicitation, des matrices de synthèse (« Matrices récapitulatives ») ont été élaborées sous Excel, afin de regrouper et synthétiser les avis d'experts. Ainsi, chaque onglet correspondait à un indicateur (CCP ou attribut), dans lequel étaient regroupées les réponses des experts pour cet indicateur lors de ce 1^{er} tour d'élicitation (4 et 7 réponses par indicateur).

À partir des réponses des experts, une pondération moyenne a été calculée, représentant ainsi l'avis moyen des experts concernés. À partir de cette moyenne, une pondération finale a été proposée, correspondant au poids s'approchant le plus de la valeur moyenne parmi les valeurs 1, 2, 5, 10. Toutefois, cette approche via une pondération moyenne ne permettait pas de représenter le degré de consensus obtenu pour chaque critère : ainsi, un code couleur a été associé à chaque pondération, permettant de visualiser le degré de consensus/discordance entre les experts :

- La **couleur verte** correspondait à un « consensus » fort, et était attribuée dans deux cas :
 - écart d'un rang entre les valeurs de pondération proposées par les experts et aucune valeur égale à zéro¹,
 - aucun écart entre les valeurs de pondération et aucune égale à zéro.
- La **couleur orange** correspondait aux "faibles divergences", attribuée dans deux cas :
 - écart de deux ou trois rangs entre les valeurs de pondération et aucune égale à zéro²,
 - écart d'un ou deux rangs entre les valeurs de pondération et au moins une de zéro.
- La **couleur rouge** correspondait aux "fortes divergences", attribuée si :
 - écart de quatre rangs entre les valeurs de pondération et aucune égale à zéro³,
 - écart d'au moins trois rangs entre les valeurs de pondération et au moins une de zéro.
- La **couleur jaune fluo** était attribuée lorsqu'il n'y avait aucun ou seulement un expert qui ayant attribué une pondération au critère, et lorsque ce critère était initialement pris en compte dans la pondération de l'indicateur dans la version 1.0 de la méthode OASIS (cases à fond coloré).

Dans les cas où seul un expert avait attribué une pondération à un critère comme participant à un indicateur (quelle que soit sa valeur), ce critère n'a pas été pris en compte dans la version 2.0 et n'a donc pas été discuté lors du second tour.

Des graphiques sous forme de diagramme en barres, (cf. annexe 4) classés par ordre décroissant afin de visualiser facilement les critères qui a priori participaient le plus à chaque indicateur ; ont été élaborés et également présentés dans ces deux "Matrices récapitulatives", suite au 1^{er} tour.

3.4.2 2^e tour d'élicitation d'experts

La partie matrice a été discutée en deuxième réunion de second tour. Elle a été abordée de la manière suivante : un rappel des consignes et de l'objectif des matrices a été effectué en premier lieu, suivi d'une présentation de la méthodologie d'analyse des matrices, avec une synthèse des résultats (issus du questionnaire) faite avec des diagrammes en barres, permettant une visualisation d'ensemble avant d'aborder les problématiques soulevées par les matrices. Ensuite, chaque indicateur a été abordé un à un, en se basant sur le code couleur établi et présenté dans la partie 3.4.1, seuls les critères faisant l'objet d'une importante divergence (code couleur associé rouge) ont été discutés avec les experts. Ces points étaient présentés dans des tableaux (pour chaque CCP et Att), reprenant l'ensemble des avis des experts de manière anonymisés et une pondération moyenne.

4. Organisation des tours d'élicitation d'experts

Au vu de la complexité des matrices de pondérations, il a été essentiel d'élaborer un guide (cf. annexe 5) suffisamment détaillé afin de permettre une compréhension claire et précise des matrices et du questionnaire par les experts. Ce guide a été essentiel afin d'orienter au mieux les experts dans la complétion des documents administrés. Il détaillait donc les différentes parties des matrices, lignes et

¹ Par exemple 1-2 ou 5-10

² Par exemple 2-10

³ Par exemple 1-10

colonnes présentées et les consignes sur la pondération des critères ; et présentait également les différentes parties du questionnaire en ligne.

Pour garantir une compréhension uniforme des indicateurs (points critiques et attributs) entre les experts, un lexique, inexistant dans la version 1.0 de la méthode, a été élaboré, proposant ainsi pour la première fois une définition des CCP et attributs évalués par la méthode OASIS (cf. annexe 6). Ont également été intégrées dans ce lexique les définitions des nouveaux attributs proposés dans le questionnaire en ligne (tel que qualité des données (QDD), interopérabilité ou encore communication par exemple). Les définitions ont été réfléchies de sorte à être les plus objectives possible, à pouvoir s'appliquer à chaque domaine d'utilisation de la méthode OASIS, et à ne pas trop se fier aux critères entrant en compte dans la version 1.0 des points critiques et attributs.

III. Résultats

Dans cette section, les résultats issus du premier et du second tour de la méthode Delphi sont présentés. Sur un total de 15 experts, sollicités, 12 ont finalement participé et ont retourné l'ensemble des documents nécessaires à l'étude. Les domaines de santé répartis de la manière suivante : 7 experts en santé animale, 4 en sécurité sanitaire des aliments et 1 en santé végétale.

1. Points critiques

1.1 Pertinence, Clarté, Redondance

Concernant la pertinence et la clarté des points critiques, les résultats du questionnaire en ligne ont montré que 66 % (8/12) des experts n'éprouvaient pas de difficulté dans la compréhension des CCP lors de la réalisation d'une évaluation OASIS. Parmi les CCP existants, « Outils » et « Recueil et circulation des données » sont ceux ayant reçu le plus de demande de reformulation et de redéfinition (respectivement 50% (6/12 experts) et 41 % (6/12 experts)) ; les avis sont plus partagés pour le maintien et la formulation du CCP « Environnement », un expert proposant de le supprimer ou de l'intégrer au CCP « Echantillonnage ». La synthèse des résultats du questionnaire en ligne sur ces thématiques est présentée dans le tableau 1.

1.2 Propositions d'ajout et de redéfinition des points critiques

La moitié des experts interrogés ont proposé des suggestions d'ajout de CCP, notamment :

- **Moyens** (n=3), qui intégrerait l'adéquation des moyens (qu'ils soient humains, financiers ou techniques) alloués au dispositif pour fonctionner, au regard des objectifs poursuivis ;
- **Collaboration** (n=2), qui concernerait la collaboration au sein du dispositif ainsi qu'à l'extérieur du dispositif ;
- **Qualité des données** (n=1), qui pourrait correspondre à l'exhaustivité, la validité et à la cohérence des données enregistrées en adéquation avec les besoins et objectifs du dispositif ;
- **Organisation institutionnelle** (n=1), qui pourrait porter spécifiquement sur l'organisation institutionnelle du dispositif, non intégrée dans le CCP actuel "Animation" ;
- **Collecte et stockage des données** (qui pourrait regrouper deux CCP actuellement existants, "Recueil des données" et "Traitement et Interprétation des données") (n=1).

Finalement plusieurs consensus ont pu être obtenus à propos des problématiques soulevées lors du premier tour et discutées lors des réunions du second tour : le CCP « Outils » a été scindé en « Outils Laboratoire » et « Outils Terrain » ; le CCP « Recueil et circulation des données » a été reformulé en « Collecte et Qualité des données », permettant d'intégrer la notion de « Qualité des données » proposée au premier tour en tant qu'attribut. De même, le CCP « Traitement et interprétation des données » a été renommé en « Analyse et interprétation des données ». La suppression du CCP « Environnement » a également été décidée, ainsi que l'ajout du CCP « Moyens », inexistant jusqu'à présent.

Tableau 1 : Synthèse des résultats du questionnaire ligne sur la pertinence, la clarté et de la redondance des points critiques

CCP	Pertinence				Clarté	Redondance
	à conserver	à reformuler	à supprimer	Sans réponse		
Objectifs	75 % (9/12)	8 % (1/12)	8 % (1/12)	8 % (1/12)		CCP redondant avec section 1 (n=1)
Echantillonnage	66 % (8/12)	25 % (3/12)	0 %	8 % (1/12)	Parler davantage des modalités de surveillance et inclure l'environnement , qui ne concerne pas que la santé animale et la faune sauvage (ex. environnement d'atelier de production en SSA).	CCP en partie redondant avec certains attributs -"Représentativité" (n=1) -"Recueil et circulation des données" (n=1) → se concentrer sur la stratégie d'échantillonnage dans le CCP « Échantillonnage » et renvoyer tout ce qui est en lien avec prélèvements et recueil des données dans le CCP "Recueil et circulation des données".
Animation	66 % (8/12)	25 % (3/12)	0 %	8 % (1/12)	- Seconde partie de la définition (mise en œuvre d'actions, manière dont les initiatives sont planifiées, réalisées dans le temps et qualité de celles-ci) confuse et peu claire - Proposition de ne pas inclure les éléments relatifs au pilotage et à l'AST. Si on veut maintenir ces deux éléments, renommer le point critique en « Gouvernance ».	
Environnement	25 % (3/12)	25 % (3/12)	16 % (2/12)	33 % (4/12)	Peu utilisé et ne semble pas correspondre à la définition d'un point critique. Plusieurs propositions : - à relier au objectifs du dispositif ? - à inclure dans le CCP Échantillonnage ? - A supprimer ? - ou à garder mais expliciter davantage ce qu'il doit contenir (lien avec dérèglement climatique ?). Il ne comporte que 2 critères actuellement, ce qui pose la question de l'homogénéité avec les autres CCP. Actuellement ce CCP ne cible que la	

					faune sauvage, l'écosystème ; il faut considérer également les autres secteurs (pas uniquement la SA), tel que proposé dans le lexique. - à simplifier ? pourrait être inclus dans un en CCP "modalités de surveillance" qui pourrait englober « Environnement » et « Echantillonnage », si on allait jusqu'à supprimer les notes de section (diagramme camembert - sortie 1).	
Outils	41 % (6/12)	50 % (6/12)	0 %	8 % (1/12)	L'expression "Analyse de données" est ambiguë, on ne sait pas s'il s'agit de l'analyse d'échantillons au laboratoire ou de l'interprétation des résultats, à clarifier. A séparer en deux CCP, en distinguant par exemple « outils de laboratoire » vs « outils de terrain » (n=2). Préciser qu'il s'agit des outils de prélèvements, mesure, laboratoires, mais pas des outils d'analyse de données ou de communication par exemple.	CCP qui concerne notamment la pertinence et qualité des outils de traitement et interprétation des données = en chevauchement avec le pt critique "Traitement et interprétation des données".
Recueil et circulation des données	50 % (6/12)	41 % (5/12)	0 %	8 % (1/12)	Clarifier si ce CCP évalue également la qualité des données.	
Traitement et interprétation des données	75 % (9/12)	16 % (2/12)	0 %	8 % (1/12)		
Diffusion de l'information	66 % (8/12)	16 % (2/12)	8 % (1/12)	8 % (1/12)	Clarifier si la qualité de l'information diffusée est aussi évaluée ou non par ce CCP.	CCP redondant avec la section 9 « Communication ».

1.3 Pondérations des points critiques

Actuellement, les huit points critiques d'OASIS sont pondérés entre eux, ce qui détermine la largeur des barres les représentant dans la sortie 2 (histogramme). L'objectif de cette pondération est de refléter l'importance de chaque point critique dans l'évaluation d'un dispositif de surveillance. A l'issue de l'analyse du questionnaire en ligne, la majorité des experts ont répondu ne pas considérer la pondération des CCP comme pertinente (Figure 2), car ils ne l'utilisent pas pour hiérarchiser les recommandations pour 50 % (6/12 experts) d'entre eux, ou pour d'autres raisons pour 40 % (5/12) d'entre eux, citées ci-après : les pondérations ne ressortent pas clairement dans la méthode et complexifient la lecture graphique en imposant une largeur différente des histogrammes de la sortie 2 (différence qui n'est généralement pas comprise), ce qui amplifie par ailleurs l'incertitude associée à l'analyse des résultats. Pour deux experts, les points critiques sont par définition critiques, donc tous d'importance majeure pour tous les dispositifs de surveillance, et ne devraient pas être hiérarchisés entre eux. Enfin, le poids qu'il faudrait attribuer aux CCP peut varier en fonction du type de surveillance et de son contexte de mise en œuvre (cela peut par exemple dépendre des ressources humaines et financières que le dispositif a à disposition pour mettre en place les recommandations) (n=2), ce qui complexifierait le choix d'une pondération "type", difficile à analyser par ailleurs.

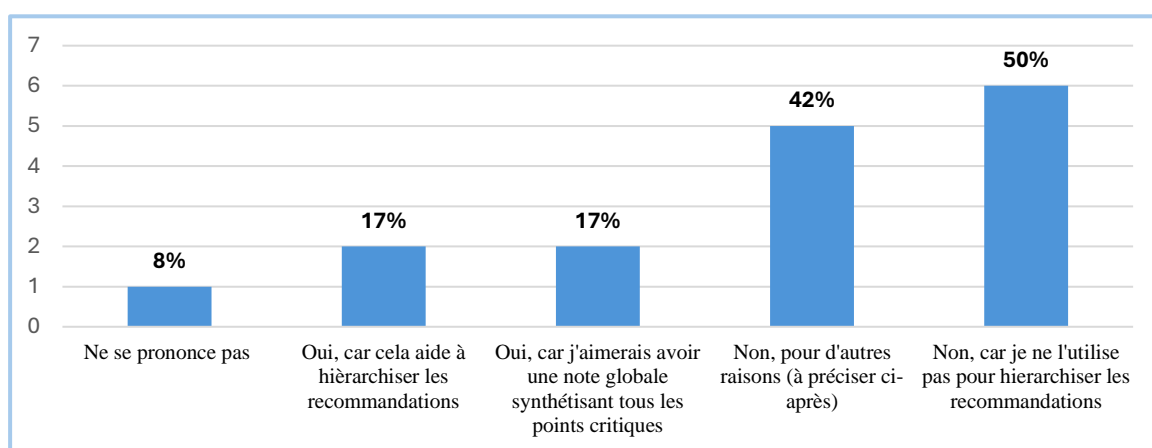


Figure 1 : Répartition du niveau de pertinence de la pondération des points critiques estimés par les experts interrogés par la méthode Delphi, à l'issue du premier tour d'élicitation (résultats du questionnaire en ligne)

1.4 Sous-catégories de Points critiques

Au total 59 % (7/12) des experts ont indiqué ne pas vouloir conserver les sous-catégories de points critiques initialement proposées, visibles uniquement dans la matrice de calcul de la méthode OASIS. Plusieurs raisons ont été citées : les sous-catégories sont considérées comme non pertinentes et non indispensables car les experts ne les utilisent pas pour faire des recommandations à la suite d'une évaluation (n=3), comme non transparentes (n=1), le détail du calcul et leur prise en compte étant peu explicité dans la méthode (n=3). Certains experts ne savaient d'ailleurs pas que ces sous-catégories existaient (n=2), leur utilité semble donc limitée (n=2). Enfin, un expert a indiqué que la méthode Oasis devait rester la plus simple possible pour être utilisée par le plus grand nombre (n=1).

1.5 Questionnements et propositions d'évolutions des points critiques

Certains résultats présentés ci-dessus ont permis de générer des problématiques sur la reformulation, la redéfinition ou encore l'ajout et la suppression de CCP, discutées lors du 2nd tour.

Les problématiques qui ressortent du questionnaires en ligne ont été discutées pour six CCP : « Outils », « Objectifs », « Animation », « Recueil et circulation des données », « Traitement et interprétation des données » et « Environnement ».

La synthèse des questionnements sur les CCP discutés ainsi que les propositions retenues sont présentées dans le tableau 2 ci-dessous.

Tableau 2 : Questionnements et propositions discutées et validées lors du 2nd tour d'élicitation d'experts concernant les points critiques de la méthode Oasis

CCP	Problématiques discutées lors du 2 nd tour	Propositions retenues
Objectifs	Redondant avec la section 1 (avec par ailleurs une méthode de calcul très proche) : est-il pertinent de conserver également un CCP sur ce thème ?	A conserver en tant que CCP malgré sa redondance avec la section 1, qui sera potentiellement renommée après avoir validé l'ensemble des critères retenus. La section 1 serait à renommer en : "Cadre de la surveillance", à la place de "Objectifs et champ de la surveillance".
Animation	Pertinence d'élargir ce CCP en intégrant également la "Gouvernance" ?	Les experts n'étaient pas favorables à la proposition d'élargir et de renommer le CCP "Animation", en intégrant des notions relatives à la "Gouvernance", et n'étaient pas favorables à la création d'un nouveau CCP "Gouvernance". La notion de gouvernance est par ailleurs déjà intégrée dans la section 2 "Organisation institutionnelle centrale", et n'est ainsi pas aussi critique que l'animation. → il est donc décidé l'intégration de notions de gouvernance (critères de la section 2) dans le CCP "Animation", mais avec de faibles poids, et l'intégration du terme "Gouvernance" dans le titre de la section 2.
Environnement	Faut-il supprimer ce CCP ou le reformuler ?	Suppression du CCP "Environnement"
Outils	Redondant avec "Traitement et interprétation des données" → A dichotomiser en deux CCP (Outils de laboratoire et de terrain) ?	Création de 2 nouveaux CCP : « Outils Laboratoire » et « Outils Terrain », au sein desquels les critères seront à redistribuer.
Recueil et circulation des données / Traitement et interprétation des données	Proposition de regroupement de ces CCP en un seul CCP, qui pourrait s'intituler "Collecte, gestion et interprétation des données".	Reformulation des 2 CCP : - "Recueil et circulation des données" → "Collecte et Qualité des données" ; - "Traitement et interprétation des données" → "Analyse (ou exploitation ?) et interprétation des données"
Moyens/Aspects Economiques	Faut-il considérer l'intégration de nouveaux CCP : Collaboration ; QDD ; Moyens ; Organisation institutionnelle ?	Ajout du CCP "Moyens" Intégration de la notion de QDD au CCP nouvellement reformulé "Collecte et circulation des données"

1.6 Révision des critères participants aux points critiques via les matrices de pondérations

1.6.1 Résultats du premier tour

L'analyse des matrices de pondérations a permis de mettre en évidence la proportion de « cases colorées » sélectionnées par les experts (c'est-à-dire critères sélectionnés lors de la révision de la méthode, qui étaient déjà pris en compte dans la version 1.0), la proportion de critères avec une pondération concordante, faiblement ou fortement discordante entre les experts, ainsi que la proportion de critères sélectionnés par un seul expert et donc non retenu dans la pondération de la version 2.0 (Tableau 3).

Tableau 3 : Résultats des pondérations des points critiques initiaux de la méthode Oasis par les experts sollicités, lors du premier tour d'élucidation

Points critiques	Nombre d'experts ayant pondéré le CCP	Pourcentage de critères inclus dans la version initiale de la méthode sélectionnés à nouveau par les experts	Pourcentage de critères nouvellement pondérés par les experts (non inclus dans la version 1.0)
Objectifs	6	100 % (5/5)	72 % (13/18)
Echantillonnage	6	100 % (8/8)	78 % (28/36)
Environnement	6	100 % (2/2)	33 % (1/3)
Animation	6	100 % (19/19)	62 % (31/50)
Outils	6	100 % (21/21)	56 % (27/48)
Recueil et circulation des données	6	100 % (15/15)	76 % (47/62)
Traitement et interprétation de l'information	7	100 % (10/10)	69 % (22/32)
Diffusion de l'information	6	100 % (4/4)	91 % (40/44)

Les résultats présentés sur le tableau 3 montrent que la totalité des critères anciennement pris en compte dans la pondération des points critiques avaient été à nouveau sélectionnés par les experts dans cette révision de la méthode. De même pour chaque CCP, la proportion de critères pondérés par les experts (qui ne l'étaient pas dans la version initiale de la méthode) est supérieure à 50 % (sauf pour le CCP « Environnement » ou elle est à 33 %, car un seul nouveau critère a été intégré par les experts). Les CCP « Diffusion de l'information », « Recueil et circulation des données » et « Échantillonnage » sont ceux pour lesquels cette proportion est la plus importante.

Le tableau 4 montre que le nombre de critères très divergents était supérieur à 10 dans la majorité des cas (sauf pour le CCP « Objectif » et « Environnement »), et était le plus important pour les CCP « Animation » et « Recueil et circulation des données » où il atteignait respectivement 25 et 26. De même, le nombre de critères faisant l'objet d'un consensus entre les experts était assez faible et inférieur à 10 pour l'ensemble des CCP. Enfin, le nombre de critères faiblement divergents (non discutés au second tour de Delphi) était inférieur à 10 pour tous les CCP, sauf pour le CCP « Outils ».

Tableau 4 : Nombre de critères participant aux points critiques dont la pondération fait l'objet d'un consensus, d'une faible ou d'une forte divergence entre les experts, et nombre de critères non retenus à l'issue du premier tour d'élucidation

Points critiques	Consensus	Faible divergence	Forte divergence	Critères non retenus (car pondéré 1 seul expert)
Objectifs	2	4	2	10
Echantillonnage	3	7	10	13
Environnement	0	0	2	1
Animation	5	9	25	11
Outils	9	10	14	14
Recueil et circulation des données	2	7	26	26
Traitement et interprétation des données	2	7	10	13
Diffusion de l'information	3	9	10	21

1.6.2 Résultats du 2nd tour

Lors du second tour, ont uniquement été discutées les pondérations de 3 CCP : « Objectifs », « Échantillonnage » et « Animation », au cours d'une réunion de 3h. Lors des discussions et débats sur les critères fortement divergents entrant ou non en compte dans ces 3 CCP, des pondérations ont été décidées collégalement (cf. annexe 7) et ont conduit à une augmentation du nombre de critères à prendre en compte dans le calcul de ces CCP (on passe de 5 à 8 critères pour le CCP « Objectifs », de 8 à 22 pour le CCP « Échantillonnage » et de 19 à 26 pour le CCP « Animation »).

Voici à titre d'exemple le tableau de discussion de second tour du CCP « **Objectif** » (Tableau 5), qui rassemble l'avis de chaque expert ayant pondéré ce CCP, la proposition de pondération moyenne proposée aux experts ainsi que la pondération consensus adoptée à l'issue des discussions.

Tableau 5 : Evolutions des pondérations proposées par les experts pour le CCP « objectifs » et décision finale validée lors du 2^{ème} tour d'élicitation

Référence du critère	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Proposition de pondération	Pondération adoptée en 2 nd tour
1.4	10	10	10	0	2	2	5	5
6.1	10	10	1	2	1	10	5	2

1.4 : Cohérence des maladies surveillées avec la situation sanitaire (maladie ou danger existants / exotiques).

6.1 : Adéquation des modalités de surveillance aux objectif du dispositif.

2. Attributs

2.1 Pertinence, Clarté, Redondance

Concernant la pertinence et la clarté des attributs, les résultats du questionnaire en ligne ont montré que la moitié des experts avaient déjà éprouvé des difficultés dans la compréhension de certains attributs, notamment pour les attributs suivants : « Simplicité » (n=2), « Fiabilité » (n=1), « Flexibilité » (n=1), « Rapidité » et « Utilité » (n=1). La synthèse des résultats du questionnaire en ligne sur ces thématiques est présentée dans le tableau 6.

Finalement lors des réunion de second tour d'élicitation, les discussions ont permis de valider un total de 3 attributs actuels à renommer : l'attribut « Rapidité » en « Réactivité », l'attribut « Stabilité » en « Durabilité » et l'attribut « Simplicité » en « Clarté ».

2.2 Pondérations des attributs

La majorité des experts ont considéré qu'il n'était pas pertinent de pondérer les attributs entre eux car ils n'utiliseraient pas ces pondérations pour hiérarchiser les recommandations pour 58 % (7/12) d'entre eux. D'autres raisons ont également été mentionnées : comme pour les points critiques, l'ajout d'une pondération entre les attributs risquerait d'amplifier l'incertitude (n=2), le poids à attribuer aux attributs peut dépendre des objectifs du dispositif de surveillance évalué (n=7), de son contexte de mise en œuvre (n=1), et du type de dispositif évalué (n=1), ce qui limite donc la possibilité de proposer un système de pondération générique pouvant être applicable à tout dispositif.

2.3 Proposition d'ajout d'attributs

Plusieurs experts ont proposé l'ajout de nouveaux attributs d'évaluation dans la méthode :

- **Qualité des données** (6/12, 50 %), qui pourrait inclure l'exhaustivité, la validité et à la cohérence des données enregistrées en adéquation avec les besoins et objectifs du dispositif ;
- **Interopérabilité** (6/12, 50 %), qui pourrait correspondre à la capacité du dispositif à fonctionner de manière transparente et efficace avec d'autres dispositifs ou plateformes de surveillance, permettant ainsi l'échange de données, d'informations et de fonctionnalités ;
- **Efficience** (3/12, 25 %), qui pourrait correspondre la capacité du dispositif à remplir ses objectifs de manière optimale en utilisant de manière judicieuse les ressources disponibles. Cela implique d'atteindre les résultats souhaités avec un minimum de ressources (coûts, temps, matériels...) et d'efforts ;
- **Bénéfice, Communication, Complétude, Efficacité** (2/12, 16 %) ;
- **Valeur Prédictive Positive (VPP), Précision et Faisabilité** (n=1/12, 8 %).

Tableau 6 : Synthèse des résultats du questionnaire en ligne sur la pertinence, la clarté et la redondance des attributs

Attributs	Pertinence			Sans réponse	Clarté	Redondance
	à conserver	à reformuler	à supprimer			
Sensibilité	83 % (10/12)	8 % (1/12)	0 %	8 % (1/12)		
Spécificité	83 % (10/12)	8 % (1/12)	0 %	8 % (1/12)		
Représentativité	91 % (11/12)	0 %	0 %	8 % (1/12)	Les dispositifs peuvent avoir une mauvaise note sans que ce soit forcément gênant dans leur fonctionnement.	
Rapidité	66 % (8/12)	25 % (3/12)	8 % (1/12)	8 % (1/12)	Mieux définir les "bornes" couvertes par cet attribut : rapidité de la procédure de la détection à la confirmation du cas ? à l'inclusion dans la base de données ? autre ? (n=1) "Réactivité" à la place de rapidité serait peut-être plus adapté. La rapidité est nécessaire dans certains cas, mais la capacité d'un réseau à pouvoir prendre le temps de l'investigation de cas complexes est aussi pertinente donc la notion de réactivité peut être complexe à évaluer.	Redondant avec « Utilité ». Proposition de garder uniquement Utilité (qui sous-entend notamment la production de résultats attendus dans les temps impartis pour permettre la prise de décision ou l'action). Pas toujours essentiel d'être rapide, l'essentiel est dans la réalisation des objectifs dans le timing prévu.
Flexibilité	75 % (9/12)	16 % (2/12)	0 %	8 % (1/12)	Difficulté de compréhension (n=1) L'attribut "Fiabilité" n'est pas clair : difficultés à voir ce que cela concerne pour un dispositif de surveillance (n=1)	
Fiabilité	83 % (10/12)	0 %	8 % (1/12)	8 % (1/12)	Trop vaste : attribut qui regroupe beaucoup de critères et peut donc être difficile à interpréter ou à utiliser pour les recommandations. → peut être à découper pour l'interpréter plus facilement ? Il faudrait séparer les notions d'harmonisation, précision, exactitude par exemple.	
Stabilité	83 % (10/12)	0 %	8 % (1/12)	8 % (1/12)	Difficultés dans la différenciation entre l'attribut "Stabilité" (présent dans la méthode Oasis) et "Durabilité" (existant dans d'autres méthodes d'évaluation) (n=1)	
Acceptabilité	91 % (11/12)	0 %	0 %	8 % (1/12)		
Simplicité	75 % (9/12)	16 % (2/12)	0 %	8 % (1/12)	Terme à revoir, un système de surveillance complexe mais bien explicite et clair avec une prise en main accompagnée peut être aussi efficient et justifié. (n=2)	

					Attribut pas très intéressant (n=1)	
					Un peu confusant car fait appel à l'existence de protocoles clairs et détaillés mais malgré tout, les circuits de collecte peuvent être bien documentés et clairs avec une info bien diffusée aux partenaires sans pour autant qu'ils soient simples (n=1)	
Utilité	66 % (8/12)	25 % (3/12)	0 %	8 % (1/12)	Nécessité de clarifier ce que recouvre cet attribut. La nouvelle section 11 sur l'utilisation des résultats de la surveillance participe à l'attribut (n=1) Pas toujours simple de savoir ce que ça recouvre : semble regrouper plusieurs choses (fiabilité, sensibilité, spécificité, représentativité...) : à clarifier davantage	"Utilité" : En partie redondant avec la nouvelle section 11 « impacts », (nouvelle section proposée par le groupe de travail « RETEX-FAX »).

Lors du second tour d'élicitation, l'ajout de l'attribut « Collaboration » a été décidé, qui reposerait en partie sur les nouveaux critères dits de « surveillance intégrée » proposés dans un groupe de travail antérieur. Comme indiqué, la « qualité des données » a été intégrée à un CCP existant. Les autres attributs n'ont pas fait l'objet de consensus, même si l'efficacité est un résultat qui semble intéressant à la plupart des experts, à valoriser éventuellement à travers de nouvelles sorties graphiques.

2.4 Catégorisation des attributs

Par ailleurs, cinq experts ont trouvé pertinent la proposition de répartition des attributs en catégories : ainsi, les catégories « Opérationnel/Fonctionnement » et « Efficacité/Impact » ont été choisies à 41 % (5/12 experts) et la catégorie « Organisationnel » à 33 % (4/12). Pour donner suite à cette question, une question conditionnelle à choix multiples permettait aux experts de répartir les attributs au sein des catégories d'attributs proposées (Tableau 7). Le détail de la répartition était le suivant pour 5/12 experts : les attributs « Sensibilité », « Spécificité », « Représentativité », « Rapidité », « Fiabilité », « Simplicité » ont été répartis en majorité dans la catégorie « Opérationnel/Fonctionnement », les attributs « Flexibilité », « Stabilité » dans la catégorie « Organisationnel » et enfin l'attribut « Utilité » dans la catégorie « Efficacité/Impact ». L'attribut « Acceptabilité » restait quant à lui assez neutre (répartition équilibré entre les catégories).

Concernant les propositions de nouveaux attributs, « Interopérabilité » serait à placer dans la catégorie « Opérationnel/Fonctionnement », « Complétude » dans « Efficacité/Impact » et « Faisabilité » dans « Organisationnel ». Ces experts ont notamment mentionné que cet ajout ressemblerait à ce qui est proposé dans l'outil OH EpiCap (outil générique destiné à évaluer, caractériser et surveiller les capacités et aptitudes de surveillance, en contribuant directement à la surveillance "One Health". Il présente trois catégories : « Organisation », « Opération », « Impact », et les représentations graphiques par catégories sont bien appréciées par les utilisateurs).

Les experts ne souhaitant pas cette catégorisation ont mis en avant les arguments suivants :

- Individuellement, chaque attribut se comprend très bien, il faut éviter de multiplier les échelons et les sous-catégories pour plus de clarté, tant que le nombre d'attributs reste limité. Cela apporterait un niveau supplémentaire de complexité, pour peu de bénéfices (n=6) ;
- Les attributs sont parfois difficiles à mettre dans des cases et peuvent appartenir à plusieurs sous catégories, ce qui risque donc de "brouiller le message" (n=1).

Tableau 7 : Répartition des attributs proposée par cinq experts au sein des catégories d'attributs proposées dans le questionnaire en ligne

	Attributs opérationnels/de fonctionnement	Attributs organisationnels	Attributs d'efficacité
Sensibilité	3	0	2
Spécificité	3	0	2
Représentativité	4	0	2
Rapidité	4	1	0
Flexibilité	2	4	0
Fiabilité	3	0	2
Stabilité	2	4	0
Acceptabilité	2	1	2
Simplicité	3	2	1
Utilité	0	0	5

Les chiffres en rouge représentent la proportion majoritaire de l'attribut.

2.5 Questionnements et propositions d'évolutions

Concernant cette partie « Attributs », peu de consensus ont été obtenus concernant les problématiques soulevées lors du premier tour. Cependant, beaucoup d'idées ont été générées et proposées :

- Intégrer un affichage facultatif de certains attributs dans les sorties graphiques, qui seraient choisis par les évaluateurs au début de l'évaluation en fonction de la pertinence des attributs par rapport au dispositif de surveillance évalué ;
- Avoir quelques attributs synthétiques (appelés attributs « fourre-tout », comme fiabilité, utilité...) et d'autres plus spécifiques, qui pourraient être utilisés différemment selon la méthode complète ou flash d'OASIS. Actuellement, certains attributs sont très généraux, (tels que « Fiabilité » ou « Flexibilité »), tandis que d'autres sont plus spécifiques, (comme « Sensibilité » ou « Spécificité »). Les attributs généraux regroupent généralement de très nombreux critères (par exemple, 48 critères composent l'attribut « Fiabilité »), ce qui peut les rendre difficiles à analyser et à interpréter (d'où l'appellation par certains de « fourre-tout »). Les attributs généraux peuvent être appliqués à tout type de dispositif, alors que les attributs spécifiques ne sont pas tous à maximiser pour chaque dispositif.

2.6 Distinction points critiques et attributs

58 % des experts (7/12) considéraient qu'il est pertinent de maintenir une distinction entre les points critiques et les attributs de manière générale dans la méthode. Cela est justifié par plusieurs éléments : les experts ont affirmé que certains attributs et CCP se recoupent, et les regrouper dans un même tableau (et une même figure) renforcerait cette impression de redondance et induirait de la confusion (n=2). Il a également été évoqué lors des réunions de second tour que les CCP représentent des aspects critiques d'un dispositif de surveillance¹, compromettant leur fonctionnement et la qualité des données produites en cas de problème. De plus, ils concernent tous les dispositifs de surveillance, quels que soient leurs objectifs ; alors que les attributs sont des indicateurs plus larges et ne sont pas tous essentiels au bon fonctionnement de tous les dispositifs de surveillance (par exemple, tous les dispositifs n'ont pas besoin d'être rapides ou spécifiques). Enfin, 58 % (7/12) des experts ont considéré qu'il était préférable de conserver une différence entre les CCP et les attributs dans les sorties graphiques car le regroupement risque de rendre les choses illisibles (n=1) et n'aurait pas un grand apport (n=1), les CCP et attributs étant deux choses complémentaires et différentes.

Un expert a néanmoins proposé un graphique en radar avec tous les CPP et attributs regroupés en 3 secteurs : « Opérationnel, Organisationnel et impact », avec affichage d'une note générale au centre (n=1).

3. Sorties graphiques

3.1 Pertinence, clarté et redondance et niveau de détails des sorties graphiques

3.1.1 Pertinence

Les sorties graphiques actuelles sont considérées comme pertinentes par la majorité des répondants (Figure 2) :

- 11/12 experts ont considéré que le niveau de pertinence de la sortie 1 au moins égal à 7 (sur une échelle de 0 à 10, 10 étant la note maximale) ;
- 8/12 experts ont considéré que le niveau de pertinence de la sortie considérée comme étant au moins égal à 7 ;
- 7/12 experts ayant considéré que le niveau de pertinence de la sortie 3 au moins égal à 7.

¹ Le principe des points critiques suit celui retrouvé dans les normes afnor iso 14001:2015 ou encore ISO 45001:2018

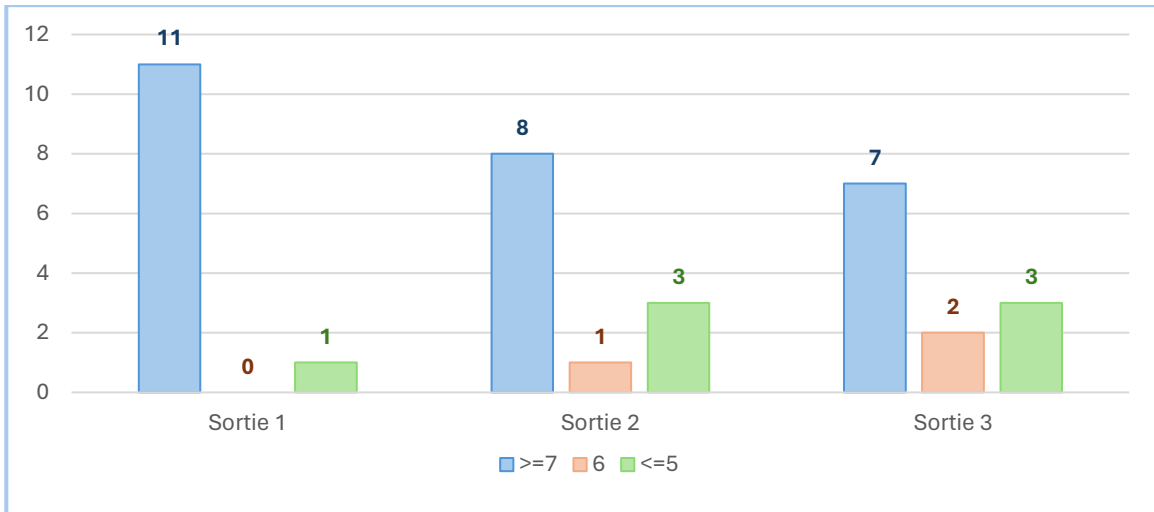


Figure 2 : Distribution de la fréquence des notes de pertinence des sorties graphiques estimées par les 12 experts répondants, sur une échelle de 0 à 10 (10 étant la note maximale)

3.1.2 Clarté

Des clarifications seraient à apporter, particulièrement pour la sortie graphique 2 (histogramme représentant le degré d'atteinte des points critiques), considérée comme la moins compréhensible.

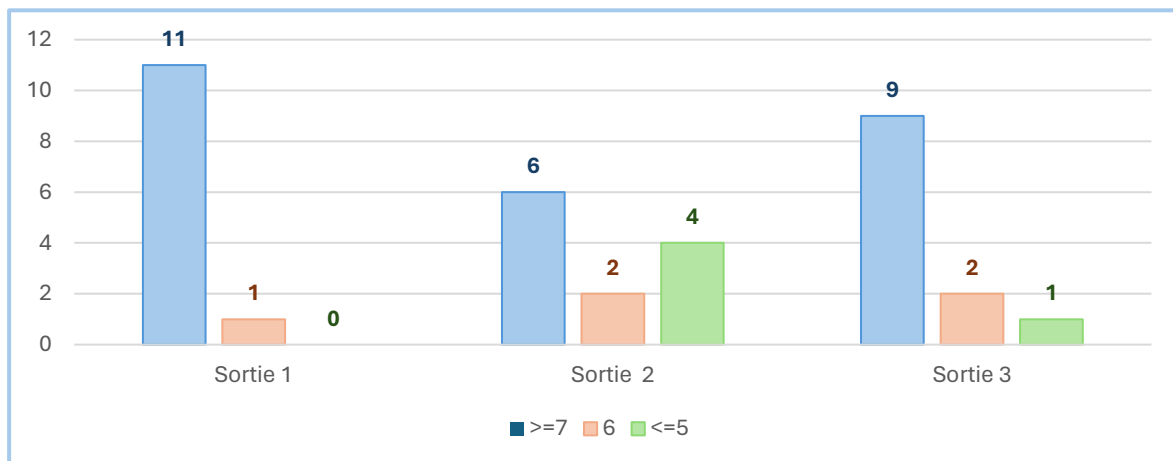


Figure 3 : Distribution de la fréquence des notes de clarté des sorties graphiques estimées par 12 experts répondants, sur une échelle de 0 à 10 (10 étant la note maximale)

3.1.3 Redondance

Concernant la redondance entre les sorties graphiques, la moitié des experts considèrent qu'il y a une redondance entre les sorties 1 et 2, et plus spécifiquement entre la section 1 et le point critique Objectif (n=2).

3.1.4 Niveau de détails

Concernant le niveau de détail des sorties graphiques :

- La moitié (6/12) des experts ont indiqué que le niveau de détail de la figure 1 n'était pas suffisant (40 % estimaient qu'il était suffisant (5/12)). Il a ainsi été proposé une figure permettant de voir le détail des notes de chaque critère (n=3), ou une figure à plusieurs niveaux (le premier niveau de lecture pourrait indiquer le niveau de satisfaction fonctionnel global pour la section, et en cliquant sur la section, un deuxième niveau pourrait permettre de visualiser les critères pris en compte et leur valeur).

- 66 % (8/12) des experts ont indiqué que le niveau de détails de figure 2 n'était pas suffisant. La largeur des barres, différente selon les points critiques, n'est en effet pas utilisée par la plupart des experts interrogés pour hiérarchiser les recommandations et complexifie la lecture de la figure. La visualisation de la satisfaction des points critiques pourrait être améliorée pour aider à la communication des résultats de l'évaluation plus clairement.
- 50 % des experts ont indiqué que le niveau de détail de la figure 3 n'était pas suffisant. Il a ainsi été proposé une meilleure visualisation incluant le détail des critères qui impactent positivement ou négativement les scores des différents attributs en complément (n=6) (par affichage permanent ou dynamique du numéro ou du mot associé aux critères concernés (n=1)).

3.2 Propositions d'évolutions

Au total, 58 % (7/12) des experts considèrent qu'il n'est pas pertinent d'ajouter une nouvelle sortie graphique à la méthode OASIS. Parmi les 33 % qui étaient pour cet ajout, les aspects que les experts souhaitaient voir représenter étaient les suivants : des aspects économiques (n=1), l'utilité des résultats (n=1), la surveillance intégrée (n=1), un score global (cf. outil OH-EPICAP) (n=1) ou encore une représentation des possibles nouvelles catégories d'attributs (n=1).

Les experts ont également proposé des évolutions des sorties graphiques existantes : faire ressortir davantage les critères qui font défaut/sur lesquels il faut cibler les recommandations ; pouvoir représenter visuellement le manque à gagner dans la note globale et dans les points critiques/attributs (n=1), présenter plutôt les points critiques et attributs au sein d'une seule sortie graphique, sous forme de deux radars côte à côte (avec par exemple à gauche les "attributs organisationnels" qui correspondraient aux CCP et à droite les « attributs fonctionnels » qui correspondraient aux attributs actuels), et en dessous un tableau des critères (incluant la note de chaque critère et son poids) (n=1), et enfin, un expert a indiqué qu'il serait intéressant de pouvoir visualiser l'évolution entre deux évaluations (voire plus) par comparaison de différents sets de scores des critères (n=1).

Le tableau 8 présente la pertinence de modifier les sorties graphiques ainsi que les principales propositions d'évolutions émises par les experts dans le questionnaire en ligne, pour chaque sortie graphique.

Tableau 8 : Synthèse des résultats sur la pertinence de la modification des sorties graphiques et propositions d'évolutions associées

Sorties	Pertinence de modification éventuelles de la sortie	Proposition d'évolutions
N°1	La moitié des experts (6/12) considèrent qu'il n'est pas pertinent de modifier la représentation des sections fonctionnelles.	<ul style="list-style-type: none"> • Fusionner les sorties 1 et 2 (qui visent toutes les deux des aspects fonctionnels du dispositifs : Objectifs, Outils, Diffusion de l'information, ...) • Proposition de transformation en diagrammes en barres représentant le pourcentage de satisfaction, ce qui pourrait faciliter l'extraction de la figure et sa lecture Proposition de transformation en histogramme circulaire permettant de visualiser tous les critères d'un seul coup d'œil, avec une mise en lumière des critères à faible score/qui font défaut. Chaque critère devrait être idéalement identifiable en un mot pour faciliter la lecture des sorties.
N°2	75 % (9/12) des experts considèrent qu'il est pertinent de modifier la représentation graphique des CCP.	<ul style="list-style-type: none"> - Harmonisation de la largeur des barres (n=4) - Graphique interactif (ou Bulle interactive, qui apparait au passage de la souris sur le point critique survolé, indiquant les critères qui font défaut pour ce point critique, sur lesquels porter en priorité les efforts de priorisation) (n=2) - Graphique en Radar (n=2)

	- Écrire en toutes lettres le nom des CCP sur la sortie 2, clarifier le détails des calculs sous-jacents pour en faciliter l'interprétation (n=1).
N°3	<p>66 % (8/12) des experts considèrent qu'il est pertinent de modifier la représentation graphique des Attributs.</p> <p>- Ajout d'un tableau sous le radar avec le détail des critères, pour pouvoir identifier plus facilement les critères "en défaut" (mal notés) pour chaque attribut, et ainsi faciliter la formulation de recommandations (n=2)</p> <p>- Diagramme en bâton (n=4) mais circulaire (n=1)</p> <p>- Radar (n=1) avec tous les CPP et attributs en 3 secteurs : "Opérationnel, Organisationnel et impact", avec affichage de la note générale au centre. (n=1).</p>

Les sorties graphiques n'ont pas été abordées dans les réunions du second tour d'élicitation.

IV. Discussion

1. Discussion de la méthode

1.1 Choix de la méthode

Depuis sa création en 2011, la méthode a été utilisée pour plusieurs dizaines d'évaluations. Après dix ans d'utilisation sans modification, il était crucial de la réévaluer pour s'assurer qu'elle reste pertinente et efficace. Le recours aux experts s'est donc avéré essentiel pour recueillir leurs avis et leur expertise dans divers domaines sanitaires. Cela permettrait à la méthode OASIS de rester applicable de manière générique à différents secteurs et types de dispositifs. En collectant ces retours, nous avons ainsi pour objectif d'optimiser la méthode et de la faire évoluer en fonction des besoins actuels et des expériences accumulées.

L'objectif de l'élicitation n'était pas seulement de recueillir des avis, mais de parvenir à un consensus (Booto Ekionea, Bernard and Plaisent, 2011) sur les évolutions à envisager, en sollicitant des experts de différents domaines et niveaux d'expertise. Il était en effet crucial que les propositions de modifications soient acceptées par le plus grand nombre, afin que la version 2.0 de la méthode reste applicable à tout type de dispositif. Cela a conduit à l'adoption d'une élicitation d'experts par la méthode Delphi à deux tours.

Ce choix nous a permis d'adopter deux approches différentes pour les deux tours : un travail individuel lors du premier tour, permettant de préparer le travail collectif du second. Cela a également permis à chaque expert de s'exprimer individuellement sans être influencé par les avis ou comportements des autres experts, ce qui était important pour garantir la neutralité des résultats et s'assurer de capter toutes les opinions, y compris celles minoritaires.

La sollicitation des experts ne s'est toutefois pas révélée suffisante, soulevant des questionnements sur la pertinence de la méthodologie adoptée. Récolter une quantité importante d'informations au premier tour, sur des thématiques variées et des questions spécifiques d'évolution de la méthode, tout en permettant aux experts de proposer et commenter leurs réponses a finalement conduit à l'impossibilité de tout discuter en seulement deux réunions de « deuxième tour » de la méthode Delphi, et donc à la finalisation des débats dans le cadre de ce stage de Master. Toutefois, dans ce travail de refonte, il était essentiel de soumettre la méthode OASIS à un questionnaire complet de ses éléments, tant sur le fond que sur la forme. Par conséquent, il est logique que la quantité de problématiques générées ait été conséquente, et le travail devra se poursuivre après ces six mois de Master 2, pour arriver à une évolution approfondie et aboutie de la méthode Oasis.

Les réunions du second tour ont permis de soulever plusieurs points importants. D'une part, la réflexion de groupe a mené à des discussions et des débats particulièrement riches, générant ainsi des idées communes et favorisant un véritable partage d'opinions et de points de vue. Ce travail de cocréation entre experts a permis de faire progresser efficacement les réflexions. D'autre part, il a été

observé que ce travail collaboratif facilitait la compréhension des débats pour certains experts et permettait de traiter plus efficacement certaines problématiques. De plus, ces échanges riches et approfondis ont conduit certains experts à réviser leur opinion initiale, comme attendu dans la méthode Delphi. Cependant, il est intéressant de noter qu'à plusieurs reprises, une opinion initialement soutenue par un seul expert a finalement été adoptée par l'ensemble du groupe après des discussions et des échanges approfondis.

La refonte de la méthode OASIS a permis de faire progresser de nombreux aspects jusqu'à présent. Néanmoins, l'ampleur des problématiques et des éléments à mettre à jour s'est rapidement révélée. Comme un effet boule de neige, la modification d'un seul élément entraîne la modification de plusieurs autres éléments et ainsi de suite. Par exemple, renommer un CCP ou redéfinir ses limites peut nécessiter des ajustements dans les critères qui lui sont attribués et dans leurs pondérations, car ces derniers peuvent ne plus avoir le même poids. Même si les experts ont été initialement sollicités pour attribuer une pondération aux CCP et attributs existants, la révision de certains d'entre eux (modification de définition, redimensionnement), peut conduire à rendre les résultats du premier tour rapidement obsolètes et donc à rediscuter les critères à prendre en compte et leur pondération de manière individuelle ou collégiale.

Cela amène à remettre en question la méthode adoptée pour cette révision. N'aurait-il pas été préférable de commencer par tout ce qui relève de la forme de la méthode OASIS, tel que les éléments du questionnaire Kobotoolbox, la compréhension des CCP et attributs, leurs définitions, leurs noms, leurs ajouts, suppression ou encore la pertinence de certains aspects comme la catégorisation des CCP (visible uniquement dans la feuille de calcul), ou la pondération des CCP, attributs ou sous-catégories éventuelles ? Puis une fois ces éléments définis, initier la réflexion sur l'identification des critères participant à chaque CCP / attribut et leur pondération associée, en prenant en compte les évolutions apportées par la partie questionnaire. Cette approche n'a pas été retenue en premier choix, car il semblait nécessaire, dans le temps imparti, d'interroger les experts sur les aspects essentiels de la méthode, y compris les pondérations, et non pas seulement sur une partie de la méthode ou sur ce qui relève de la forme.

1.2 Choix des experts

Le choix des experts conditionne la qualité des résultats (Adler and Ziglio, 1996), c'est pourquoi le choix des expert a été crucial. La méthode choisie a permis de solliciter un panel d'experts issus de domaines variés, permettant ainsi d'obtenir des opinions diversifiées et de façonner une nouvelle version d'OASIS inclusive, pouvant continuer à être appliquée à différents secteurs sanitaires.

Initialement, 15 experts ont été sollicités, mais seuls 12 ont finalement participé à la révision de la méthode. Parmi les 3 experts qui n'ont pas pu participer à cette refonte, l'un provenait du secteur de la santé animal, et deux autres du même secteur mais travaillaient également sur l'antibiorésistance. Ces derniers n'ont pas pu participer par faute de temps ou par manque d'expérience avec la méthode. Malgré l'absence de leur contribution, les débats et les informations collectées lors de l'élicitation semblent montrer que l'absence de réponse de ces experts n'a pas eu un impact significatif, car finalement la qualité des données collectées lors de ces deux tours repose surtout sur la qualité de l'expertise et des compétences des experts. De même, elle repose également sur de leur intention de coopérer sur une longue période et durant les débats (P Baillette, Fallery and Girard, 2013), ce qui a été le cas pour cette révision. De plus, une hétérogénéité dans la représentativité des experts a été observée, avec une sous-représentation du domaine de la santé végétale. Cette disparité pourrait potentiellement introduire des biais. Elle peut être attribuée au fait que la méthode est plus largement utilisée et développée dans le domaine de la santé animale, par rapport à d'autres secteurs.

Ensuite, il est important de mentionner une limite liée aux possibilités de sollicitation des experts. Il n'est en effet pas possible de les sur-solliciter, par ailleurs déjà très occupés dans leurs travaux actuels. Il a donc été nécessaire d'adapter l'éllicitation à la disponibilité et au temps que les experts pouvaient y consacrer. Le travail sur les matrices a dû être allégé, passant d'une matrice complète à remplir (d'un total de 18 CCP et attributs), à seulement 6 indicateurs (CCP ou attributs) à remplir.

Il est également pertinent de considérer l'importance de choisir des experts ayant un niveau d'expérience suffisant avec la méthode OASIS. En effet, certains experts se sont engagés pleinement dans le processus, tandis que d'autres ont rencontré des difficultés à se positionner en raison de l'inadéquation de leur domaine d'expertise dans certains cas (par exemple durant la pondération, certains experts n'ont pas pu pondérer un bon nombre de critères qui ne s'appliquaient pas à leurs dispositifs de surveillance).

1.3 Questionnaire en ligne

Le questionnaire de sollicitation des experts pour le premier tour de l'éllicitation d'experts par la méthode Delphi a été élaboré sur la plateforme Kobotoolbox. Même si cela a grandement facilité la collecte de données, certains aspects techniques ont quelque peu compliqué sa réalisation. Le logiciel a en effet offert un large éventail de possibilités pour les questions conditionnelles, permettant l'affichage de différents tableaux en fonction des choix effectués en amont (cf. annexe 2). Ces fonctionnalités ont permis de capturer fidèlement la trame de questionnaire initialement prévue sur format Word, de manière plus efficace que d'autres plateformes envisagées, telles que Lime Survey ou Sphinx par exemple.

Néanmoins, il est important de noter que les tableaux initialement prévus dans le questionnaire (cf. annexe 2), permettant de récolter des informations sur la pertinence, la pondération et les commentaires concernant les CCP et Att ont été difficiles à réaliser sur cette plateforme. Le logiciel n'a en effet pas permis de trier et de regrouper correctement les données issues de ces tableaux, nécessitant leur extraction individuelle, ce qui a rendu l'analyse des données assez chronophage.

De même, la mise en page a été particulièrement complexe à mettre en place. En effet, la présentation d'un questionnaire aussi complexe doit être la plus simple et compréhensible possible pour les experts, et ces difficultés ont entravé la clarté des questions dans le questionnaire ayant possiblement influencé sa complétion par les experts.

Il est également important de souligner l'anonymat des réponses paramétrées pour le questionnaire en ligne, qui n'a pas permis de faire le lien entre les réponses ou commentaires provenant du même expert. En effet, certaines questions pouvaient avoir un lien entre elles, notamment les questions 5, 9 et 10 du questionnaire (cf. annexe 2) : deux questions portaient sur l'intérêt d'ajouter des CCP ou d'attributs à évaluer, et une autre sur la pertinence d'une liste proposée de nouveaux attributs identifiés dans la littérature. Des réponses similaires ont été identifiées à plusieurs endroits du questionnaire, posant un problème pour le comptage des propositions, car une même idée répétée par un expert pouvait être comptée plusieurs fois à tort.

Bien qu'il aurait été possible de faire une analyse individuelle de chaque questionnaire rempli, cela aurait été très chronophage dans le temps imparti. De plus, au vu de la quantité de documents produits à la suite du premier tour, cette analyse n'aurait pas respecté les limites fixées pour le retour d'analyse aux experts avant le second tour. Néanmoins, les résultats des réunions du second tour ont montré que ce ne sont pas toujours les idées proposées par la majorité des experts qui ont finalement été retenues à la suite des discussions. Par exemple, la décision de séparer le CCP « Outil » en 2 (« Outils Laboratoire » et « Outils Terrain ») prise lors des réunions du second tour n'avait été proposée que par un seul expert.

Enfin au vu de la quantité de données générées par le questionnaire et les matrices (zones de commentaires permettant aux experts de s'exprimer sur divers aspects de la méthode en plus des questions), il a fallu faire un tri et établir des priorités dans les sujets à aborder avec les experts, même si chaque question aurait mérité une discussion plus approfondie.

1.4 Matrice de pondération des critères associés aux points critiques ou attributs

L'échelle allant de 1, 2, 5 et 10 a été sélectionnée pour plusieurs raisons. Tout d'abord, le choix d'un nombre pair vise à éviter l'attribution de notes moyenne par défaut, lors d'un avis tranché, incitant ainsi les experts à exprimer leur position de manière plus claire. Cela permet d'éviter les valeurs moyennes difficiles à interpréter. De plus, il est important de fournir des résultats permettant de distinguer clairement les différents niveaux d'avis ou de préférence.

L'un des inconvénients de ce choix de pondération peut résider dans l'inégalité des écarts entre les valeurs. Par exemple, l'écart entre 10 et 5 est beaucoup plus significatif que celui entre 2 et 5, ou entre 1 et 2, ce qui pose questions sur le poids relatif de chaque valeur. Toutefois, cette méthode a été délibérément conçue avec des extrêmes marqués, (par exemple, un poids maximal de 10 est 10 fois plus élevé qu'un poids minimal de 1), évitant ainsi les poids intermédiaires (tels que 2, 3 ou 5). Ainsi, l'attribution d'un poids de 10 était réservée à des situations extrêmes où le critère était considéré comme nettement plus important que les autres, afin de le mettre en évidence dans l'analyse des données. Cette réflexion remet en question notre échelle.

Dans la version initiale de la méthode, seules trois valeurs de poids étaient proposées (1, 2 ou 3), le poids maximal étant ainsi d'une part assez faible et par ailleurs peu éloigné de la valeur moyenne, rendant la discrimination entre critères moins importante, et pouvant donc restreindre la proposition de recommandations pertinentes. De même, certaines études mentionnent que les échelles avec le moins de catégories ont tendance à produire des scores et réponses moins fiables (Preston and Colman, 2000). L'échelle choisie dans la révision de la méthode permet de mieux différencier les critères dans leurs poids et importance réelle dans la notation d'un CCP ou d'un attribut.

Le choix d'utiliser une matrice comme celle choisie pour l'élicitation d'experts soulève plusieurs points de réflexion. Tout d'abord, il est important de rappeler que ces matrices présentaient l'ensemble des critères présents dans la méthode (critères « historiques » et nouveaux critères sur les aspects collaboratifs et d'impacts, proposés dans le cadre de groupes de travail en 2022-2023), et, pour les critères historiques, identifiaient à l'aide d'un fond de case coloré les critères participant à chaque point critique ou attribut dans la version initiale de la méthode. Cela a pu influencer les experts lors de leur révision de la méthode, les incitant à attribuer en priorité des poids à ces critères, et peut-être à délaissier davantage les critères non initialement sélectionnés (s'ils n'étaient pas inclus initialement, pourquoi devraient-ils l'être maintenant ?).

Pour évaluer cette influence, l'impact des cases à fond coloré sur la capacité des experts à s'écarter des lignes directrices de l'ancienne méthode a été analysé. Il a été constaté que chaque expert a fait un réel travail de réflexion sur les matrices et ne s'est pas uniquement fié aux critères existants. Les experts non pas hésité à proposer de nouveaux critères pour les notations, bien qu'ils aient été moins enclins à supprimer les anciens critères.

Initialement, les matrices devaient être complétées entièrement par tous les experts. Cependant, en raison de leur complexité et de leur remplissage chronophage (94 critères, 8 points critiques, 10 attributs), il a été décidé de ne solliciter les experts que sur seulement 6 points critiques ou 6 attributs chacun. Finalement, chaque indicateur (CCP ou attribut) a été noté par 4 à 7 experts, ce qui n'a pas rendu hétérogène le traitement des données, et conduit à une précision différente dans la pondération moyenne obtenue, soulevant également la question de la suffisance des données collectées. Ce point

a toutefois certainement en partie été neutralisé par le 2ème tour de la méthode Delphi, qui a permis de valider l'ensemble des résultats obtenus.

Les résultats des matrices (pondération par critère pour chaque point critique et attribut) n'ont pu être discutés et débattus qu'à la deuxième réunion du second tour de la méthode Delphi. Cela met en évidence un problème rencontré dans la planification initiale des réunions. Le temps consacré à la discussion des critères présentant une forte divergence a été beaucoup plus long qu'estimé initialement, en raison de la richesse des échanges et des arguments présentés. De plus, certains critères proposés par seulement 2 experts sur 5 ou 6, souvent avec des poids faibles, ont également été discutés. Contrairement aux attentes initiales, ces discussions ont parfois conduit les autres experts à changer d'avis et à retenir ces critères, alors qu'ils semblaient destinés à être éliminés. Ainsi les réunions se sont avérées très chronophages, en particulier pour certains points issus du questionnaire Kobotoolbox, tel que les discussions sur la suppression, reformulation et redéfinition de certains points critiques et attributs. Néanmoins, la richesse des opinions et des idées générées lors des discussions sur la pondération de ces 3 CCP n'est pas négligeable.

Les résultats de l'analyse, présentés dans la partie résultats, suggèrent qu'une « règle » de pondération des critères divergents aurait pu être établie et appliquée pour des cas similaires, permettant de gagner du temps lors des réunions à la vue du nombre de critères très divergents à discuter. Cependant nous n'avons pas pu appliquer de règle « type » et tous les critères ont dû être discutés avec les experts, même avec des profils de pondération issues du 1er tour similaires : les choix de consensus n'étaient finalement pas les mêmes, la décision finale dépendant intrinsèquement des discussions entre experts pour chaque critère, et non des pondérations initialement proposées individuellement par chaque expert.

Plusieurs cas de figure ont été identifiés : d'une part les cas où une minorité d'experts s'éloignait de la tendance de pondération générale. Dans ces cas, le groupe parvenait souvent à convaincre cette minorité de se retrancher vers la pondération proche de la tendance. Il y a toutefois eu des cas inverses où une minorité d'experts qui s'éloignait du choix de pondération soutenu par la majorité, réussissait finalement à convaincre les autres avec des arguments solides d'opter et de se ranger du côté de la pondération minoritaire ou du moins de choisir une pondération qui s'en rapprochait plus que la pondération moyenne.

Considérant ces deux situations, le temps passé pour choisir la pondération finalement retenue, qui fasse consensus auprès des tous les experts, semblait ainsi justifié et légitime en tout point.

Un débat et une discussion ont permis la liberté et le partage d'idées, en prenant soin d'inclure l'avis des différents domaines représentés en temps réel, ce qui a d'ailleurs été mentionné à la fin de ces deux réunions.

2. Discussion des résultats

2.1 Questionnaire en ligne

Un réel travail de tri des sujets et problématiques à aborder en second tour a dû être effectué, tous les aspects ne pouvant être abordés en quelques réunions. Les résultats du questionnaire en ligne et des réunions montrent également que la partie CCP a obtenu beaucoup plus de consensus que la partie attribut. Cela pourrait s'expliquer par la difficulté à comprendre concrètement ce que représentent les Attributs, alors que les points critiques sont clairement identifiés comme des éléments essentiels au bon fonctionnement de tout dispositif. Il est plus facile de déterminer ce qui constitue un point critique que de définir des attributs, certains étant très larges (regroupe beaucoup de critères). De plus, de nombreuses idées et propositions ont été recueillies pour la partie attributs, rendant difficile l'obtention d'un consensus en premier tour et nécessitant une poursuite des réflexions au second tour.

Les experts étaient globalement d'accord sur le fait que les points critiques constituent une partie importante de la méthode, sur laquelle tous les dispositifs de surveillance doivent atteindre la meilleure note possible, reflétant ainsi leur bon fonctionnement. La suppression du CCP « Environnement » peut finalement s'interpréter par le fait qu'il n'était constitué que de deux critères qui finalement peuvent s'intégrer au sein d'autres CCP. De même, l'ajout du CCP « Moyens » ainsi que la redéfinition des points critiques « Recueil et circulation des données » et « Traitement et interprétation des données » en « Collecte et Qualité des données » et « Analyse et interprétation des données » ont permis d'affiner ce qui est essentiel à l'évaluation des dispositifs de surveillance au vu des experts, garantissant une meilleure compréhension des CCP, et une meilleure couverture d'un ensemble complet de points indispensables au bon fonctionnement d'un dispositif de surveillance.

La suppression des catégories de CCP et de la pondération associée (déterminant la largeur des barres de la sortie 2) s'intègre également dans cette logique de meilleure compréhension et de fluidité des CCP. L'objectif était de simplifier la méthode sans pour autant sacrifier le niveau de détail, la rendant au contraire la plus complète et détaillée possible. Il est indispensable de garantir une méthode à la fois simple et compréhensible pour les évaluateurs. Quant à la redondance, elle sera principalement résolue par une redéfinition et une répartition plus claire des critères participant à la notation des CCP et attributs.

Les sorties graphiques s'intègrent dans cette même logique de simplification. Le questionnaire a permis d'obtenir une certaine richesse de résultats et d'idées proposées permettant d'imaginer les ébauches des sorties graphiques de la méthode 2.0. Les résultats ont montré des difficultés de compréhension plus importantes pour la sortie 2 (barres d'histogramme représentant le degré de satisfaction des points critiques) et globalement, un niveau de détails des sorties graphiques insuffisant. En effet, les experts souhaiteraient pouvoir faire afficher sur les graphiques, si possible de manière dynamique au passage de la souris, les critères faisant le plus défaut pour chaque point critique ou attribut considéré, permettant ainsi de mieux cibler les recommandations à formuler. Ainsi, 58 % des experts ne considéraient pas comme pertinent l'ajout d'une nouvelle sortie graphique, soulignant plutôt la nécessité de modifier les sorties graphiques existantes. Cela implique une volonté des experts d'atteindre un niveau de détail supérieur et d'intégrer alors une visualisation des critères (défaillants dans les dispositifs).

Enfin concernant les attributs, dans la logique d'enlever les catégories de CCP, les catégories d'attributs n'ont pas été choisies comme une problématique assez importante à débattre en second tour et ne seront donc pas gardées dans la version 2.0. Les discussions se sont plutôt focalisées sur la distinction entre des attributs globaux souvent considérés comme « fourre-tout » (terme utilisé par les experts), et les attributs plus spécifiques, comme la sensibilité ou la spécificité. Les attributs globaux contiennent un nombre important de critères et caractérisent la qualité générale du dispositif, tandis que les attributs spécifiques traitent des aspects plus détaillés. Il a été souligné par les experts qu'un dispositif n'avait pas besoin que tous les attributs aient une note d'évaluation élevée pour fonctionner correctement. Par exemple, des attributs comme la rapidité ou la simplicité ne sont pas indispensables pour qu'un dispositif soit performant, et ne sont pas forcément à maximiser selon le type de dispositif. Un dispositif efficace peut être complexe ou ne pas être particulièrement rapide. Cela souligne l'importance de hiérarchiser les attributs en fonction de leur pertinence pour chaque dispositif. De plus, une évolution notable de la méthode est proposée : conserver des attributs généraux à évaluer systématiquement, bien qu'ils soient complexes à analyser en raison de leur fondement sur de nombreux critères, car ils fournissent une vue d'ensemble sur le dispositif ; et choisir, en amont de l'évaluation, les attributs spécifiques à évaluer selon les objectifs du dispositif.

2.2 Matrice de pondération des critères permettant l'évaluation des indicateurs

Concernant la partie résultats des matrices, comme évoqué plus haut, la pondération finale n'a été discutée et obtenue que pour 3 CCP. Les discussions et débats concernant les critères très divergents lors des réunions du second tour ont semblé essentiels pour tous les experts, et ont permis de mettre en lumière leur importance dans la méthode de Delphi.

Pour pouvoir interpréter correctement les résultats, il a été crucial de prendre en compte les potentiels biais d'interprétation, notamment en ce qui concerne les tendances et la sensibilité des experts aux notes extrêmes (1 ou 10 sur l'échelle de pondération proposée). En effet, l'analyse des profils des experts a montré que ceux-ci pouvaient être très différents : certains experts ont en effet eu tendance à rester sur les notes centrales (experts 2, 5, 7, 9) et d'autres ont davantage eu tendance à s'en éloigner et à utiliser facilement l'extrême représenté par la note de 10 (experts 3, 6, 8, 10). Globalement, chaque expert avait un profil unique, mais tous ont peu utilisé le poids de 1. Des études montrent que lorsque les participants ont tendance à choisir des réponses extrêmes ou, au contraire, à les éviter systématiquement, et cela peut fausser les résultats (Liu *et al.*, 2017). Cependant, ce n'est pas forcément le cas dans notre étude. Les experts présentent des tendances de réponse assez variées, utilisant à la fois des réponses extrêmes et centrales. Seul l'expert 9 se distingue par une tendance marquée à éviter systématiquement les réponses extrêmes.

De même, les experts ont joué le jeu et ont tous proposé un certain nombre de nouveaux critères, permettant d'enrichir les discussions du second tour. Cela témoigne de la liberté de réflexion individuelle laissée aux experts qui se sont parfaitement prêtés au jeu.

Tableau 9 : Répartition de la fréquence de pondération des cinq experts ayant pondérés les indicateurs

	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	Exp 9	Exp 10	Exp 11
Poids 10	22 %	7 %	32 %	10 %	4 %	42 %	9 %	34 %	4 %	39 %	0 %
Poids 5	34 %	28 %	29 %	32 %	56 %	27 %	27 %	32 %	37 %	39 %	31 %
Poids 2	28 %	48 %	26 %	23 %	41 %	25 %	46 %	21 %	53 %	16 %	2 %
Poids 1	16 %	17 %	12 %	35 %	0 %	6 %	11 %	13 %	6 %	2 %	68 %
Pourcentage de cases non coloré parmi l'ensemble des cases pondérées*	29 %	50 %	68 %	55 %	26 %	52 %	41 %	44 %	15 %	43 %	68 %
Pondération la plus fréquente	5	2	10	1	5	10	2	10	2	10 et 5	1

* Pourcentage de critères nouvellement pondérés par les experts (non inclus dans la version 1.0)

À la suite du travail de pondération des experts du premier tour, une quantité importante de critères ont été pondérés (Tableau 10). Pour beaucoup de critères pour lesquels la pondération aurait pu être à débattre (critères divergents), une pondération proche de la moyenne a été choisie et proposée, sans être discutée lors des réunions de second tour (même si chacun avait accès à la matrice d'analyse des résultats, permettant à tous les experts d'évoquer un critère non prévu à la discussion). De même, pour les critères n'ayant reçu qu'une seule pondération, il est impossible de savoir si une discussion aurait pu ouvrir un débat pertinent et finalement prouver l'importance de ces critères. Un total assez conséquent de critères de ce type n'a donc pas été évoqués (environ une dizaine par point critique). On pourrait également se demander si le nombre d'experts participant à l'élicitation avait été plus élevé, si cela aurait conduit à éliminer moins de critères d'office, en supposant que, peut-être, davantage d'experts les auraient sélectionnés.

L'objectif initial étant de simplifier la méthode, pour les trois CCP examinés en détail, la révision de la méthode a plutôt conduit à augmenter le nombre de critères intégrer dans leur pondération : on

passé de 5 à 8 pour le CCP « Objectifs », de 8 à 22 pour le CCP « Échantillonnage », et de 19 à 26 (39 retenus au départ moins 13 non retenus au second tour) pour le CCP « Animation ». Cette augmentation n'a pas simplifié la formule de calcul, ce qui signifie que l'objectif initial n'a pas été pleinement atteint. Néanmoins tous les critères conservés semblaient importants à prendre en compte par les experts.

Tableau 10 : Répartition des critères pour les trois points critiques discutés durant le second tour de Delphi

Points critiques	Critères faisant consensus	Critères moyennement divergents	Critères très divergents	Critères non retenus (sélectionné par un seul expert, quel que soit sa pondération)	Nombres total de critères pondérés
Objectifs	2	4	2	10	18
Echantillonnage	3	7	10	13	33
Animation	5	9	25	11	49

3. Perspectives et Recommandations

En perspective, compte tenu de la quantité importante de données collectées et de la richesse des échanges entre les experts, il apparaît nécessaire de consacrer plus de temps que prévu initialement à la refonte de la méthode OASIS. Les experts ont décidé collectivement de poursuivre les échanges sous forme de réunions, malgré le temps que cela requiert, car c'est le seul moyen de véritablement approfondir les résultats et les discussions.

Une poursuite des réunions est donc prévue et permettra de confirmer et d'obtenir un consensus pour la nouvelle pondération des CCP et attributs, qui n'évoluent pas énormément (CCP assez proche de la version 1.0). De plus, une nouvelle sollicitation des experts sur les matrices est prévue, pour pondérer les nouveaux points critiques (« Moyens », « Outils Laboratoire », « Outils Terrain », « Collecte et Qualité des données ») et attributs (« Collaboration »), ainsi que les attributs à redimensionner, (« Fiabilité », « Utilité » et « Simplicité »), en individuel, avant de discuter à nouveau des aspects discordants en réunion.

Un travail de mise à jour globale de la méthode est prévu. En effet, certains changements ayant obtenu consensus, comme la suppression des catégories de CCP ou encore la suppression du CCP « Environnement » seront intégrés. De même, contrairement aux CCP, qui sont tous indispensables pour le bon fonctionnement d'un dispositif, ce n'est pas toujours le cas des attributs, (par exemple, certains attributs doivent être spécifiques, tandis que d'autres, moins). En revanche, certains attributs généraux, comme la fiabilité et la flexibilité, peuvent s'appliquer à tout dispositif. Il pourrait donc être envisagé que les évaluateurs sélectionnent les attributs à évaluer au début de l'évaluation, permettant au logiciel de calculer automatiquement et d'intégrer uniquement les attributs nécessaires pour les sorties graphiques.

Il est également prévu un travail de vérification de l'ensemble des critères et d'identification de ceux qui ne participent potentiellement plus à aucun CCP ni attribut, une fois fini le travail de révision de la méthode. Un aspect important, ressorti lors des réunions, est que l'objectif n'est pas de répartir l'ensemble des critères actuels au sein des CCP et attribut, mais de sélectionner les critères les plus pertinents pour donner un sens au CCP et attribut. Ainsi, si certains critères ne sont plus utilisés dans au moins un CCP ou attribut après la nouvelle pondération, leur rôle et intérêt dans la méthode pourront alors être remis en question. Il est important de notifier également que certains experts ont proposé de nouveaux critères ou reformulations de critères, qui seront traités hors stage, cela sera repris dans le cadre du GS EDS. Un travail considérable de paramétrage du calcul des résultats et sorties d'OASIS sera nécessaire pour intégrer les nouvelles pondérations de la partie attribut, qui

avaient déjà des pondérations attribuées à leurs critères, ainsi que pour attribuer pour la première fois une pondération aux critères des CCP.

De plus, l'échelle de pondération choisie pour la sollicitation des experts (de 1, 2, 5, ou 10, 0 représentant un poids nul) diffère de celle initialement utilisée dans la version historique de la méthode OASIS (échelle de 1, 2 et 3). Il sera donc nécessaire de débattre du choix de l'échelle de pondération à adopter pour la version 2.0. Bien que la question n'ait pas été abordée dans le détail, il semble logique de conserver l'échelle utilisée dans les matrices de cette sollicitation, car cela éviterait toute adaptation ou transformation entre les nouvelles pondérations des CCP et attributs et la pondération actuelle d'OASIS. Le changement ne concernerait alors que les feuilles de calculs. De même, la pondération choisie pour la refonte présente des avantages et des inconvénients, mais elle a été retenue sur la base d'arguments solides comme l'avantage d'avoir un nombre pair des valeurs, et de permettre un choix plus large de pondération que celui proposé par la version 1.0.

On ne peut aborder la question de la partie pondération sans évoquer les sorties graphiques qui permettront de présenter les résultats à l'issue des évaluations. Cette partie étant complexe et longue à traiter, il a été décidé de la finaliser après le stage pour y consacrer le temps nécessaire. D'autant plus qu'il fallait d'abord stabiliser les réflexions sur les CCP et attributs à conserver et redimensionner. Un groupe de travail sera possiblement créé dans ce but, afin de se concentrer et d'élaborer au mieux de nouvelles sorties graphiques, qui jouent un rôle crucial dans la visualisation des résultats des évaluations OASIS et, par conséquent, dans la formulation des recommandations. Il a été jugé qu'il est préférable de finaliser tout ce qui concerne les matrices et la pondération en premier lieu, afin de maintenir la dynamique progressivement instaurée avec les experts.

Enfin, tous les commentaires présents sur les matrices et dans le questionnaire n'ont pas pu être pris en compte dans leur intégralité, mais ont été regroupés dans des documents distincts et seront traités par les groupes de travail du GS EDS dans les prochains mois.

Conclusion

Ce mémoire a exploré la refonte de la méthode OASIS, qui s'est appuyée sur une élicitation d'experts dans un but d'en améliorer l'utilité et la pertinence sur le terrain. Cette démarche a conduit à la production d'une quantité importante de données et à la révision de nombreux aspects de la méthode actuelle.

Parmi les modifications majeures, on peut citer la suppression des pondérations entre les points critiques, la suppression des sous-catégories de points critiques, et l'ajout de nouveaux indicateurs tels que « Moyen » et « Collaboration ». Plusieurs noms d'indicateurs ont été changés et redéfinis, comme « Outils » qui a été scindé en « Outils Laboratoire » et « Outils Terrain », « Recueil et circulation des données » qui est devenu « Collecte et qualité des données », et « Traitement et interprétation des données » qui est devenu « Analyse et interprétation des données ». Le point critique « Environnement » a également été supprimé.

Néanmoins, il est évident qu'une poursuite de l'élicitation d'expert est inévitable. Il reste nécessaire de clarifier plusieurs aspects, notamment le changement d'échelle de pondération, qui n'est pas encore certain de rester tel que proposé dans la méthode révisée (échelle 1, 2, 5, 10). Il faudra également discuter de la pondération des autres indicateurs, dont certains n'ont pas beaucoup changé par rapport à la méthode 1.0 (tel que « Analyse et interprétation des données », « Diffusion de l'information », « Sensibilité », « Spécificité », « Représentativité », « Acceptabilité », « Flexibilité/Adaptabilité », « Rapidité », « Stabilité/Durabilité »).

En outre, une redéfinition et un redimensionnement des indicateurs ayant beaucoup évolué (comme « Collecte et qualité des données », « Fiabilité », « Utilité/Utilisation », « Simplicité/Clarté ») sont indispensables. Il sera également essentiel de discuter des sorties graphiques, qui n'ont pas pu être abordées lors du second tour, ainsi que du paramétrage du calcul des résultats et des sorties de la méthode. La pertinence de certains critères, qui n'ont pas été discutés lors du premier tour, devra également être évaluée.

Ainsi, bien que des progrès significatifs aient été réalisés, la refonte de la méthode OASIS nécessite encore des efforts soutenus pour atteindre une version finalisée et pleinement opérationnelle. La poursuite des discussions entre experts sera cruciale pour affiner et valider les ajustements nécessaires, garantissant ainsi que la méthode reste pertinente et adaptable aux besoins actuel et futur des dispositifs de surveillance.

Liste des références bibliographiques

Adler, M. and Ziglio, E. (eds) (1996) *Gazing into the oracle: the Delphi method and its application to social policy and public health*. London: Jessica Kingsley Publishers.

Booto Ekionea, J.-P., Bernard, P. and Plaisent, M. (2011) ‘Consensus par la méthode Delphi sur les concepts clés des capacités organisationnelles spécifiques de la gestion des connaissances’, *Recherches qualitatives*, 29(3), p. 168. Available at: <https://doi.org/10.7202/1085878ar>.

Bordier, M. *et al.* (2019) ‘One Health Surveillance: A Matrix to Evaluate Multisectoral Collaboration’, *Frontiers in Veterinary Science*, 6, p. 109. Available at: <https://doi.org/10.3389/fvets.2019.00109>.

Boulkedid, R. *et al.* (2017) ‘Combien faut-il d’experts dans une enquête Delphi ? Étude de simulation’, *EPICLIN 2017 11e Conférence francophone d’Épidémiologie Clinique 24e Journée des statisticiens des Centres de lutte contre le cancer*, 65, p. S74. Available at: <https://doi.org/10.1016/j.respe.2017.03.056>.

Bronner, A. *et al.* (2015) ‘Quantitative and qualitative assessment of the bovine abortion surveillance system in France’, *Preventive Veterinary Medicine*, 120(1), pp. 62–69. Available at: <https://doi.org/10.1016/j.prevetmed.2015.02.019>.

Calba, C. *et al.* (2015) ‘Surveillance systems evaluation: a systematic review of the existing approaches’, *BMC Public Health*, 15(1), p. 448. Available at: <https://doi.org/10.1186/s12889-015-1791-5>.

Drewe, J.A. *et al.* (2015) ‘SERVAL: a new framework for the evaluation of animal health surveillance’, *Transboundary and Emerging Diseases*, 62(1), pp. 33–45. Available at: <https://doi.org/10.1111/tbed.12063>.

Dufour, B., Hendriks, P. and Thonnat, J. (2011) *Surveillance épidémiologique en santé animale*. 3e éd. Versailles: Éditions Quæ.

Dufour, B., Hendriks, P. and Toma, B. (2005) ‘Élaboration et mise en place de systèmes de surveillance épidémiologique des maladies à haut risque dans les pays développés.’, *Revue Scientifique et Technique Office International des Epizooties*. [Preprint].

Hendriks, P. *et al.* (2011) ‘OASIS: an assessment tool of epidemiological surveillance systems in animal health and food safety’, *Epidemiology and Infection*, 139(10), pp. 1486–1496. Available at: <https://doi.org/10.1017/S0950268811000161>.

Hoinville, L.J. *et al.* (2013) ‘Proposed terms and concepts for describing and evaluating animal-health surveillance systems’, *Preventive Veterinary Medicine*, 112(1–2), pp. 1–12. Available at: <https://doi.org/10.1016/j.prevetmed.2013.06.006>.

Liu, M. *et al.* (2017) ‘The Effect of Extreme Response and Non-extreme Response Styles on Testing Measurement Invariance’, *Frontiers in Psychology*, 8, p. 726. Available at: <https://doi.org/10.3389/fpsyg.2017.00726>.

Mesure de l'état de santé des populations et notification des données (2000). Rapport du Secrétariat EB107/8. Organisation Mondiale de la santé.

Muellner, P. *et al.* (2018) 'SurF: an innovative framework in biosecurity and animal health surveillance evaluation', *Transboundary and Emerging Diseases*, 65(6), pp. 1545–1552. Available at: <https://doi.org/10.1111/tbed.12898>.

P Baillette, Fallery, B. and Girard, A. (2013) *Baillette, P., Fallery, B., & Girard, A. (2013). La méthode Delphi pour définir les accords et les controverses : applications à l'innovation dans la traçabilité et dans le e-recrutement.*

Preston, C.C. and Colman, A.M. (2000) 'Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences', *Acta Psychologica*, 104(1), pp. 1–15. Available at: [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5).

Principes d'évaluation des réseaux de santé (1999). Paris: Agence nationale d'accréditation et d'évaluation en santé.

'Qu'est-ce que la salmonellose et comment s'en prémunir?' (2022) *ANSES* [Preprint]. Available at: <https://www.anses.fr/fr/content/qu%E2%80%99est-ce-que-la-salmonellose-et-comment-s%E2%80%99en-pr%C3%A9munir>.

Rodríguez-Prieto, V. *et al.* (2015) 'Systematic review of surveillance systems and methods for early detection of exotic, new and re-emerging diseases in animal populations', *Epidemiology and Infection*, 143(10), pp. 2018–2042. Available at: <https://doi.org/10.1017/S095026881400212X>.

Smith, P. (2019) 'Environmental Health', in *Encyclopedia of Environmental Health (Second Edition)*,.

'Systèmes de contrôle des aliments' (no date) *Organisation des Nations Unies pour l'alimentation et l'agriculture*, Sécurité sanitaire et qualité des aliments.

Tegegne, H.A. *et al.* (2023) 'OH-EpiCap: a semi-quantitative tool for the evaluation of One Health epidemiological surveillance capacities and capabilities', *Frontiers in Public Health*, 11, p. 1053986. Available at: <https://doi.org/10.3389/fpubh.2023.1053986>.

Annexes

Annexe 1 : Sorties graphiques de la méthode OASIS.

Annexe 2 : Questionnaire Kobotoolbox.

Annexe 3 : Matrice Excel de pondération des indicateurs.

Annexe 4 : Diagramme en barre de visualisation des critères participant au CCP « Objectifs » classés par ordre décroissant

Annexe 5 : Notice d'utilisation des matrices.

Annexe 6 : Guide d'utilisation du questionnaire et des matrices.

Annexe 7 : Lexique de définitions des points critiques, attributs et propositions d'attributs

Annexe 8 : Résultats des discussions de pondérations des CCP « Echantillonnage » et « Animation ».











Annexe 1.

Sorties graphiques de la méthode OASIS

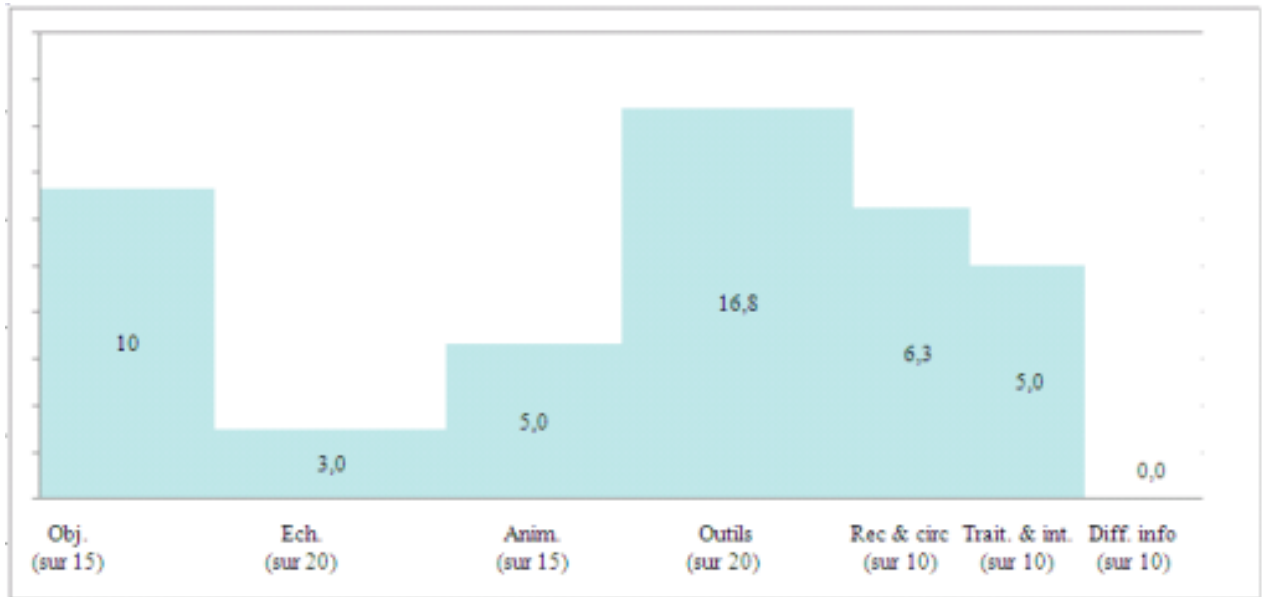
Les trois sorties graphiques de la méthode OASIS

Il s'agit de résultats fictifs

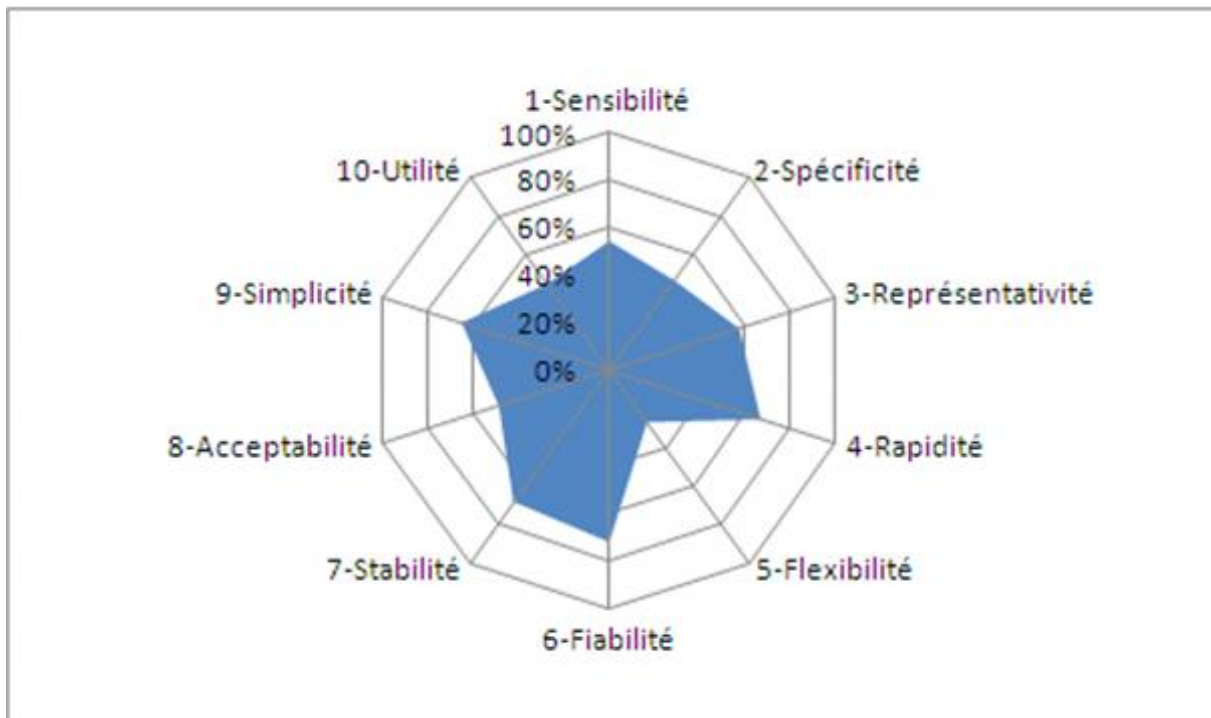
Sortie graphique 1 (sections fonctionnelles) : Camemberts

Section 1 : Objectifs et champ de la surveillance	
Section 2 : Organisation institutionnelle centrale	
Section 3 : Organisation institutionnelle de terrain	
Section 4 : Laboratoire	
Section 5 : Outils de surveillance	
Section 6 : Modalités de surveillance	
Section 7 : Gestion des données	
Section 8 : Formation	
Section 9 : Communication	
Section 10 : Evaluation	

Sortie graphique 2 (points critiques) : Histogrammes



Sortie graphique 3 (attributs) : Graphique en toile d'araignée (ou radar)



Annexe 2.

Questionnaire Kobotoolbox

Questionnaire OASIS 2.0

Ce questionnaire porte sur la révision de la méthode OASIS, et questionne plus particulièrement sur la pertinence et l'utilisation des points critiques, attributs, et de leur valorisation sous forme graphique. Merci de répondre à ce questionnaire, qui vous prendra environ 30 minutes, puis de compléter les matrices Excel qui vous ont été envoyées par ailleurs.

I - Points critiques

La méthode OASIS contient actuellement huit points critiques : **Objectifs, Échantillonnage, Animation, Outils, Environnement** (ce point critique n'est actuellement calculé que dans la grille des dangers sanitaires exotiques), **Recueil et circulation des données, Traitement et interprétation des données, Diffusion de l'information**. Leur compréhension, notation et pertinence sont primordiales pour la qualité des évaluations des systèmes de surveillance. Si besoin, les définitions de ces points critiques sont disponibles dans le lexique envoyé par ailleurs.

Q1. Durant vos évaluations OASIS, avez-vous déjà éprouvé des difficultés dans la compréhension de certains points critiques ?

- Oui
- Non

Si oui, précisez lesquels et pourquoi : ...

Q2. Actuellement, les huit points critiques d'OASIS ont une pondération, qui détermine la largeur des barres représentant chacun d'eux dans la sortie 2 (histogramme). Cette pondération est calculée en faisant le rapport du total des poids des critères qui constituent chaque point critique, sur le total des poids de tous les critères pour tous les points critiques. L'objectif de cette pondération est de refléter l'importance de chaque point critique dans l'évaluation d'un dispositif de surveillance. Pensez-vous que cette pondération soit pertinente ? Si besoin, les définitions de ces points critiques sont disponibles dans le lexique envoyé par ailleurs. (*Question à choix multiples*)

- Non, car je ne l'utilise pas pour hiérarchiser les recommandations
- Non, pour d'autres raisons (à préciser ci-après)
- Oui, car cela aide à hiérarchiser les recommandations
- Oui, car j'aimerais avoir une note globale synthétisant tous les CPP
- Oui, pour d'autres raisons à préciser ci-après)
- Ne se prononce pas

(Pour d'autres raisons) précisez :

Commentaires : ...

Si oui : affichage du tableau avec la répartition des poids (sur 100) à préciser en Q3

Q3. Pour chaque point critique, vous êtes invité(e) à considérer leur pertinence (**à supprimer, à conserver tel quel, à conserver mais à redéfinir/reformuler, ne se prononce pas**) et à répartir un

total de 100 points entre les points critiques selon l'importance que vous leur attribuez dans le cadre de l'évaluation d'un dispositif de surveillance. La somme des poids attribués à tous les points critiques doit être égale à 100 (une pondération équitable pour ces 8 points critiques correspondant à 12,5/100). Si besoin, les définitions de ces points critiques sont disponibles dans le lexique envoyé par ailleurs.

Points critiques	Pertinence	Répartition des poids (sur 100)	Justification (si suppression, ou redéfinition)
Objectifs		/100	
Echantillonnage		/100	
Animation		/100	
Outils		/100	
Environnement		/100	
Recueil et circulation des données		/100	
Traitement et interprétation des données		/100	
Diffusion de l'information		/100	

Commentaires :

si non : affichage du tableau sans la répartition des poids (sur 100).

Q3. Pour chaque point critique, vous êtes invité(e) à considérer leur pertinence (**à supprimer, à conserver tel quel, à conserver mais à redéfinir/reformuler, ne se prononce pas**).

Points critiques	Pertinence	Justification (si suppression, ou redéfinition)
Objectifs		
Echantillonnage		
Animation		
Outils		
Environnement		
Recueil et circulation des données		
Traitement et interprétation des données		
Diffusion de l'information		

Commentaires :

Q4. Actuellement, dans la méthode OASIS, les points critiques sont subdivisés en sous-catégories. Par exemple, le CCP « Animation » est subdivisé en quatre sous-catégories : **Mise en œuvre d'une action spécifique, Entretien de la sensibilisation, Modalités de l'animation et Adéquation du temps imparti.** Ou encore, le CCP « Outils » est subdivisé aussi en quatre sous-catégories : **Outils de mesure, Prélèvements, Analyses de laboratoire, Réactifs utilisés et Laboratoire.** Des pondérations sont attribuées à ces sous-catégories en fonction de leur importance estimée dans le point critique correspondant. Pensez-vous qu'il soit pertinent de conserver ces sous-catégories ?

- Oui, je pense qu'il est pertinent de conserver ces sous-catégories
- Non, je pense qu'il n'est pas pertinent de conserver ces sous-catégories
- Ne se prononce pas

Justifier :

Q5. Avez-vous des suggestions de points critiques à intégrer à OASIS ?

- Oui
- Non

Si oui : Indiquez lesquels ainsi qu'une proposition de définition associée. (*Pensez à reporter les suggestions de points critiques dans la colonne Suggestions de la matrice Excel « Points critiques » et à identifier dans cette matrice les critères qui d'après vous contribueraient à ces nouveaux points critiques ainsi que leur poids respectif*) :

II - Attributs

La méthode OASIS contient actuellement 10 attributs qui sont les suivants : **Sensibilité, Spécificité, Représentativité, Rapidité, Flexibilité, Fiabilité, Stabilité, Acceptabilité, Simplicité, Utilité.** Ils illustrent le niveau de qualité du système de surveillance évalué. Si besoin, les définitions de ces attributs sont disponibles dans le lexique envoyé par ailleurs.

Q6. Durant vos évaluations OASIS, avez-vous déjà éprouvé des difficultés dans la compréhension de certains attributs ?

- Oui
- Non

Si oui, précisez lesquels et pourquoi : ...

Q7. Actuellement, les 10 attributs d'OASIS n'ont aucune pondération attribuée (sortie graphique n°3 d'OASIS, en radar). Pensez-vous que pondérer les attributs en fonction de leur importance estimée dans un système de surveillance soit pertinent ? (*Question à choix multiples*)

- Non car je ne l'utiliserais pas pour hiérarchiser les recommandations
- Non pour d'autres raisons (à préciser ci-après)
- Oui cela aiderait à hiérarchiser les recommandations
- Oui car j'aimerais avoir une note globale synthétisant tous les attributs
- Oui, pour d'autres raisons (à préciser ci-après)
- Ne se prononce pas

(pour d'autres raisons), précisez :

Commentaires : ...

Si oui : affichage du tableau avec la répartition des poids (sur 100) à préciser en Q8

Q8. Pour chaque attribut, vous êtes invité(e) à considérer leur pertinence (**à supprimer, à conserver tel quel, à conserver mais à redéfinir, ne se prononce pas**) et à répartir **un total de 100 points** entre les différents attributs selon l'importance que vous leur attribuez dans le cadre de l'évaluation d'un dispositif de surveillance. La somme des poids attribués à tous les attributs doit être égale à 100 (une pondération équitable pour ces 10 attributs correspondant à 10/100).

Attributs	Pertinence	Répartition des poids (sur 100)	Justification (si suppression, ou redéfinition)
Sensibilité		/100	
Spécificité		/100	
Représentativité		/100	
Rapidité		/100	
Flexibilité		/100	
Fiabilité		/100	
Stabilité		/100	
Acceptabilité		/100	
Simplicité		/100	
Utilité		/100	

Commentaires :

Si non : affichage du tableau sans la répartition des poids (sur 100).

Q8. Pour chaque attribut, vous êtes invité(e) à considérer leur pertinence (à supprimer, à conserver tel quel, à conserver mais à redéfinir, ne se prononce pas).

Attributs	Pertinence	Justification (si suppression, ou redéfinition)
Sensibilité		
Spécificité		
Représentativité		
Rapidité		
Flexibilité		
Fiabilité		
Stabilité		

Acceptabilité		
Simplicité		
Utilité		

Commentaires : ...

Q9. Certains attributs sont actuellement utilisés dans d'autres méthodes d'évaluation, mais ne font pas partie de la méthode OASIS. Parmi la liste proposée ci-dessous, certains vous sembleraient-ils pertinents à ajouter à la méthode OASIS ? Si besoin, des propositions de définitions sont disponibles pour ces nouveaux attributs dans le lexique qui vous a été envoyé. (*Question à choix multiples*)

- Qualité des données
- Valeur Prédictive Positive (VPP)
- Communication
- Bénéfice
- Interopérabilité
- Complétude
- Faisabilité
- Efficience
- Efficacité
- Précision
- Je ne pense pas que cela soit pertinents d'ajouter de nouveau attributs à la méthode OASIS

Justifier (Pensez à reporter les possibles suggestions d'attributs dans la partie Suggestion de la matrice Excel « Attributs », et d'identifier les critères qui d'après vous contribuent à ces nouveaux attributs, ainsi que leur poids respectif) :

Q10. Souhaitez-vous ajouter d'autres attributs que ceux proposés ci-dessous ?

- Oui
- Non

Indiquez lesquels, ainsi qu'une proposition de définition associée. (*Pensez à reporter les suggestions d'attributs dans la partie Suggestion de la matrice Excel « Attributs », et d'identifier les critères qui d'après vous contribuent à ces nouveaux attributs ainsi que leur poids respectif*) : ...

Q11. Dans la littérature, les attributs sont parfois regroupés en catégories (par exemple : attributs d'efficacité, de fonctionnement, ...), mais ce n'est pas le cas de la méthode OASIS. Pensez-vous qu'il serait pertinent de regrouper les attributs en catégories, ou non ?

- Oui, je pense qu'il est pertinent de regrouper les attributs au sein de catégories
- Non, je ne pense pas qu'il soit pertinent de regrouper les attributs au sein de catégories
- Ne se prononce pas

Q11-bis Parmi les catégories présentées ci-dessous, identifiées dans la littérature, lesquelles considérez-vous comme pertinentes ? (*Question à choix multiples*)

- Opérationnel / fonctionnement
- Organisationnel
- Efficacité / Impacts
- Autres suggestions

Autres suggestions :

Commentaires :

Si les catégories ont été considérées pertinentes en question 11, on passe à la question 11bis sinon, question 12 directement.

Q11.5. Indiquez à quelle catégorie d'attribut vous attribueriez chaque attribut actuel de la méthode OASIS. Merci de ne sélectionner qu'une catégorie pour chaque attribut. Si certaines catégories ne vous conviennent pas, merci de l'indiquer dans les commentaires en bas de cette question.

Catégories /Attributs	Attributs opérationnels / de fonctionnement	Attributs organisationnels	Attributs d'efficacité / d'impacts
Sensibilité			
Spécificité			
Représentativité			
Rapidité			
Fiabilité			
Flexibilité			
Acceptabilité			
Simplicité			
Utilité			
Stabilité			

Commentaires : ...

Q12. Actuellement l'outil OASIS fait une distinction entre les points critiques et les attributs. Pensez-vous qu'il soit pertinent de continuer à faire cette distinction dans la méthode et dans la présentation des résultats ?

- Oui, je pense qu'il est pertinent de distinguer points critiques et attributs (dans des tableaux et figures différentes)
- Non, je pense qu'il serait plus pertinent de regrouper (dans un même tableau et dans une même figure) les points critiques et les attributs
- Ne se prononce pas

Justifier :

III - Sorties graphiques

La méthode OASIS contient actuellement trois sorties graphiques qui sont les suivantes : la sortie 1 sous forme de camemberts (correspondant à la représentation des sections fonctionnelles), la sortie 2 sous forme d'histogramme (correspondant à la représentation des points critiques) et la sortie 3 sous forme de graphique en toile d'araignée ou radar (correspondant à la représentation des attributs).

L'objectif de cette partie est de discuter de leur pertinence, maintien et des éventuelles perspectives d'évolutions.

SORTIE 1 : sections fonctionnelles

Q13. Sur une échelle de 0 à 10, à quel niveau estimez-vous le niveau de compréhension de la sortie graphique 1 de la méthode OASIS (la note 10 correspondant à une figure parfaitement compréhensible) ?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Q14. Sur une échelle de 0 à 10, à quel niveau estimez-vous le niveau de pertinence de la sortie graphique 1 (la note 10 correspondant à une figure parfaitement pertinente) ?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Q15. Pensez-vous qu'il soit pertinent de modifier la représentation graphique des sections fonctionnelles (sortie 1) (i.e. en représentant les mêmes données mais en utilisant une autre forme graphique)

- Oui
- Non
- Ne se prononce pas

(Si oui), Avez-vous des suggestions ? :

Q16. Pensez-vous que le niveau de détails fourni par la figure 1 soit suffisant ? Serait-il nécessaire d'afficher également les résultats par critères constituant les sections, pour avoir une figure plus complète.

- Oui
- Non
- Ne se prononce pas

(Si non), Précisez :

Commentaires : ...

SORTIE 2 : points critiques

Q17. Sur une échelle de 0 à 10, à quel niveau estimez-vous le niveau de compréhensibilité de la sortie graphique 2 de la méthode OASIS (la note 10 correspondant à une figure parfaitement compréhensible) ?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Q18. Sur une échelle de 0 à 10, à quel niveau estimez-vous le niveau de pertinence de la sortie graphique 2 sur les points critiques (la note 10 correspondant à une figure parfaitement pertinente) ?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Q19. Pensez-vous qu'il soit pertinent de modifier la représentation graphique des points critiques (i.e. en représentant les mêmes données mais en utilisant une autre forme graphique) ?

- Oui
- Non
- Ne se prononce pas

(Si oui), Avez-vous des suggestions ? :

Q20. Pensez-vous que le niveau de détails fourni par la figure 2 soit suffisant ?

- Oui
- Non
- Ne se prononce pas

(Si non), Précisez :

Commentaires : ...

SORTIE 3 : attributs

Q21. Sur une échelle de 0 à 10, à quel niveau estimez-vous le niveau de compréhension de la sortie graphique 3 de la méthode OASIS (la note 10 correspondant à une figure parfaitement compréhensible) ?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Q22. Sur une échelle de 0 à 10, à quel niveau estimez-vous le niveau de pertinence de la sortie graphique 3 concernant les attributs (la note 10 correspondant à une figure parfaitement pertinente) ?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Q23. Pensez-vous qu'il soit pertinent de modifier la représentation graphique des attributs (i.e. en représentant les mêmes données mais en utilisant une autre forme graphique) ?

- Oui
- Non
- Ne se prononce pas

(Si oui), Avez-vous des suggestions ? : ...

Q24. Pensez-vous que le niveau de détails fourni par la figure 3 soit suffisant ?

- Oui
- Non
- Ne se prononce pas

(Si non), Précisez :

Commentaires : ...

IV - Commentaires généraux sur les sorties graphiques

Q25. Pensez-vous qu'il soit pertinent d'ajouter une(des) nouvelle(s) sortie(s) graphique(s) ?

- Oui
- Non
- Ne se prononce pas

(Si oui), Quel(s) aspect(s) souhaiteriez-vous illustrer graphiquement ? : ...

Commentaires : ...

Q26. Pensez-vous qu'il y ait une redondance entre les différentes sorties graphiques ?

- Oui
- Non
- Ne se prononce pas

(Si oui), Merci de spécifier les sorties graphiques qui sont concernées et de proposer des améliorations/évolutions :

Q27. Pensez-vous qu'il soit pertinent que les points critiques et attributs soient regroupés ou séparés dans les sorties graphiques ?

- Regroupés
- Différenciés
- Ne se prononce pas

(Si non), Justifier :

Commentaires : ...

Annexe 3.

Matrice Excel de pondération des points critiques

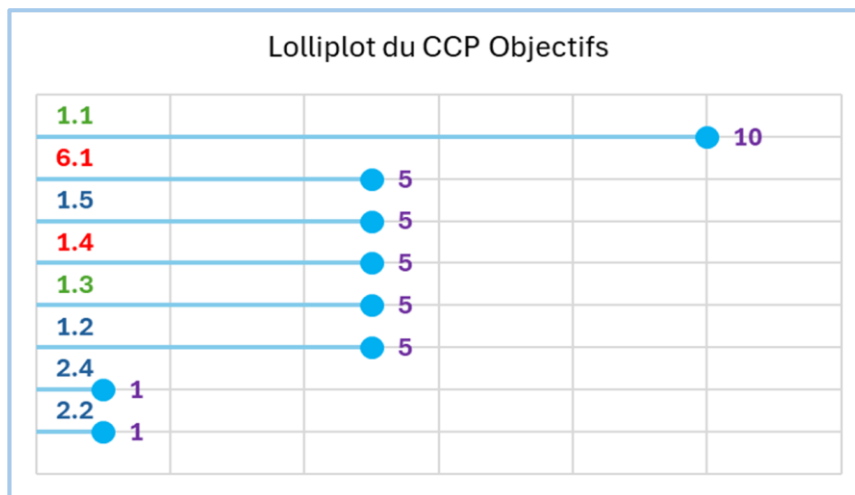
Matrice Excel de pondération des points critiques (exemple de la section 1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
		Pertinence des critères (Champ à compléter uniquement si vous pensez qu'il faut supprimer ou clarifier le critère)	Objectifs	Echantillonnage	Animation	Environnement	Outils	Recueil et circulation des données	Traitement et interprétation des données	Diffusion de l'information	Suggestion de point critique :	Suggestion de point critique :	Suggestion de point critique :	Commentaires
1														
2	Section 1 : Objectifs et champ de la surveillance													
3	1.1 Pertinence des objectifs de surveillance													
4	1.2 Niveau de détail, de précision et de formalisation des objectifs													
5	1.3 Prise en compte de l'attente des partenaires													
6	1.4 Cohérence des maladies surveillée avec situation sanitaire (maladies ou dangers existants / exotiques)													
7	1.5 Pertinence des objectifs du dispositif vis-à-vis de l'ensemble du système de surveillance													

Annexe 4.

Diagramme en barre de visualisation des critères participants au CCP « Objectifs » classés par ordre décroissant

Diagramme en barre (lollipop) de visualisation des critères participants au CCP « Objectifs » classés par ordre décroissant



Annexe 5.

Guide d'utilisation du questionnaire et des matrices

Guide d'utilisation du questionnaire et des matrices pour la refonte de la méthode OASIS (1er tour d'élicitation)

Cette étude est réalisée dans le cadre d'une démarche visant à améliorer l'outil OASIS, méthode semi-quantitative d'évaluation de dispositifs de surveillance. L'objectif premier est d'enrichir et de perfectionner cet outil, et nous sollicitons votre expertise et celle d'autres utilisateurs d'OASIS expérimentés afin d'y parvenir collectivement. Vos avis, fondés sur votre expérience de l'outil OASIS et votre domaine d'expertise spécifique (santé animale, santé végétale, sécurité de la chaîne alimentaire), nous permettront d'identifier les points à renforcer et les ajustements nécessaires afin de garantir qu'OASIS reste un outil de qualité pour l'évaluation des systèmes de surveillance.

Nous vous invitons donc à examiner attentivement chaque aspect de l'outil OASIS présenté dans le questionnaire en ligne et les matrices Excel et à partager vos réflexions, vos suggestions et vos critiques de manière constructive.

Ce document constitue un guide pour vous aider dans le déroulement du premier tour de la méthode Delphi utilisée pour solliciter vos avis, et qui repose d'une part sur un questionnaire à remplir en ligne et d'autre part sur des matrices multicritères concernant les points critiques et attributs de la méthode à compléter sous format Excel, et à renvoyer à porsan.fiona@gmail.com d'ici le **vendredi 22 mars 2024**.

Instructions générales - Partie 1 - Questionnaire

En partie 1 vous trouverez un questionnaire à compléter en ligne, divisé en trois parties distinctes :

- Points critiques
- Attributs
- Sorties graphiques

Le questionnaire est constitué de différents types de questions, notamment des QCM, des QROC et des matrices de notation.

Pour cette partie, il vous est demandé de répondre aux questions de manière aussi complète que possible (même si les questions ne sont pas à réponse obligatoire), et à apporter d'éventuels compléments ou précisions dans les zones libres de texte proposées.

Le questionnaire est accessible au lien suivant : <https://ee.kobotoolbox.org/x/ZlmmLDOE>

Instructions générales - Partie 2 - Matrices

En partie 2, nous vous demandons de compléter des matrices multicritères permettant d'identifier les critères, ainsi que leurs poids respectifs, qui contribuent aux points critiques et attributs identifiés dans la méthode (voire à de nouveaux indicateurs que vous pouvez proposer).

Pour cela, un fichier Excel vous a été envoyé par mail, comprenant trois onglets :

- le premier contient une notice succincte (à lire uniquement)
- le second concerne les points critiques (à compléter)
- le troisième concerne les attributs (à compléter)

Afin de répartir au mieux la charge de travail, nous avons décidé de vous assigner à chacun 6 attributs **OU** 6 points critiques de manière aléatoire (l'un ou l'autre uniquement). Les attributs/points critiques qui vous sont attribués sont en colonnes et sont indiqués en **bleu vif** sur les matrices (couleur de remplissage des en-têtes de colonnes concernées) : merci de travailler spécifiquement sur ces indicateurs qui vous sont attribués **en priorité**. Si vous en avez la possibilité, vous pouvez bien sûr compléter d'autres colonnes en supplément de celles qui vous ont été attribuées ! Nous vous en remercions par avance, car plus nous recevrons d'avis pour chaque point critique et attribut, plus l'élicitation sera robuste.

MATRICE POINTS CRITIQUES

Dans la matrice "Points critiques" (2e onglet Excel), nous souhaiterions avoir votre avis sur le choix et la pertinence de l'inclusion des 94 critères de notation OASIS dans l'évaluation des points critiques, ainsi que sur leurs importances respectives (poids).

- Les points critiques sont présentés en colonne, et les critères de la méthode OASIS en ligne. Afin de permettre une compréhension uniforme des points critiques entre répondants, leurs définitions sont disponibles dans le document joint intitulé "Lexique Points critiques et Attributs d'OASIS".
- Les cases avec un fond coloré (rosé) correspondent aux critères actuellement pris en compte dans le calcul de chaque point critique, dans la version actuelle de la méthode OASIS. Ils sont rappelés à titre d'information, vous n'êtes pas obligés d'en tenir compte dans votre réflexion.
- **Pour chaque point critique (colonne) :**
 1. Identifiez les critères qui selon vous participent au point critique considéré ;
 2. Attribuez à chaque critère que vous avez identifié une note (poids) représentant l'importance du critère pour le point critique considéré : vous avez le choix entre les poids **1, 2, 5 et 10** (10 étant le poids le plus important pouvant être attribué). Pour les critères qui ne vous semblent pas pertinents à prendre en compte pour un point critique, laissez la case vide (correspondant à un poids nul).
 3. Ainsi, remplir une case avec une notation revient à considérer le(s) critère(s) comme entrant en compte dans la notation du point critique. N'attribuez pas de notation dans les cases colorées si vous considérez que le critère (pris en compte dans la notation actuelle) ne devrait pas être pris en compte dans la notation révisée.
- Plusieurs lignes "**Suggestion de critère**" vous laissent la possibilité de proposer un à plusieurs critères à intégrer à OASIS (trois lignes vous sont proposées en bas du tableur mais il vous est possible d'en ajouter d'autres si nécessaire). Il vous sera également demandé dans ce cas de renseigner les poids qu'auraient ces critères pour chacun des points critiques auxquels ils devraient contribuer selon vous.
- Des colonnes "**Suggestion de point critique**" vous laissent la possibilité de proposer un à plusieurs points critiques supplémentaires à intégrer à OASIS (trois colonnes vous sont proposées sur la droite du tableur mais il vous est possible d'en ajouter d'autres si nécessaire). Vous pouvez donc y reporter vos suggestions de points critiques proposés dans la **partie Points Critiques du questionnaire** (question 5). Il vous sera également demandé dans ce cas d'attribuer un poids dans les cases correspondantes aux critères que vous considérez comme entrant en compte dans la notation des points critiques proposés.
- Une colonne "**Pertinence des critères**" (colonne B) vous permet, grâce à une liste déroulante, de juger si des critères devraient être supprimés ou clarifiés dans leurs formulations ou définitions. A noter que c'est la pertinence intrinsèque du critère qui est jugée ici (indépendamment de sa contribution à divers attributs et points critiques). Si le critère vous paraît pertinent en l'état, laissez cette colonne vide.

- Une colonne et une ligne “**Commentaires**” vous permettent de laisser des commentaires éventuels concernant les critères (en bas du tableur) et les points critiques (à droite du tableur) si vous le souhaitez.

MATRICE ATTRIBUTS

Dans la matrice “Attributs” (3e onglet Excel), nous souhaiterions avoir votre avis sur les critères à prendre en compte pour chaque attribut étudié, ainsi que leurs importances respectives (poids).

- Les attributs sont présentés en colonne, les critères de la méthode OASIS en ligne. Afin de permettre une compréhension uniforme des attributs entre répondants, leurs définitions sont disponibles dans le document joint intitulé “Lexique Points critiques et Attributs d’OASIS”.
- Les cases avec un fond coloré (rosé) correspondent aux critères actuellement pris en compte dans le calcul de chaque attribut, dans la version actuelle de la méthode OASIS. Ils sont rappelés à titre d’information, vous n’êtes pas obligés d’en tenir compte dans votre réflexion.
- **Pour chaque attribut (colonne) :**
 1. Identifiez les critères qui selon vous participent à cet attribut ;
 2. Attribuez à chaque critère que vous avez identifié une note (poids) représentant l’importance du critère pour l’attribut considéré : vous avez le choix entre les poids **1, 2, 5 et 10** (10 étant le poids le plus important pouvant être attribué). Pour les critères qui ne vous semblent pas pertinents à prendre en compte pour un attribut, laissez la case vide (correspondant à un poids nul).
 3. Ainsi, remplir une case avec une notation revient à considérer le(s) critère(s) comme entrant en compte dans la notation de l’attribut. N’attribuez pas de notes dans les cases colorées si vous considérez que le critère (pris en compte dans la notation actuelle) ne devrait pas être pris en compte dans la méthode de calcul révisée.
- Plusieurs lignes “**Suggestion de critère**” vous laissent la possibilité de proposer un à plusieurs critères à intégrer à OASIS (trois lignes vous sont proposées en bas du tableur mais il vous est possible d’en ajouter si nécessaire). Il vous sera également demandé dans ce cas d’attribuer un poids à ces critères pour chaque attribut auquel ils devraient contribuer selon vous.
- Des colonnes “**Suggestion d’attribut**” vous laissent la possibilité de proposer un ou plusieurs attributs à intégrer à OASIS (trois colonnes vous sont proposées sur la droite du tableur mais il vous est possible d’en ajouter d’autres si nécessaire). Vous pouvez ainsi reporter dans cette matrice vos suggestions d’attributs proposés dans la **partie Attributs du questionnaire** (questions 9 et 10). Il vous sera également demandé dans ce cas d’attribuer une note de pondération dans les cases correspondant aux critères que vous considérez comme entrant en compte dans la notation de(s) l’attribut(s) proposé(s).
- Une colonne “**Pertinence des critères**” (colonne B) vous permet, grâce à une liste déroulante, de juger si des critères devraient être supprimés ou clarifiés dans leurs formulations ou définitions. À noter que c’est la pertinence intrinsèque du critère qui est jugée ici (indépendamment de sa contribution à divers attributs et points critiques). Si le critère vous paraît pertinent en l’état, laissez cette colonne vide.
- Une colonne et une ligne “**Commentaires**” vous permettent de laisser un commentaire concernant les critères (en bas du tableur) et les attributs (à droite du tableur) si vous le souhaitez.

Nous vous remercions sincèrement pour votre participation et votre engagement dans cette démarche collaborative d’amélioration d’OASIS.

Annexe 6.

Lexique des définitions des points critiques, attributs et propositions d'attributs

Points critiques :

<u>Points Critiques</u>	<u>Définitions</u>
Objectifs	Ce point critique correspond à la pertinence, la cohérence et la précision des objectifs du système de surveillance, au regard de la situation à surveiller et des attentes des partenaires.
Échantillonnage	Ce point critique correspond à la qualité du processus de sélection et de collecte d'un échantillon des données d'intérêts à partir d'une population cible, au regard des objectifs du système de surveillance.
Animation	Ce point critique concerne l'implication de l'équipe chargée du bon fonctionnement et de la coordination du système de surveillance et la mise en œuvre d'actions, ainsi que la manière dont ces initiatives sont planifiées et réalisées dans le temps et la qualité de ces initiatives.
Environnement	Ce point critique correspond à la contribution du système de surveillance à la surveillance de l'environnement (pouvant inclure faune/flore sauvage, vecteurs, milieu naturel, environnements de production (hors matière première et produits finis), etc.).
Outils utilisés	Ce point critique correspond aux outils nécessaires à la gestion et à l'analyse des données, et plus généralement au fonctionnement du système de surveillance. Cela comprend tout ce qui est associé à la pertinence et à la qualité des outils de mesure, aux prélèvements, aux analyses de laboratoire, aux outils de communication entre acteurs, aux outils de traitement et d'interprétation des données, etc.
Recueil et circulation des données	Ce point critique correspond à la manière dont les données sont collectées, stockées et diffusées à travers le système de surveillance, au regard de ses objectifs. Il inclut les ressources disponibles (humaines et matérielles) associées.
Traitement et interprétation des données	Ce point critique concerne la manière dont les données sont traitées, nettoyées, et interprétées, au regard des objectifs du système de surveillance. Il inclut les ressources disponibles (humaines et matérielles) associées.
Diffusion de l'information	Ce point critique correspond à la manière dont les données sont diffusées à travers le système de surveillance et en externe (qualité de l'information, fréquence de diffusion, portée de la diffusion).

Attributs :

<u>Attributs</u>	<u>Définitions</u>
Sensibilité	<p>La sensibilité correspond à la capacité du système de surveillance à détecter les cas de la (des) maladie(s) ou les contaminations associées au(x) danger(s) sanitaire(s) surveillé(s) par le système. Selon l'objectif du système, la sensibilité peut se définir par la proportion d'unités épidémiologiques infectées détectées par le système de surveillance, ou par la capacité du système à détecter au moins un cas si la maladie/le danger atteint un certain niveau de prévalence.</p>
Spécificité	<p>La spécificité correspond à la capacité d'un système de surveillance à classer correctement les unités épidémiologiques indemnes (i.e. à bien les considérer indemnes). La spécificité est d'autant plus élevée que le système considère peu de cas à tort (peu de "faux positifs").</p> <p>Elle peut être calculée par la proportion d'unités épidémiologiques indemnes bien identifiées comme tel par le système (classées non infectées/non contaminées).</p> <p>Le complément à 1 de la spécificité est également souvent étudié, à savoir la proportion d'unités indemnes classées à tort comme infectées/contaminées par le système de surveillance (proportion de "faux-positifs").</p>
Représentativité	<p>La représentativité correspond à la capacité d'un système de surveillance à décrire correctement la survenue d'un phénomène de santé au cours du temps, ainsi que sa distribution dans la(les) population(s) étudiée(s), tant géographique que temporelle et en fonction des caractéristiques individuelles des unités surveillées.</p> <p>Un système de surveillance est représentatif s'il permet de collecter des données de surveillance qui reflètent les caractéristiques de la ou des population(s) concernée(s) au regard du danger surveillé.</p> <p>Un défaut de représentativité peut refléter d'éventuels biais inhérents au système de surveillance. Le caractère aléatoire de l'échantillonnage favorise une bonne représentativité.</p>
Rapidité	<p>La rapidité d'un système de surveillance peut correspondre à la rapidité d'exécution entre les différentes étapes de la surveillance.</p> <p>Elle correspond également à sa capacité à être réactif, à détecter et à signaler rapidement les cas ou événements surveillés, à assurer une transmission des informations avec le moins de délai possible, aussi bien de manière ascendante, que descendante ou transversale au sein du système de surveillance.</p>
Flexibilité	<p>La flexibilité d'un système de surveillance correspond à sa capacité à s'adapter et à répondre efficacement à l'évolution des besoins ou à des situations variées (changement dans la définition de cas, dans les outils utilisés, dans certaines étapes du fonctionnement du système, de la situation sanitaire, de la réglementation...) avec un minimum de temps, de personnel ou de fonds supplémentaires.</p>

Fiabilité	La fiabilité d'un système de surveillance correspond à celle des résultats qu'il fournit. Cela signifie que les données collectées et résultats produits par le système doivent être exactes, précises et cohérentes.
Stabilité	La stabilité d'un système de surveillance correspond à sa capacité à fonctionner de manière constante, efficace et sans défaillance pendant une période prolongée. Elle correspond également à la capacité d'un système de surveillance à être résilient face aux changements et perturbations externes, et à maintenir des performances constantes et fiables dans le temps.
Acceptabilité	L'acceptabilité correspond à l'adhésion des personnes et des organisations partenaires aux objectifs, à l'organisation, au fonctionnement et aux résultats du système de surveillance et à leur volonté de participer aux activités de ce système. Elle reflète l'implication ou l'engagement des parties prenantes (acteurs du système de surveillance) dans la planification, la conception et la mise en œuvre de la surveillance et ce, quelle que soit la situation (y compris défavorable pour l'intervenant). L'acceptabilité inclut également l'adhésion des utilisateurs des résultats de la surveillance et celle des bailleurs éventuels (manière dont les utilisateurs des données de surveillance et les bailleurs de fonds perçoivent, acceptent et soutiennent les résultats et les conclusions de la surveillance).
Simplicité	La simplicité d'un système de surveillance est liée à la facilité de compréhension et d'utilisation du système et de ses outils par les personnes chargées de collecter, de traiter et d'analyser les données (protocole et circuit de collecte clairs et définis, définitions de cas, procédures claires de gestion, centralisation et diffusion des données, etc.). Elle intègre également les outils et canaux de communication et les modalités de fonctionnement des instances. La simplicité d'un système de surveillance peut influencer l'acceptabilité de celui-ci par les acteurs, sa rapidité, sa stabilité et son efficacité, ainsi que la fiabilité, l'exhaustivité et l'harmonisation des données et des résultats produits. Un système de surveillance simple est un système dont les circuits de collecte, de centralisation des données et de diffusion de l'information permettent de comprendre rapidement et facilement les données collectées.
Utilité	L'utilité d'un système de surveillance se réfère à sa capacité à fournir des informations pertinentes pour soutenir la prise de décision et les actions de santé publique, (au sens large, santé animale et environnementale/végétale incluse), visant à prévenir, contrôler ou traiter la maladie ou le danger surveillé. Un système de surveillance utile doit être en mesure de fournir des données facilement exploitables qui permettent aux décideurs de comprendre la dynamique du danger sanitaire, d'identifier les tendances émergentes, de déterminer les populations à risque et d'évaluer l'efficacité des interventions sanitaires afin de les faire évoluer si nécessaire.

Suggestions d'attributs (existant dans la littérature et étant utilisés dans d'autres méthodes d'évaluations de dispositifs de surveillance)

<u>Attributs</u>	<u>Définitions</u>
Qualité des données	<p>La qualité des données correspond à l'exhaustivité, la validité et à la cohérence des données enregistrées en adéquation avec les besoins et objectifs du système de surveillance.</p> <p>Une donnée dite de qualité est une donnée adéquate pour l'usage que l'on souhaite en faire. Elle a plusieurs dimensions :</p> <ul style="list-style-type: none"> • la complétude ; • la validité (degré de conformité des données à la réalité, aux règles ou contraintes définies) dont les sous-composantes d'actualité, d'unicité et d'uniformité ; • la cohérence des données entre elles ; • la vitesse de circulation des données ; • la précision.
Valeur prédictive positive (VPP)	<p>La valeur prédictive positive (VPP) est la proportion d'unités détectées par le système de surveillance comme infectées ou contaminées, qui le sont réellement. En d'autres termes, la VPP mesure la capacité du système à identifier correctement les événements d'intérêt parmi tous ceux qui sont détectés comme tels (à tort ou à raison).</p>
Communication	<p>L'attribut "communication" d'un système de surveillance correspond à une communication interne et externe efficace, fluide et de qualité, incluant notamment la diffusion des informations aux décideurs et le partage d'informations en cas de suspicion/d'événements particuliers comme en "temps de paix".</p>
Bénéfice	<p>Avantages directs et indirects produits par le système de surveillance. Non limités à l'argent, les avantages peuvent inclure toute perte évitée grâce aux informations fournies par le système de surveillance.</p> <p>Cela peut correspondre à une meilleure utilisation des ressources, une amélioration de la production animale, végétale ou agroalimentaire, une amélioration de la santé publique (au sens large, santé animale et environnementale/végétale incluse) , de la biodiversité, une meilleure compréhension d'un danger sanitaire, l'amélioration de la détection précoce des menaces, l'augmentation des échanges commerciaux...</p>
Interopérabilité	<p>L'interopérabilité d'un système de surveillance correspond à sa capacité à fonctionner de manière transparente et efficace avec d'autres systèmes ou plateformes de surveillance, permettant ainsi l'échange de données, d'informations et de fonctionnalités entre eux.</p> <p>Un système interopérable est capable de communiquer, de partager des données et de coopérer avec d'autres systèmes, qu'ils soient similaires ou différents, sans nécessiter de modifications importantes.</p>

Complétude	La complétude d'un système de surveillance correspond à sa capacité à collecter et à fournir un ensemble exhaustif et pertinent de données en fonction des objectifs de surveillance établis. Cela implique que le système soit en mesure de rassembler toutes les informations nécessaires pour obtenir une compréhension approfondie et précise du phénomène sous surveillance, sans omettre des aspects significatifs ou des éléments essentiels.
Faisabilité	La faisabilité d'un système de surveillance correspond à sa capacité à être réalisable, c'est-à-dire à pouvoir être mis en œuvre et fonctionner de manière pratique, réaliste et efficace de manière durable, en considérant les ressources et compétences disponibles et les défis.
Efficience	L'efficience d'un système de surveillance correspond à sa capacité à remplir ses objectifs de manière optimale en utilisant de manière judicieuse les ressources disponibles. Cela implique d'atteindre les résultats souhaités avec un minimum de ressources (coûts, temps, matériels...) et d'efforts.
Efficacité	L'efficacité correspond à la capacité d'un système de surveillance à atteindre ses objectifs de manière satisfaisante et à produire les résultats attendus, en fonction de ses ressources et de ses contraintes, en répondant aux besoins des utilisateurs et en fournissant des informations pertinentes et fiables.
Précision	La précision d'un système de surveillance est sa capacité à fournir des résultats avec un intervalle de confiance étroit. La précision est influencée par la taille de l'échantillon, le niveau de confiance choisi, l'exhaustivité et l'exactitude des données.

Annexe 7.

Résultats issus des discussions de pondérations des CCP « Echantillonnage » et « Animation »

Evolutions de pondérations validées pour le points critique « Echantillonnage »

Référence du critère	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Proposition de pondération	Pondération adoptée en 2 nd tour
1.4	5	0	0	0	1	0	1	?
2.3	5	0	x	2	1	0	2	2
2.4	10	10	1	2	1	10	5	2
2.6	0	2	0	0	1	0	1	1
3.2	2	5	x	0	0	0	1	2
6.1	5	0	0	2	0	0	1	2
6.3	10	10	10	0	2	2	5	5
8.1	2	0	0	0	5	0	1	2

1.4 : Cohérence des maladies surveillées avec la situation sanitaire (maladie ou danger existants / exotiques).

2.3 : Existence d'un comité scientifique du dispositif

2.4 : Organisation et fonctionnement du réseau prévus par la réglementation, une charte ou convention entre partenaires

2.6 : Mise en place d'une supervision par l'échelon central

3.2 : Rôle actif des unités intermédiaires dans le fonctionnement du réseau (validation, animation, retour d'information)

6.1 : Adéquation des modalités de surveillance aux objectifs du dispositif

6.3 : Existence d'actions de sensibilisation des sources de données en réseau passif (événementiel)

8.1 : Niveau de compétence satisfaisant en épidémiologie des membres de l'unité centrale

X : représente le choix d'un expert de ne pas pondérer mais de la prise en compte de ce critère dans la notation du CCP

Tableau des évolutions de pondérations validées pour le points critique « Animation » 1/2

Référence du critère	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Proposition de pondération	Pondération adoptée en 2 nd tour
2.2	10	10	0	2	5	5	5	5
2.3	5	0	x	2	1	0	2	0
2.4	10	10	1	2	1	10	5	0
2.7	10	2	X	0	2	1	2	2
2.8	10	2	0	1	0	1	2	0
3.2	10	5	x	0	2	2	5	5
3.3	5	5	x	0	2	2	2	2
3.5	10	1	0	0	0	1	2	2
4.1	5	5	0	1	5	5	2	2

4.6	5	2	0	5	2	0	2	0
7.3	0	0	x	0	5	1	1	0
8.1	5	5	0	0	10	10	5	2
8.2	5	5	0	5	5	10	5	5
8.3	2	0	0	0	0	5	1	0
8.4	5	2	0	2	2	5	2	2
9.1	5	0	0	2	0	0	1	0
9.2	5	2	0	0	0	2	1	0
9.3	5	5	0	0	5	5	2	2
9.4	10	5	0	5	5	5	5	5
9.5	10	5	0	2	2	0	2	2
9.6	0	0	0	1	5	0	1	0

2.2 Existence d'une structure de pilotage fonctionnelle et représentative des partenaires (comité de pilotage)

2.8. Niveau de collaboration du niveau central

2.7 Suffisance des moyens matériels et financiers de l'échelon central

3.2 Rôle actif des unités intermédiaires dans le fonctionnement du réseau (validation, animation, retour d'information)

3.3 Mise en place d'une supervision par l'échelon intermédiaire

3.5 Suffisance des moyens matériels et financiers des UI

4.1 Intégration effective du laboratoire dans le dispositif de surveillance

4.6 Existence d'une équipe d'investigation pour appuyer les agents de terrain

7.3 Personnel spécifique disponible et qualifié pour la saisie, la gestion et l'analyse des données

8.1 Niveau de compétence satisfaisant en épidémiologie des membres de l'unité centrale

8.2 Formation initiale mise en œuvre pour tous les agents de terrain à leur entrée dans le dispositif

8.3. Objectifs et contenu de la formation initiale des acteurs de terrain du dispositif en adéquation avec les besoins opérationnels de la surveillance

8.4 Formations de perfectionnement régulières

9.1 Edition régulière de rapports et articles scientifiques sur les résultats de la surveillance

9.3 Diffusion régulière d'un bulletin d'information pertinent

9.4 Restitution systématique des bilans de résultats de terrain (hors bulletin)

9.5 Présence d'un système d'échange d'informations organisé transversalement et verticalement entre les acteurs de terrain (mail et /ou web)

9.6 Politique de communication externe solide

Tableau des évolutions de pondérations validées pour le points critique « Animation » 2/2

Référence du critère	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Proposition de pondération	Pondération adoptée en 2 nd tour
9.7	5	2	0	1	0	0	1	1
9.8	5	1	0	0	0	0	1	0
9.9	5	1	0	0	0	0	1	0
10.1	5	5	0	0	0	1	1	1
10.2	5	2	0	0	0	0	1	2
10.3	5	1	0	0	5	0	2	5
10.4	5	10	0	0	5	0	2	5

10.5	5	1	0	0	0	0	1	0
-------------	---	---	---	---	---	---	---	---

9.7 Suffisance des moyens humains, matériels et financiers pour la communication

9.8 Communication commune avec d'autres dispositifs de surveillance

9.9 Communication transversale entre dispositifs de surveillance

10.1 Système d'indicateurs de performance développé et validé par les responsables du réseau

10.2 Indicateurs de performance régulièrement calculés, interprétés et diffusés

*10.4 Mise en œuvre des mesures correctrices**

10.3 Evaluation externes effectuées