anses
agence nationale de sécurité sanitaire
alimentation, environnement, travail

Connaître, évaluer, protéger

Illustrations et actualisation des recommandations pour l'évaluation du poids des preuves et l'analyse d'incertitude à l'Anses

Avis de l'Anses
Rapport d'expertise collective

Octobre 2017 Édition scientifique

anses
agence nationale de sécurité sanitaire
alimentation, environnement, travail

Connaître, évaluer, protéger

Illustrations et actualisation des recommandations pour l'évaluation du poids des preuves et l'analyse d'incertitude à l'Anses

Avis de l'Anses
Rapport d'expertise collective

Octobre 2017 Édition scientifique

Le directeur général

Maisons-Alfort, le 18 octobre 2017

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

relatif au rapport « Illustrations et actualisation des recommandations pour l'évaluation du poids des preuves et l'analyse d'incertitude à l'Anses »

L'Anses met en œuvre une expertise scientifique indépendante et pluraliste.
L'Anses contribue principalement à assurer la sécurité sanitaire dans les domaines de l'environnement, du travail et de l'alimentation et à évaluer les risques sanitaires qu'ils peuvent comporter.
Elle contribue également à assurer d'une part la protection de la santé et du bien-être des animaux et de la santé des végétaux et d'autre part à l'évaluation des propriétés nutritionnelles des aliments.

Elle fournit aux autorités compétentes toutes les informations sur ces risques ainsi que l'expertise et l'appui scientifique technique nécessaires à l'élaboration des dispositions législatives et réglementaires et à la mise en œuvre des mesures de gestion du risque (article L.1313-1 du code de la santé publique).

Ses avis sont publiés sur son site internet.

L'Anses s'est autosaisie le 31 mars 2015 pour la réalisation d'une analyse critique sur les approches d'évaluation des niveaux de preuve à l'étape d'identification des dangers et la prise en compte des incertitudes dans les évaluations des risques sanitaires à l'Anses.

1. CONTEXTE ET OBJET DE LA SAISINE

L'Anses est chargée de mettre en œuvre une expertise scientifique indépendante et pluraliste, afin d'évaluer les risques dans les domaines de l'environnement, du travail, de l'alimentation, de la protection de la santé des végétaux, et de la protection de la santé et du bien-être des animaux. Les questions instruites par l'Agence sont traitées dans le cadre du processus d'évaluation des risques sanitaires, constitué des quatre étapes suivantes :

1. Identification des dangers,
2. Caractérisation des dangers (incluant l'établissement de la relation dose-réponse),
3. Evaluation de l'exposition,
4. Caractérisation des risques.

Les données utilisées pour répondre à ces questions proviennent de sources multiples (bases de données de la littérature scientifique, expertises de professionnels etc.), sont souvent de nature (*in vitro*, *in vivo*, toxicologiques, épidémiologiques, etc.) et de qualité hétérogènes. Elles peuvent également être contradictoires et parfois insuffisantes pour appréhender complètement la ou les questions posées.

L'utilisation des données dans le contexte de l'évaluation des risques pose deux difficultés méthodologiques majeures. La première est d'évaluer la pertinence et la validité des données disponibles, puis de combiner ces données en tenant compte de leur hétérogénéité pour répondre aux questions posées. La seconde est d'identifier les manques et les limites dans les

connaissances disponibles, et de décrire, quantifier et analyser les incertitudes associées aux résultats. La première difficulté relève de l'évaluation du poids des preuves, et la seconde de l'analyse d'incertitude. Ces sujets ont fait l'objet de deux rapports publiés par l'Anses en 2016.

La littérature fait principalement référence à l'évaluation du poids des preuves à l'étape d'identification du danger, dont l'objectif est d'identifier le type et la nature des effets néfastes qu'un agent (biologique, chimique ou physique) peut causer sur un organisme, un système ou une population. L'analyse d'incertitude concerne, par contre, les quatre étapes du processus d'évaluation des risques, et ne se focalise pas exclusivement sur l'étape d'identification des dangers.

Un aspect central de l'évaluation du poids des preuves concerne la recherche et la sélection des études pertinentes. Diverses approches peuvent être mobilisées dans ce but, notamment la revue systématique dont l'objectif est d'assembler, d'évaluer et de synthétiser de manière exhaustive toutes les études pertinentes, parfois contradictoires, qui abordent une question précise.

Une évaluation du poids des preuves inclut nécessairement une analyse d'incertitude, mais la réciproque n'est pas vraie. Lors de l'expression des conclusions concernant le poids des preuves, l'incertitude doit être caractérisée, soit sous forme qualitative (ex : score qualitatif ou texte explicatif), soit sous forme quantitative (ex : intervalle de confiance associé à une taille d'effet estimée par méta-analyse). Une analyse d'incertitude est donc réalisée pour toute évaluation du poids des preuves. Par contre, une analyse d'incertitude n'est pas systématiquement associée à une évaluation du poids des preuves. Elle peut être en effet réalisée en dehors du contexte du poids des preuves, par exemple lors de l'estimation d'une dose seuil à partir d'une étude expérimentale, ou lors de l'utilisation d'un modèle mathématique pour simuler un niveau d'exposition.

En résumé, l'évaluation du poids des preuves inclut une analyse d'incertitude, mais ne se réduit pas à ce type d'analyse ; elle s'intéresse également à d'autres aspects concernant la recherche et la sélection d'études pertinentes, et la combinaison de données hétérogènes en vue de déterminer le niveau de plausibilité d'hypothèses, essentiellement à l'étape d'identification du danger. L'analyse d'incertitude se focalise sur les manques et les limites dans les connaissances disponibles pour évaluer une situation en vue d'une prise de décision. Elle doit être considérée à toutes les étapes de l'évaluation des risques.

2. ORGANISATION DE L'EXPERTISE

2.1 Modalités de traitement des saisines

L'Anses a confié au groupe de travail « Méthodologie d'évaluation des risques » (GT MER), rattaché au conseil scientifique (CS), l'instruction de ces saisines.

Un rapport d'étape a été rédigé pour chaque autosaisine :

- Un rapport d'étape sur la saisine 2015-SA-0089 a été publié en juillet 2016 sur l'« évaluation du poids des preuves à l'Anses : revue critique de la littérature et recommandations à l'étape d'identification des dangers ».
- Un rapport d'étape sur la saisine 2015-SA-0090 a été publié en décembre 2016 sur la « prise en compte de l'incertitude en évaluation des risques : revue de la littérature et recommandations pour l'Anses ».

Ces rapports ont permis de répondre aux objectifs suivants : décrire les pratiques actuelles de l'Anses et les comparer avec celles d'autres organismes ou agences sanitaires, faire une revue

critique des différentes approches et élaborer des recommandations spécifiques pour les différentes étapes du processus d'évaluation du poids des preuves et d'analyse d'incertitude.

L'objectif du présent rapport est de clarifier et d'illustrer les recommandations grâce à des études de cas réalisées avec les collectifs d'experts de l'Anses en 2016-2017 et au séminaire du 20 janvier 2017 (qui comportait des ateliers participatifs pour les experts et les scientifiques de l'Anses sur les recommandations des deux rapports).

Le présent rapport a été validé par le GT MER. Il a été soumis au conseil scientifique le 20 juin 2017 et tient compte des observations et éléments complémentaires transmis par les membres du CS.

Le travail d'expertise présenté dans ce rapport a été conduit par un collectif d'experts intervenant dans les différents domaines de l'Agence et ayant des compétences dans les méthodes d'évaluation des risques.

L'expertise a été réalisée dans le respect de la norme NF X 50-110 « Qualité en expertise – Prescriptions générales de compétence pour une expertise (Mai 2003) ».

2.2 Prévention des risques de conflits d'intérêts.

L'Anses analyse les liens d'intérêts déclarés par les experts avant leur nomination et tout au long des travaux, afin d'éviter les risques de conflits d'intérêts au regard des points traités dans le cadre de l'expertise.

Les déclarations d'intérêts des experts sont publiées sur le site internet de l'Anses (www.anses.fr).

3. ANALYSE ET CONCLUSIONS DU CONSEIL SCIENTIFIQUE

3.1 La démarche d'évaluation du poids des preuves

Le poids des preuves est une synthèse formalisée de lignes de preuves, éventuellement de qualités hétérogènes, dans le but de déterminer le niveau de plausibilité d'hypothèses.

La démarche d'évaluation du poids des preuves est représentée dans la figure ci-dessous :

Figure 1 : Démarche d'évaluation du poids des preuves (étapes et sous-étapes)

3.2 La démarche d'analyse d'incertitude

L'incertitude est un manque ou une limite dans les connaissances disponibles pour évaluer une situation en vue d'une prise de décision. Le terme « décision » fait référence à une décision en vue d'une gestion des risques, les travaux menés par l'Anses s'inscrivant dans la perspective d'une aide à la décision en matière de gestion des risques.

La démarche d'analyse d'incertitude est présentée dans la figure ci-dessous :

Figure 2 : Démarche d'analyse d'incertitude

3.3 Les recommandations mises à jour

Afin de faciliter l'appropriation des concepts et méthodes sur l'évaluation du poids des preuves et l'analyse d'incertitude par les collectifs d'experts, le GT MER recommande :

- de veiller à la définition et à l'harmonisation du vocabulaire entre les différents collectifs,
- de mettre à disposition des collectifs d'experts un soutien méthodologique pour faciliter la mise en œuvre des méthodes d'évaluation du poids des preuves et d'analyse d'incertitude,
- de mettre en place un système d'information pour capitaliser les travaux d'expertise,
- d'intégrer explicitement les éléments relatifs à l'évaluation du poids des preuves et à l'analyse d'incertitude dans les documents du système qualité Anses.

3.3.1 <u>Planification de l'évaluation</u>

Que ce soit dans le cadre d'une évaluation du poids des preuves ou dans celui d'une analyse d'incertitude, le GT MER recommande la mise en place d'un processus complet pour la planification de l'évaluation des risques en 3 sous-étapes :

1. **Cadrage :** rédaction d'un document de cadrage de l'expertise précisant les enjeux, le contexte décisionnel et les ressources mobilisées *a priori*. En particulier, il convient d'expliciter de manière transparente les éléments susceptibles de guider les choix méthodologiques.
2. **Formulation de la ou des question(s),** en se basant sur un schéma conceptuel qui comprend une description graphique et une description narrative des principaux facteurs à prendre en compte.
3. **Choix et description de la méthode d'évaluation :** rédaction d'un plan décrivant les données, les informations, les méthodes et les modèles à utiliser. L'expression des résultats et modalités de communication doivent également être précisées à ce stade.

Le GT MER recommande d'adopter une démarche rétroactive entre les trois sous-étapes pour pouvoir éventuellement modifier le cadrage (sous-étape 1) et la formulation des questions (sous-étape 2) de manière à s'assurer que les questions puissent être correctement traitées à l'issue de la sous-étape 3.

Dans le cadre d'une évaluation du poids des preuves :

Le GT MER recommande, pour la sous-étape 2, de formaliser la ou les question(s) au moyen d'une structure de description présentée dans la littérature (PICO, PECO, etc.) ;

Le GT MER recommande, pour la sous-étape 3, de spécifier au minimum (à partir d'une revue préliminaire de la littérature et si besoin d'auditions) :

- Le choix du type de revue de la littérature : revue systématique ou revue approfondie pour traiter la ou les questions définies lors de la sous-étape 2. Ce choix doit tenir compte des éléments issus du cadrage, de l'analyse socio-économique de l'Anses et des critères proposés par l'EFSA : impact potentiel du résultat de la revue systématique, quantité et qualité des données, source et confidentialité des données, transparence nécessaire et controverse du sujet de la saisine, ressources à mettre en œuvre.
- Les critères utilisés pour évaluer la qualité des études par type d'étude.
- Les méthodes retenues pour évaluer un ensemble d'études et pour intégrer les lignes de preuves (approches qualitatives ou quantitatives).
- Les modalités d'expression des conclusions concernant le poids des preuves.

Dans le cadre d'une analyse d'incertitude :

Le GT MER recommande, pour la sous-étape 2, d'identifier dans le schéma conceptuel les éléments qui pourraient ne pas être pris en compte dans l'évaluation mais essentiels à l'analyse d'incertitude.

Le GT MER recommande, pour la sous-étape 3, de décrire les méthodes d'analyse d'incertitude mobilisées et les modalités d'expression des conclusions concernant l'incertitude.

3.3.2 <u>Etablissement des lignes de preuves</u>

Pour la sous-étape 1, recherche et sélection des études, le GT MER recommande de :

- développer une fiche de lecture adaptée à la question posée et au type d'étude scientifique. Un exemple est présenté en annexe du rapport associé à cet avis. La fiche inclura des éléments concernant la pertinence des études par rapport à la question posée et des éléments descriptifs de l'étude.
- recourir à au moins deux experts pour remplir les fiches de lecture.

- Lorsqu'une revue systématique est réalisée, d'appliquer certains grands principes : l'utilisation d'au moins deux bases de données, la sélection des études par deux personnes indépendantes, et la définition des critères de sélection et d'exclusion des études en amont. Les collectifs d'experts pourront s'appuyer sur le document guide EFSA (2010).
- Dans le cas où une revue systématique n'est pas réalisée (en cas de ressources insuffisantes par exemple), de réaliser une revue approfondie de la littérature et de décrire de manière aussi précise que possible dans le rapport d'expertise la procédure de recherche, de sélection et d'extraction, en se rapprochant des pratiques d'une revue systématique.

Pour la sous-étape 2, évaluation de la qualité des études, le GT MER recommande :

- l'utilisation de listes de critères formalisés par type d'étude (épidémiologiques, toxicologiques, etc.) de manière à assurer la transparence du processus d'évaluation. Un exemple est présenté en annexe du rapport associé à cet avis. Ces listes devront être élaborées en coordination avec les CES, en exploitant par exemple les critères mentionnés dans les méthodes OHAT et GRADE. Lorsque des études sont exclues, les critères d'exclusion doivent être explicites et définis par le collectif d'experts.
- lorsque des études de synthèse sont considérées dans le processus d'expertise, l'utilisation de la méthode AMSTAR et R-AMSTAR pour évaluer leur qualité.

Pour la sous-étape 3, analyse d'un ensemble d'études par type d'étude en vue d'établir des lignes de preuves, le GT MER recommande :

- lors de l'utilisation de méthodes qualitatives avec ou sans notation (ex : arguments en faveur de la causalité de Bradford Hill, méthode du CIRC, méthode dérivée de l'approche de Lewandovski et Rhomberg, OHAT) et dans le but d'assurer un niveau de transparence élevé au processus d'expertise, de définir aussi précisément que possible les critères utilisés et la signification des notes attribuées à ces critères par les experts.
- de privilégier la méta-analyse et d'en réaliser sur des sujets stratégiques, en tenant compte des enjeux en termes de risques sanitaires et des ressources disponibles (ressources humaines, disponibilité des données). Ces méta-analyses devraient être menées en s'inspirant des documents guides disponibles, notamment du document guide Cochrane.
- lors de l'utilisation de méthodes quantitatives, d'analyser la sensibilité des résultats aux paramètres (ou variables) d'entrée.

3.3.3 <u>Intégration des lignes de preuves</u>

- Il est important que les experts décrivent en détails et justifient les méthodes utilisées pour rendre l'expertise scientifique transparente.
- Pour la combinaison des lignes de preuves, le GT MER recommande l'utilisation des méthodes qualitatives proposées par le CIRC, le WCRF (en incluant explicitement une liste permettant de vérifier la prise en compte des critères de Hill) ou l'OHAT.
- Le GT MER souligne l'intérêt de la modélisation statistique pour combiner différentes lignes de preuve (notamment de l'approche bayésienne). Cette approche nécessitant des compétences spécifiques et un investissement en temps relativement important, la faisabilité de son utilisation devra cependant être évaluée sur des exemples concrets.
- Lorsqu'un modèle statistique est mobilisé, le GT MER recommande d'analyser la sensibilité des résultats aux hypothèses du modèle et à ses paramètres d'entrée.
- Dans le domaine spécifique de l'évaluation du mode d'action, la méthode d'évaluation comparative du poids des preuves pourrait être utilisée.

3.3.4 Identification et description des sources d'incertitude

Le GT MER propose d'adopter une typologie commune des sources d'incertitude composée des quatre classes suivantes :

I. sources d'incertitude liées au contexte, comprenant deux sous-classes (*cadrage* et *formulation de la question*),

II. sources d'incertitude liées au *corpus* de connaissances, incluant celles liées à la variabilité (intrinsèque et liée à l'hétérogénéité) des caractéristiques d'intérêt. Cette classe comprend trois sous-classes (*état des connaissances*, *méthode de collecte de données* et *modèles existants*),

III. sources d'incertitude liées à la méthode d'évaluation, comprenant trois sous-classes (*données sélectionnées*, *méthodes d'intégration des données* et *interprétation des résultats*)

IV. sources d'incertitude liées à la communication des résultats comprenant deux sous classes (*présentation des résultats* et *expression des résultats*).

Le GT MER recommande à l'étape 1 « Identifier et décrire l'ensemble des incertitudes » :

- d'effectuer un recensement le plus complet possible des sources d'incertitude en se basant sur la typologie proposée. Ce recensement inclura également les sources d'incertitude dont l'impact ne peut être évalué (par exemple la restriction d'une évaluation à quelques substances dans le cadre d'une pollution complexe ou un recueil de données insuffisant en raison d'un manque de temps ou de budget).
- de présenter ce recensement par un tableau d'incertitude construit à partir de la typologie proposée (lignes) et de dimensions (colonnes) à définir en collaboration avec les collectifs d'experts à travers des exemples. En particulier, l'impact des sources d'incertitude évaluées doit être mentionné dans le tableau d'incertitude.

3.3.5 Évaluation individuelle des sources d'incertitude

Pour estimer l'étendue des valeurs possibles, le GT MER recommande de distinguer les incertitudes liées à des hypothèses du modèle de celles liées à des paramètres ou variables d'entrée du modèle. Lorsque cela est possible, le GT MER recommande d'estimer les incertitudes individuelles à l'aide de méthodes statistiques et/ou probabilistes.

Le GT MER recommande à l'Anses de veiller à la définition et à l'harmonisation du vocabulaire entre les différents collectifs d'experts sur cette question, notamment lorsque des méthodes qualitatives, basées sur des descriptions nominales ou des échelles ordinales, sont appliquées.

3.3.6 Évaluation de l'impact combiné des sources d'incertitude

Le GT MER recommande :

- d'évaluer quantitativement l'impact combiné des incertitudes sur le résultat, lorsque les données le permettent ;
- de tenir compte des dépendances entre les variables dans l'évaluation de l'impact combiné de leurs incertitudes afin de se limiter aux combinaisons réalistes. Lorsque cela est possible, le GT MER recommande d'estimer les dépendances entre les variables à l'aide de méthodes statistiques.

3.3.7 Hiérarchisation des sources d'incertitude selon leur contribution à l'incertitude globale

Le GT MER recommande de réaliser, dans la mesure du possible et en fonction des besoins de l'évaluation, une analyse de sensibilité pour classer les sources d'incertitude.

3.3.8 Expression des conclusions

Le GT MER recommande de présenter systématiquement les conclusions d'une évaluation du poids des preuves et les conclusions d'une analyse d'incertitude dans des chapitres spécifiques, non seulement dans l'avis de l'Anses, mais aussi dans le rapport d'expertise collective associé.

Dans le cadre d'une évaluation du poids des preuves :

Le GT MER recommande :

- d'exprimer les conclusions concernant le poids des preuves soit de manière quantitative, soit de manière qualitative selon une classification en 4 niveaux correspondant à des niveaux de preuve croissants. Une classe supplémentaire « évaluation impossible » peut également être considérée. Le GT MER recommande que chaque niveau soit défini précisément dans les rapports d'expertise,
- de rédiger un texte explicatif répondant à la question posée, notamment lorsque le poids des preuves est exprimé de manière quantitative,
- de structurer le contenu de la conclusion en adaptant les travaux de la collaboration Cochrane et de la méthode GRADE aux domaines de l'agence. Un rapprochement avec la Collaboration Cochrane pourrait être envisagé afin de bénéficier de son expérience,
- de caractériser l'incertitude dans la conclusion concernant le poids des preuves, soit sous forme qualitative, soit sous forme quantitative, selon la méthode utilisée.

Dans le cadre d'une analyse d'incertitude :

Le GT MER recommande

- de présenter dans la conclusion l'impact combiné des incertitudes ainsi que les principales sources évaluées, en listant les sources d'incertitude qu'il n'a pas été possible d'analyser. Les points essentiels doivent être résumés à chaque étape de l'évaluation afin d'éclairer efficacement les parties prenantes,
- de présenter le recensement et la description des incertitudes dans un tableau,
- lorsque cela est possible et important pour l'évaluation, de présenter les résultats sous forme probabiliste telle que la probabilité de dépassement, l'intervalle de crédibilité, l'intervalle de confiance et à défaut, d'utiliser une échelle ordinale uniformisée.

3.4 Conclusions et recommandations du conseil scientifique

Le conseil scientifique souligne l'intérêt pour les collectifs d'experts de l'Agence des travaux du GT « Méthodologie de l'évaluation des risques » et endosse ce rapport. Il apprécie le caractère participatif de l'élaboration du rapport. Enfin, il estime nécessaire de valoriser ce rapport à travers des publications et de prévoir une communication spécifique auprès des parties prenantes.

4. CONCLUSIONS ET RECOMMANDATIONS DE L'AGENCE

L'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail adopte les conclusions de son Conseil scientifique et le rapport du groupe de travail « Méthodologie de l'évaluation des risques ».

Elle souligne le caractère stratégique et pionnier de ce travail initié dans le cadre d'une auto-saisine de l'Anses.

Les conclusions des experts validées par le Conseil scientifique ont vocation à structurer profondément la démarche et le cadre méthodologique nécessaires en matière d'évaluation du poids des preuves et d'analyse des incertitudes concernant les travaux d'évaluation des risques portés par l'Agence.

La structuration renforcée de ces approches s'avère en effet fondamentale au regard d'une part des pratiques parfois diverses selon les thématiques d'expertise et d'autre part de la nature des travaux de l'agence se déployant le plus souvent sur des questions profondément empreintes d'incertitudes. Cette structuration renforcera encore la transparence des travaux portés par l'Anses.

La mise en œuvre des conclusions du Conseil dans les pratiques de l'Agence suppose désormais leur intégration dans les référentiels et procédures structurant le travail d'expertise.

Un travail d'adaptation de ces procédures sera engagé rapidement. Au-delà des principes généraux, il devra aussi prendre en compte les modalités pratiques de pilotage des travaux au regard de la nature des questions posées mais aussi des calendriers de leur instruction ou encore des ressources et des compétences mobilisables.

Dr Roger Genet

MOTS-CLÉS

Analyse d'incertitude, étude de cas, incertitude, ligne de preuves, planification de l'évaluation, poids des preuves, source d'incertitude, transparence

KEY WORDS

Case study, line of evidence, planning and scoping, uncertainty, uncertainty analysis, source of uncertainty, transparency, weight of evidence

anses
agence nationale de sécurité sanitaire
alimentation, environnement, travail

Connaître, évaluer, protéger

Illustrations et actualisation des recommandations pour l'évaluation du poids des preuves et l'analyse d'incertitude à l'Anses

Saisines « n°2015-SA-0089 » et « n°2015-SA-0090 »

RAPPORT
d'expertise collective

Groupe de travail « Méthodologie d'évaluation des risques »

Juin 2017

Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail,
14 rue Pierre et Marie Curie, 94701 Maisons-Alfort Cedex
Téléphone : + 33 (0)1 49 77 13 50 - Télécopie : + 33 (0)1 49 77 26 26 - www.anses.fr

Mots clés

Analyse d'incertitude, étude de cas, incertitude, ligne de preuves, planification de l'évaluation, poids des preuves, source d'incertitude, transparence

Key Words

Case study, line of evidence, planning and scoping, uncertainty, uncertainty analysis, source of uncertainty, transparency, weight of evidence

Présentation des intervenants

PRÉAMBULE : Les experts membres de comités d'experts spécialisés, de groupes de travail ou désignés rapporteurs sont tous nommés à titre personnel, *intuitu personae,* et ne représentent pas leur organisme d'appartenance.

GROUPE DE TRAVAIL « MÉTHODOLOGIE D'ÉVALUATION DES RISQUES » (GT MER)

Président

M. David MAKOWSKI – Directeur de recherche à l'INRA – Évaluation des risques, statistique, modélisation

Membres

Mme Isabelle ALBERT – Chargée de recherche à l'INRA – Évaluation des risques alimentaires, statistique, microbiologie, modélisation

Mme Nathalie BONVALLOT – Enseignante chercheuse à l'EHESP – Évaluation des risques, toxicologie

Mme Soraya BOUDIA – Professeure à l'Université Paris Descartes – Histoire et sociologie des risques

Mme Céline BROCHOT – Chercheuse à l'INERIS – Modélisation, toxicocinétique, statistique

M. Olivier BRUYERE – Professeur à l'Université de Liège – Méthodologie, épidémiologie, santé publique

M. Philippe GLORENNEC – Enseignant chercheur à l'EHESP et à l'IRSET – UMR Inserm 1085 – Méthodologie d'évaluation des expositions et des risques sanitaires environnementaux

M. Pierre MARTIN – Chercheur au CIRAD – Santé végétale, knowledge management, modélisation informatique

Mme Bette MEEK – Directrice associée à l'Université d'Ottawa – Évaluation des risques des produits chimiques, toxicologie

M. Claude SAEGERMAN – Professeur à la Faculté de médecine vétérinaire, Université de Liège – Épidémiologie et analyse de risques appliquées aux sciences vétérinaires

Mme Mathilde TOUVIER – Chercheuse à l'INSERM – Épidémiologie nutritionnelle

Mme Laurence WATIER – Chargée de recherche à l'INSERM – Biostatistique, épidémiologie, santé publique

CONSEIL SCIENTIFIQUE

Les travaux, objets du présent rapport, ont été suivis et adoptés par le Conseil Scientifique le 20 juin 2017.

Présidente

Mme Isabelle MOMAS – Professeure des universités, directrice du département "Produits de santé et Santé publique" à la Faculté de Pharmacie de Paris, Université Paris Descartes

Membres de droit

La présidente du conseil scientifique de l'Agence nationale de sécurité du médicament et des produits de santé (ANSM), Mme Annick ALPEROVITCH, représentée par M. Robert BAROUKI,

professeur des universités, directeur d'unité à l'Institut national de la santé et de la recherche médicale (INSERM)

Le président du conseil scientifique de Santé Publique France (ANSP), représenté par M. Christian DUCROT, directeur de recherche à l'Institut national de la recherche agronomique (INRA)

Personnalités scientifiques compétentes

Mme Geneviève BENOIST – Directrice-médecin conseil à l'AFOMETRA.

M. Olivier BONATO – Chargé de recherche à l'Institut de recherche pour le développement.

Mme Soraya BOUDIA – Professeure de sociologie des risques à l'Université Paris Descartes.

M. Patrick BROCHARD – Professeur des universités - Praticien hospitalier en médecine et santé au travail à l'université de Bordeaux.

Mme Véronique COXAM – Directrice de recherche à l'Institut national de la recherche agronomique.

M. Jean-Pierre CRAVEDI – Directeur de recherche à l'Institut national de la recherche agronomique.

Mme Pascale DUCHÉ – Professeur en sciences et techniques des activités physiques et sportives à l'université Blaise Pascal de Clermont-Ferrand

Mme Brigitte ENRIQUEZ – Docteur vétérinaire, Professeure de Pharmacie-Toxicologie à l'Ecole nationale vétérinaire d'Alfort.

Mme Jeanne GARRIC – Directrice de recherche à l'Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture, directrice adjointe scientifique du département Eaux.

M. Philippe GIORDANENGO – Professeur en entomologie à l'Université de Picardie Jules Verne.

M. Jean-Paul HETTELINGH – Directeur du Centre de coordination pour les effets à l'Institut national de Santé Publique et de l'environnement aux Pays Bas (RIVM).

Mme Marie-Line ISKRA-CARUANA – Directrice de recherche au Centre de coopération internationale en recherche agronomique pour le développement (Cirad), chef d'équipe à l'Unité Mixte de Recherche Biologie et Génétique des Interactions Plante-Parasite.

M. Pierre-Benoit JOLY – Directeur de recherche à l'Institut national de la recherche agronomique.

M. Alain KAUFMANN – Directeur du département « Interface Sciences – Société » à l'Université de Lausanne.

M. Pierre KERKHOFS – Directeur général du Centre d'études et de recherches vétérinaires et agrochimiques (CODA-CERVA).

Mme Bette MEEK – Directrice associée d'évaluation des risques des produits chimiques à l'université d'Ottawa.

M. Christophe NGUYEN-THÉ – Directeur de recherche à l'Institut national de la recherche agronomique.

M. Philippe QUENEL – Directeur de laboratoire à l'École des hautes études en santé publique, directeur d'équipe UMR 1085 Inserm IRSET.

M. Jean-Marc SOULAT – Professeur des universités - praticien hospitalier en médecine et santé au travail à l'Université Toulouse III.

M. Noël TORDO – Chef de l'unité stratégies antivirales à l'Institut Pasteur de Paris et directeur de l'institut Pasteur de Guinée.

M. Xavier VAN HUFFEL – Directeur de la direction d'encadrement pour l'évaluation des risques à l'Agence fédérale belge pour la sécurité de la chaîne alimentaire (AFSCA).

PARTICIPATION ANSES

Coordination scientifique

Mme Claire BLADIER – Chef de projet scientifique – Direction de l'évaluation des risques

Contribution scientifique

Mme Claire BLADIER – Chef de projet scientifique – Direction de l'évaluation des risques

Mme Amélie CREPET – Chef de projet scientifique – Direction de l'évaluation des risques

Mme Sandrine FRAIZE-FRONTIER – Chef de projet scientifique – Direction de l'évaluation des risques

M. Moez SANAA – Conseiller scientifique pour l'évaluation des risques - Direction de l'évaluation des risques

Secrétariat administratif

Mme Virginie SADE – Anses

COMITE INTERNE DE METHODOLOGIE

Ces contributions concernent notamment la mise à disposition de données.

- Agence Nationale du Médicament Vétérinaire (ANMV)

M. Thierry GODARD – Expert Toxicologue

- Direction des produits réglementés (DEPR)

M. Yannick BAYONA - Evaluateur scientifique - Unité écotoxicologie et environnement des intrants du végétal

M. Jérémy DE SAINT JORES – Evaluateur scientifique - Unité évaluation toxicologie des produits réglementés

Mme Alexia FAURE - Evaluateur scientifique - Unité résidus et sécurité des aliments

- Direction de l'évaluation des risques (DER)

M. Laurent BODIN - Coordinateur scientifique – Unité d'évaluation des risques liés aux substances chimiques

Mme Karine PETIT – Coordinateur scientifique – Unité d'évaluation des risques liés à la santé, à l'alimentation et au bien-être des animaux

M. Matteo REDAELLI – Coordinateur scientifique – Unité d'évaluation des risques liés à l'air

Mme Anne THEBAULT – Chef de projet scientifique – Unité méthodologie et études

- Laboratoire de Santé des Végétaux (LSV) d'Angers

Mme Christine TAYEH – Coordinateur scientifique – Unité Expertise - Risques Biologiques

SOMMAIRE

Présentation des intervenants 3
Sigles et abréviations 8
Glossaire 9
Liste des tableaux 10
Liste des figures 10

1 Contexte et objectifs 11

1.1 Contexte 11
1.2 Objectifs 12
1.3 Modalités de traitement : moyens mis en œuvre et organisation 12
1.4 Prévention des risques de conflits d'intérêts. 13

2 Démarches d'évaluation du poids des preuves et d'analyse d'incertitude proposées par le GT MER 14

2.1 Evaluation du poids des preuves 14
2.2 Analyse d'incertitude 17

3 Méthode d'expertise : études de cas et séminaire 20

3.1 Organisation des études de cas 20
3.1.1 Etudes de cas sur l'évaluation du poids des preuves 20
3.1.2 Etudes de cas sur l'analyse d'incertitude 21
3.2 Organisation du séminaire 22
3.2.1 Des ateliers participatifs 22
3.2.2 Des posters pour illustrer les recommandations 23
3.2.3 Synthèse des méthodes mobilisées dans les études de cas et les posters 23

4 Mise à jour des recommandations du GT MER pour l'évaluation du poids des preuves et l'analyse d'incertitude 28

4.1 Planification de l'évaluation 28
4.1.1 Recommandations initiales 28
4.1.2 Clarification et exemples d'application 29
4.1.3 Recommandations mises à jour 32
4.2 Établissement des lignes de preuve 33
4.2.1 Recherche et sélection des études 33
4.2.1.1 Recommandations initiales 33
4.2.1.2 Clarification et exemples d'application 33
4.2.1.3 Recommandations mises à jour 34
4.2.2 Évaluation de la qualité des études 34
4.2.2.1 Recommandations initiales 34
4.2.2.2 Clarification et exemples d'application 34
4.2.2.3 Recommandations mises à jour 35
4.2.3 Analyse d'un ensemble d'études par type d'étude en vue d'établir des lignes de preuve 35
4.2.3.1 Recommandations initiales 35
4.2.3.2 Clarification et exemples d'application 36
4.2.3.3 Recommandations mises à jour 36

4.3 Intégration des lignes de preuves 37
4.3.1 Recommandations initiales 37
4.3.2 Clarification et exemples d'application 37
4.3.3 Recommandations mises à jour 38
4.4 Identification et description des sources d'incertitude 38
4.4.1 Recommandations initiales 38
4.4.2 Clarification et exemple d'application 38
4.4.3 Recommandations mises à jour 41
4.5 Évaluation individuelle des sources d'incertitude 42
4.5.1 Recommandations initiales 42
4.5.2 Clarification et exemples d'application 42
4.5.3 Recommandations mises à jour 43
4.6 Évaluation de l'impact combiné des sources d'incertitude 44
4.6.1 Recommandations initiales 44
4.6.2 Clarification et exemples d'application 44
4.6.3 Recommandations mises à jour 45
4.7 Hiérarchisation des sources d'incertitude selon leur contribution à l'incertitude globale 45
4.7.1 Recommandations initiales 45
4.7.2 Clarification et exemples d'application 45
4.7.3 Recommandation mise à jour 46
4.8 Expression des conclusions 46
4.8.1 Recommandations initiales 46
4.8.2 Clarification et exemples d'application 46
4.8.3 Recommandations mises à jour 47
4.9 Harmonisation des pratiques de l'Anses et partage d'expériences 48
4.9.1 Harmonisation du vocabulaire 48
4.9.2 Partage d'expériences 48
4.9.3 Relations entre les recommandations du GT MER et le système qualité 49

5 Conclusions du groupe de travail 50

6 Bibliographie 51

ANNEXES 54

Annexe 1 : Liste des recommandations (résumées) et des posters associés 55
Annexe 2 : Principales demandes de clarification par thématique (séminaire du 20/01/2017) 56
Annexe 3 : Applicabilité des recommandations par thématique 59
Annexe 4 : Grille d'évaluation des études épidémiologiques pour l'élaboration de VLEP 60

Sigles et abréviations

ARP	Analyse de risque phytosanitaire
CS	Conseil scientifique
DER	Direction de l'évaluation des risques
DEPR	Direction de l'évaluation des produits réglementés
GT	Groupe de travail
GT MER	Groupe de travail « méthodologie de l'évaluation des risques »
LSV	Laboratoire de la santé des végétaux
OHAT	Office of Health Assessment and Translation (office d'évaluation sanitaire et de traduction)
PDP	Poids des preuves
PECO	Patient, problem or population ; Exposure ; Comparison, control or comparator ; Outcomes (patient, problème ou population ; exposition ; comparaison, contrôle ou comparateur ; effets)
PECOTS	Populations, Exposures, Comparators, Outcomes, Timings, Settings of interest (Populations ; expositions ; comparateurs ; effets ; temps; localisation d'intérêt)
PICO	Patient, problem or population ; Intervention ; Comparison, control or comparator ; Outcomes (patient, problème ou population ; intervention ; comparaison, contrôle ou comparateur ; effets)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses (lignes directrices pour l'écriture des revues systématiques et des méta-analyses)
UERA	Unité d'évaluation des risques liés à l'air
UERALIM	Unité d'évaluation des risques liés aux aliments
UERAP	Unité d'évaluation des risques liés aux agents physiques
UERB	Unité expertise sur les risques biologiques
UERSABA	Unité d'évaluation des risques liés à la santé, à l'alimentation et au bien-être des animaux
UESC	Unité d'évaluation des substances chimiques
UETPR	Unité évaluation toxicologie des produits réglementés
UME	Unité méthodologie et études
VLEP	Valeur limite d'exposition professionnelle

Glossaire

- **Analyse d'incertitude** (Anses, 2016b) : dans un contexte d'évaluation des risques, une analyse d'incertitude est un processus ayant pour objectif d'identifier, décrire, quantifier et communiquer les incertitudes associées aux résultats.
- **Elicitation des connaissances d'un expert** : l'élicitation d'un expert est une action qui permet d'aider un expert à formaliser ses connaissances sur un objet (ex : un paramètre d'un modèle). Dans le cadre d'une élicitation probabiliste, le résultat est présenté sous la forme d'une distribution de probabilité.
- **Incertitude** (ANSES 2016c) : l'incertitude est un manque ou une limite dans les connaissances disponibles pour évaluer une situation en vue d'une prise de décision. Le terme « décision » fait référence à une décision en vue d'une gestion des risques, les travaux menés par l'Anses s'inscrivant dans la perspective d'une aide à la décision en matière de gestion des risques.
- **Ligne de preuves** (ANSES 2016b) : une ligne de preuves est un ensemble d'informations de même nature, intégrées pour évaluer une hypothèse.
- **Méta-analyse** : deux types de méta-analyse sont distingués (Tudur Smith et al. 2016)
    - Une **méta-analyse sur données agrégées** est une démarche statistique combinant les résultats d'une série d'études indépendantes sur une thématique donnée. Elle permet une synthèse systématique et quantifiée de données agrégées (ex : moyenne et variances des données élémentaires) portant sur cette thématique. Elle est systématique car elle implique une recherche exhaustive de toutes études disponibles sur un sujet donné. Elle est quantifiée car elle se base sur des analyses statistiques.
    - La **méta-analyse sur données individuelles** est une analyse poolée de données individuelles non agrégées issues d'un ensemble des essais portant sur un sujet d'intérêt. Ce type de méta-analyse est donc réalisé directement à partir des informations concernant les individus eux-mêmes, et pas à partir des données résumées relatives à des groupes de sujets. Ce type de méta-analyse suppose que les données individuelles soient disponibles.
- **Poids des preuves** (ANSES 2016b) : le poids des preuves est une synthèse formalisée de lignes de preuves, éventuellement de qualités hétérogènes, dans le but de déterminer le niveau de plausibilité d'hypothèses.
- **Revue systématique** (ANSES 2016b) : une revue systématique de la littérature scientifique consiste à assembler, évaluer et synthétiser de manière exhaustive toutes les études pertinentes, parfois contradictoire, qui abordent une question précise. Une revue systématique est basée sur la rédaction d'un protocole détaillé au préalable favorisant la transparence de la démarche et sa reproductibilité.
- **Schéma conceptuel** : un schéma conceptuel est une représentation graphique accompagnée d'un texte explicatif, comportant les objets de la question posée, à savoir les agents d'intérêt, les populations exposées, les effets sanitaires et les relations entre ces différents éléments.
- **Source d'incertitude** : En évaluation des risques, la source – i.e. l'origine ou la cause – de l'incertitude peut être due à un manque ou une limite des connaissances sur les voies d'exposition considérées, la population cible de l'évaluation, le scénario d'exposition, les modèles utilisés, les données utilisées, etc.
- **Variabilité** (ANSES 2016c) : la variabilité d'une caractéristique dans un référentiel prédéfini (individu, groupe ou population) reflète la différence observée entre les valeurs de cette caractéristique dans le référentiel choisi.

Liste des tableaux

Tableau 1 : Méthodes recensées pour chaque étape du processus d'évaluation du poids des preuves 16

Tableau 2 : Liste des méthodes pouvant être utilisées, pour chaque étape de l'analyse d'incertitude, selon le type de connaissances traitées..... 19

Tableau 3 : Etudes de cas par étape de l'évaluation du poids des preuves 21

Tableau 4 : Etudes de cas par étape de l'analyse d'incertitude 22

Tableau 5 : liste des posters illustrant les recommandations par domaine de l'Anses 23

Tableau 6 : Applications de différentes méthodes relatives à l'évaluation du poids des preuves à l'Anses... 24

Tableau 7 : Applications de différentes méthodes relatives à l'analyse d'incertitude à l'Anses..... 26

Tableau 8 : Exemple de protocole de méta-analyse (étude de cas PDP2) 31

Tableau 9 : exemple de présentation des incertitudes relatives à l'identification des dangers dans le cadre de l'évaluation d'un produit réglementé (étude de cas AI3) 40

Tableau 10 : exemple de présentation des incertitudes relatives à l'évaluation de l'exposition dans le cadre de l'évaluation d'un produit réglementé (étude de cas AI3) 40

Tableau 11 : exemple de tableau d'incertitudes (étude de cas AI4) 41

Liste des figures

Figure 1 : Démarche d'évaluation du poids des preuves (étapes et sous-étapes) 15

Figure 2 : Démarche d'analyse d'incertitude 17

Figure 3 *:* PECOTS évaluation des effets sanitaires des particules de l'air ambiant (Etude de cas PDP5) .. 30

Figure 4 : principaux éléments du processus d'entrée potentielle de *Salmonella* (étude de cas AI1)..... 31

Figure 5 : Schéma conceptuel du maillon portuaire (étude de cas AI 1)..... 31

Figure 6 : Choix d'une méthode d'élicitation pour l'autosaisine 2015-SA-0087 (poster 11) 43

Figure 7: schéma conceptuel séparant variabilité et incertitude (poster 13)..... 44

1 Contexte et objectifs

1.1 Contexte

L'Anses est chargée de mettre en œuvre une expertise scientifique indépendante et pluraliste, afin d'évaluer les risques dans les domaines de l'environnement, du travail, de l'alimentation, de la protection de la santé des végétaux, et de la protection de la santé et du bien-être des animaux. Les questions instruites par l'Agence sont traitées dans le cadre du processus d'évaluation des risques sanitaires, constitué des quatre étapes suivantes :

i. Identification des dangers,

ii. Caractérisation des dangers (incluant l'établissement de la relation dose-réponse),

iii. Evaluation de l'exposition,

iv. Caractérisation des risques.

Les données utilisées pour répondre à ces questions proviennent de sources multiples (bases de données de la littérature scientifique, expertises de professionnels etc.), sont souvent de nature (*in vitro*, *in vivo*, toxicologiques, épidémiologiques, etc.) et de qualité hétérogènes. Elles peuvent également être contradictoires et parfois insuffisantes pour appréhender complètement la ou les questions posées.

L'utilisation des données dans le contexte de l'évaluation des risques pose deux difficultés méthodologiques majeures. La première est d'évaluer la pertinence et la validité des données disponibles, puis de combiner ces données en tenant compte de leur hétérogénéité pour répondre aux questions posées. La seconde est d'identifier les manques et les limites dans les connaissances disponibles, et de décrire, quantifier et analyser les incertitudes associées aux résultats. La première difficulté relève de l'évaluation du poids des preuves, et la seconde de l'analyse d'incertitude. Ces sujets ont fait l'objet de deux rapports publiés par l'Anses en 2016 (ANSES 2016b, c)

La littérature fait principalement référence à l'évaluation du poids des preuves à l'étape d'identification du danger, dont l'objectif est d'identifier le type et la nature des effets néfastes qu'un agent (biologique, chimique ou physique) peut causer sur un organisme, un système ou une population (IPCS 2004). C'est également pour **l'étape d'identification du danger** que l'Anses a rédigé son rapport sur le poids des preuves en 2016 (ANSES 2016b). L'analyse d'incertitude concerne, par contre, les quatre étapes du processus d'évaluation des risques, et ne se focalise pas exclusivement sur l'étape d'identification des dangers.

Un aspect central de l'évaluation du poids des preuves concerne la recherche et la sélection des études pertinentes. Diverses approches peuvent être mobilisées dans ce but, notamment la revue systématique dont l'objectif est d'assembler, d'évaluer et de synthétiser de manière exhaustive toutes les études pertinentes, parfois contradictoires, qui abordent une question précise (ANSES 2016b).

Une évaluation du poids des preuves inclut nécessairement une analyse d'incertitude, mais la réciproque n'est pas vraie. Lors de l'expression des conclusions concernant le poids des preuves, l'incertitude doit être caractérisée, soit sous forme qualitative (ex : score qualitatif ou texte explicatif), soit sous forme quantitative (ex : intervalle de confiance associé à une taille d'effet estimée par méta-analyse). Une analyse d'incertitude est donc réalisée pour toute évaluation du poids des preuves. Par contre, une analyse d'incertitude n'est pas systématiquement associée à

une évaluation du poids des preuves. Elle peut être en effet réalisée en dehors du contexte du poids des preuves, par exemple lors de l'estimation d'une dose seuil à partir d'une étude expérimentale, ou lors de l'utilisation d'un modèle mathématique pour simuler un niveau d'exposition.

En résumé, l'évaluation du poids des preuves inclut une analyse d'incertitude, mais ne se réduit pas à ce type d'analyse ; elle s'intéresse également à d'autres aspects concernant la recherche et la sélection d'études pertinentes, et la combinaison de données hétérogènes en vue de déterminer le niveau de plausibilité d'hypothèses, essentiellement à l'étape d'identification du danger. L'analyse d'incertitude se focalise sur les manques et les limites dans les connaissances disponibles pour évaluer une situation en vue d'une prise de décision. Elle doit être considérée à toutes les étapes de l'évaluation des risques.

1.2 Objectifs

Dans l'objectif d'améliorer la transparence du processus d'expertise requise pour un organisme certifié ISO 9001, l'Anses a confié au Groupe de Travail « Méthodologie de l'Évaluation des Risques » (GT MER) la conduite d'une analyse critique sur les approches d'évaluation du poids des preuves et d'analyse d'incertitude.

Par décision en date du 31 mars 2015, l'Anses s'est autosaisie sur ces deux questions.

Un rapport d'étape a été rédigé pour chaque autosaisine :

- Un rapport d'étape sur la saisine 2015-SA-0089 a été publié en juillet 2016 sur l'« évaluation du poids des preuves à l'Anses : revue critique de la littérature et recommandations à l'étape d'identification des dangers ».
- Un rapport d'étape sur la saisine 2015-SA-0090 a été publié en décembre 2016 sur la « prise en compte de l'incertitude en évaluation des risques : Revue de la littérature et recommandations pour l'Anses ».

Ces rapports ont permis de répondre aux objectifs suivants : décrire les pratiques actuelles de l'Anses et les comparer avec celles d'autres organismes ou agences sanitaires, faire une revue critique des différentes approches et élaborer des recommandations spécifiques pour les différentes étapes du processus d'évaluation du poids des preuves et d'analyse d'incertitude.

L'objectif du présent rapport est de clarifier et d'illustrer les recommandations grâce à des études de cas réalisées avec les collectifs d'experts de l'Anses en 2016-2017 et au séminaire du 20 janvier 2017 (qui comportait des ateliers participatifs pour les experts et les scientifiques de l'Anses sur les recommandations des deux rapports).

1.3 Modalités de traitement : moyens mis en œuvre et organisation

L'Anses a confié au groupe de travail « Méthodologie d'évaluation des risques » (GT MER), rattaché au Conseil scientifique (CS), l'instruction de cette saisine.

Le présent rapport a été validé par le GT MER. Il a été soumis au conseil scientifique le 20 juin 2017 et tient compte des observations et éléments complémentaires transmis par les membres du CS.

Ces travaux sont issus d'un collectif d'experts aux compétences complémentaires.

L'expertise a été réalisée dans le respect de la norme NF X 50-110 « Qualité en expertise – prescriptions générales de compétence pour une expertise (Mai 2003) » (AFNOR 2003).

1.4 Prévention des risques de conflits d'intérêts.

L'Anses analyse les liens d'intérêts déclarés par les experts avant leur nomination et tout au long des travaux, afin d'éviter les risques de conflits d'intérêts au regard des points traités dans le cadre de l'expertise.

Les déclarations d'intérêts des experts sont publiées sur le site internet de l'agence (www.anses.fr).

2 Démarches d'évaluation du poids des preuves et d'analyse d'incertitude proposées par le GT MER

Les démarches d'évaluation du poids des preuves et d'analyse d'incertitude présentées ci-dessous ont été extraites des rapports Anses publiés en 2016 « Évaluation du poids des preuves à l'Anses : revue critique de la littérature et recommandations à l'étape d'identification des dangers » et « Prise en compte de l'incertitude en évaluation des risques : Revue de la littérature et recommandations pour l'Anses » **(ANSES 2016b, c)**. Les principales définitions ont été reprises dans le glossaire.

2.1 Evaluation du poids des preuves

Le GT MER définit le poids des preuves comme une synthèse formalisée de lignes de preuves, éventuellement de qualités hétérogènes, dans le but de déterminer le niveau de plausibilité d'hypothèses. Une ligne de preuves correspond à un ensemble d'informations de même nature (par exemple des études épidémiologiques, ou animales) intégrées pour évaluer une hypothèse. Les quatre étapes de l'évaluation du poids des preuves sont les suivantes :

(1) Planification de l'évaluation

(2) Etablissement des lignes de preuves

(3) Intégration des lignes de preuves

(4) Expression du poids des preuves

La planification de l'évaluation (étape 1) a pour objectif de définir le périmètre de l'évaluation du risque, de décrire les enjeux et d'identifier les méthodes à mettre en œuvre, incluant celles pour l'évaluation du poids des preuves. Elle comporte trois sous-étapes opérationnelles :

- **Cadrage** : définition du contexte de conduite de l'évaluation notamment en termes de ressources et de compétences
- **Formulation de la ou des question(s) à évaluer**
- **Développement** (description) **de la méthode d'évaluation**

L'établissement des lignes de preuve (étape 2) comporte également trois sous-étapes opérationnelles :

- **Recherche et sélection des études (individuelles ou de synthèse)** : la recherche a pour objectif de récupérer l'ensemble des études pertinentes pour la question posée.
- **Évaluation de la qualité des études (individuelles ou de synthèse) :** cette sous-étape vise à évaluer la qualité des études sélectionnées en utilisant une série de critères explicites en fonction des objectifs, comme par exemple le protocole d'échantillonnage, la présence de biais méthodologiques ou le traitement statistique.
- **Analyse d'un ensemble d'études par type d'étude en vue d'établir des lignes de preuve :** il s'agit ici d'intégrer les résultats de plusieurs études d'un même type (ex : études expérimentales *in vitro*, études épidémiologiques, études toxicologiques) dans le but d'estimer une quantité d'intérêt (ex : taux d'incidence d'une pathologie au sein d'une population, valeur de référence) ou de tester une hypothèse (ex : existence d'une relation causale).

L'intégration des lignes de preuves pour établir le poids des preuves (étape 3) a pour objectif de synthétiser les lignes de preuves disponibles dans le but de déterminer le niveau de validité d'hypothèses ou d'estimer des quantités d'intérêt.

L'étape d'**expression des conclusions sur le poids des preuves (étape 4)** permet d'exprimer le poids des preuves sous forme d'un résultat clair et explicite pour l'aide à la prise de décision.

La démarche d'évaluation du poids des preuves est représentée dans la figure ci-dessous :

Figure 1 : Démarche d'évaluation du poids des preuves (étapes et sous-étapes)

Des méthodes ont été recensées pour chaque étape du processus d'évaluation du poids des preuves (Tableau 1).

Tableau 1 : Méthodes recensées pour chaque étape du processus d'évaluation du poids des preuves

		Etablissement des lignes de preuve				
MÉTHODE	**Planification de l'évaluation**	Recherche et sélection des études	Évaluation de la qualité des études	Analyse d'un ensemble d'études par type d'étude en vue d'établir des lignes de preuve	**Intégration des lignes de preuves**	**Expression des conclusions sur le poids des preuves**
(R-)AMSTAR			X			
Analyse multicritères			X	X	X	X
Arbre de décision				X	X	
B. Hill initial				X	X	
B. Hill pondéré				X	X	
CIRC			X	X	X	X
Epid-Tox			X	X	X	X
Évaluation comparative du PDP					X	X
Évaluation du PDP fondée sur les hypothèses					X	X
FDA (2009)		X	X	X		
GRADE	X		X	X		X
Hope et Clarkson	X		X	X	X	X
ILSI 2010			X	X		
Klimisch			X			
Méta-analyse				X		
Méthode bayésienne					X	X
Navigation Guide		X	X	X	X	
NRC (2014)	X					X
OHAT	X	X	X	X	X	X
PRISMA			X			
RS-Cochrane	X	X	X			X
RS-EFSA	X	X				X
SCENIHR			X	X	X	X
STROBE			X			
WCRF			X	X	X	

Pour faciliter l'appropriation des méthodes d'évaluation du poids des preuves par les collectifs d'experts, des fiches méthodes ont été mises à disposition dans le rapport d'étape ainsi que les références bibliographiques clés.

2.2 Analyse d'incertitude

Dans un contexte d'évaluation des risques, le GT MER définit l'analyse d'incertitude comme un processus ayant pour objectif d'identifier, décrire, quantifier et communiquer les incertitudes qui entourent les résultats de l'évaluation (ANSES 2016c).

Pour y répondre, il préconise une démarche qui se veut générique, s'adaptant aux différents champs de compétences de l'Anses en matière d'évaluation des risques (santé-alimentation, santé-environnement ou santé au travail, risques biologiques, chimiques ou physique, contexte règlementaire, décisionnel, *etc*.). La démarche préconisée est structurée en cinq étapes, précédées par une phase de planification (Figure 2). La phase préliminaire de planification a pour but de définir le périmètre de l'analyse de l'incertitude et plus largement de l'évaluation des risques dans laquelle elle s'inscrit, d'une part, de choisir les méthodes et outils appropriés au regard de ce périmètre, d'autre part. Les cinq étapes de l'analyse d'incertitude sont :

(1) d'identifier et de décrire toutes les sources d'incertitude rencontrées lors de l'évaluation, puis de sélectionner celles faisant l'objet des étapes ultérieures,

(2) d'évaluer individuellement les sources d'incertitude retenues,

(3) d'évaluer l'impact combiné des sources d'incertitude retenues sur les résultats de l'évaluation,

(4) de hiérarchiser si nécessaire les sources d'incertitude, par exemple, pour retenir les sources d'incertitude à évaluer de manière individuelle ou pour sélectionner les sources d'incertitude dont l'impact doit être évalué de manière combinée ou pour déterminer les actions à mener pour réduire les incertitudes,

(5) de communiquer les résultats de l'analyse d'incertitude.

La démarche d'analyse d'incertitude est présentée dans la figure ci-dessous :

Figure 2 : Démarche d'analyse d'incertitude

Par ailleurs, le GT MER insiste sur l'importance de proportionner l'effort d'analyse aux besoins de l'évaluation. En effet, aujourd'hui pour réaliser une analyse d'incertitude, une palette de méthodes et d'outils sont disponibles (Maxim 2014, EFSA 2016). Le Tableau 2 en présente les principaux. Les étapes de l'analyse d'incertitude dans lesquelles ces méthodes et ces outils peuvent servir sont également indiquées. Certaines méthodes mettent en œuvre des approches descriptives, d'autres, des approches probabilistes et/ou statistiques qui nécessitent parfois l'utilisation de techniques et d'outils de modélisation et de simulations complexes, demandant des compétences spécifiques. Leurs pertinences dépendent du niveau de complexité de l'évaluation des risques.

Par exemple, une évaluation des risques sommaire ne requière pas, en principe, une analyse d'incertitude quantitative approfondie. Par ailleurs, toutes les incertitudes ne peuvent pas être appréciées de façon identique. Par exemple, les incertitudes liées à des choix ou à des hypothèses (choix de scénarios, choix de modèles, extrapolations, *etc.*) ne peuvent généralement être évaluées qu'à l'aide de méthodes qualitatives. Enfin, les connaissances disponibles n'offrent pas toujours la possibilité d'utiliser des méthodes quantitatives complexes. Pour proportionner l'effort d'analyse, le GT MER recommande de procéder par itération ; une première méthode sera sélectionnée lors de la phase de planification et celle-ci sera si nécessaire adaptée aux besoins et contraintes du collectif.

Tableau 2 : Liste des méthodes pouvant être utilisées, pour chaque étape de l'analyse d'incertitude, selon le type de connaissances traitées

MÉTHODES	Identifier et décrire les sources d'incertitude	Évaluer individuellement les sources d'incertitude retenues	Évaluer l'impact combiné des sources d'incertitude sur le résultat	Hiérarchiser les sources d'incertitude selon leur contribution	Communiquer les résultats de l'analyse d'incertitude
Méthodes utilisées pour traiter exclusivement des dires d'expert					
Expression narrative de l'incertitude	X	X	X	X	X
Echelle ordinale (ANSES 2016d, IPCC 2005)	X	X	X		X
Combinaison d'échelles ordinales (EFSA 2016)		X	X		X
Tableau des incertitudes (Knol et al. 2009)	X	X	X		X
Analyse par intervalles (IPCS 2008)		X	X		
Elicitation de connaissances d'expert (semi-formelle) (EFSA 2014)	X	X	X	X	
Elicitation formalisée de connaissances d'expert (O'Hagan et al. 2006)		X	X		
Méthodes utilisées pour traiter des données issues d'études (expérimentales, épidémiologiques, systèmes de surveillance, campagnes de mesures, *etc.*)					
Calculs déterministes avec hypothèses conservatoires (valeurs par défaut) (EFSA 2007)		X	X		
Inférence statistique - intervalles de confiance (Cox 2006)		X	X		
Inférence statistique – bootstrap (Efron and Tibshirani 1993)		X	X		
Méthodes utilisant les deux types de connaissances (dires d'experts ou données d'études)					
Inférence statistique - inférence bayésienne (Bernier, Parent, and Boreux 2000, Parent and Bernier 2007)		X	X		
Simulations de Monte Carlo simulation (1D-MC et 2D-MC) (Cullen and Frey 1999)			X	X	
Analyse de sensibilité et par scénario (Saltelli, Chan, and Scott 2000, Saltelli et al. 2004)				X	

3 Méthode d'expertise : études de cas et séminaire

Deux sources d'information ont été utilisées pour étudier l'applicabilité des recommandations aux différentes étapes du processus d'évaluation du poids des preuves et d'analyse d'incertitude :

- Les études de cas conduites par des collectifs d'experts, des coordinateurs scientifiques Anses et des membres du GT MER. Ces études de cas ont été choisies pour répondre à un besoin exprimé par les collectifs de l'ANSES pour aider à traiter leurs questions. Le GT MER s'est assuré de la représentativité des domaines, des méthodes, et des étapes, en tenant compte des contraintes de temps et de ressources dans les collectifs. Les enjeux traités dans ces études de cas correspondent soit à ceux des saisines en cours, soit à des enjeux sociétaux importants, soit à un questionnement méthodologique.
- Le séminaire Anses sur les recommandations des deux rapports d'étape du GT MER, organisé en janvier 2017. Ce séminaire comportait plusieurs ateliers participatifs impliquant de nombreux experts et scientifiques Anses. A cette occasion, les recommandations ont été illustrées par des exemples issus des études de cas ou de travaux antérieurs Anses, et présentés sous forme de posters.

3.1 Organisation des études de cas

3.1.1 Etudes de cas sur l'évaluation du poids des preuves

Sept études de cas ont permis de tester les recommandations du rapport sur l'évaluation du poids des preuves dans deux domaines de l'Anses :

Domaine Santé travail – Direction de l'évaluation des risques :

- PDP1. Grille d'évaluation des études épidémiologiques du CES Valeur Limite d'Exposition Professionnelle (VLEP) (analyse des risques de biais par la méthode OHAT) – Unité d'évaluation des risques liés aux substances chimiques (UESC)/CES VLEP

Domaine Santé environnement – Direction de l'évaluation des risques :

- PDP2. Planification de la méta-analyse chez les sujets électrohypersensibles (EHS) - Unité d'évaluation des risques liés aux agents physiques (UERAP) /GT Radiofréquences
- PDP3. Sensibilité des conclusions à la méthode qualitative pour combiner les lignes de preuves (arbre de décision Bisphénol A / méthode CIRC adapté) – UERAP/CES Agents physiques
- PDP4. Evaluation de la revue de synthèse par R-AMSTAR – Unité d'évaluation des risques liés à l'air (UERA) /GT Particules
- PDP5. Etablissement et intégration des lignes de preuve – UERA/GT Particules
- PDP6. Evaluation de la qualité des revues et des publications ultérieures (méthodes R-AMSTAR et OHAT) – UERA/GT Silice cristalline
- PDP7. Etablissement d'une grille de cotation pour les effets perturbateurs endocriniens (intégration des données épidémiologiques, expérimentales, *in silico* et de mode d'action) – UESC/GT Perturbateurs endocriniens

Les études de cas ont été croisées avec l'ensemble des étapes du processus d'évaluation du poids des preuves pour vérifier leur représentativité (Tableau 3). Les études de cas (PDP1,PDP2,

PDP4, PDP6) sont liées à la mise en œuvre d'une méthode spécifique (OHAT, R-AMSTAR) tandis que d'autres s'intéressent à une ou plusieurs étapes de la démarche.

Tableau 3 : **Etudes de cas par étape de l'évaluation du poids des preuves**

Etapes de l'évaluation du poids des preuves (PDP)	1 Planification de l'évaluation			2 Etablissement des lignes de preuves			3 Intégration	4 Expression
	Cadrage	Formulation de la/des question(s)	Dévelop-pement de la méthode d'évaluation	Recherche, sélection des études	Evaluation de la qualité des études	Analyse d'un ensemble d'études par type d'étude		
Titre de l'étude de cas								
PDP1 : grilles des études épidémiologiques CES VLEP					X OHAT			
PDP2 : Méta-analyse sur EHS	X	PECOTS	X	(X)				
PDP3 : sensibilité des conclusions à la méthode qualitative pour combiner les lignes de preuves (BPA/CIRC)						X	X	X
PDP4 : Evaluation de la revue de référence par R-Amstar GT Particules		PECOTS			x R-AMSTAR			
PDP5 : Etablissement et intégration des lignes de preuves GT Particules		PECOTS			OHAT	OHAT	(X)	(x)
PDP6 : Evaluation de la qualité des revues et des publications ultérieures GT Silice Cristalline		PECO			R-AMSTAR OHAT	(X)	(X)	
PDP7 : Etablissement d'une grille de cotation pour les effets perturbateurs endocriniens GT PE							X	

Toutes les études de cas ont déjà produit des résultats, présentés dans le chapitre suivant. Trois d'entres elles (PDP2, PDP5, PDP7) ne sont pas terminées (x)..

3.1.2 Etudes de cas sur l'analyse d'incertitude

Quatre études de cas ont permis de tester les recommandations du rapport sur l'analyse d'incertitude dans quatre domaines de l'Anses :

Domaine Santé animale – Direction de l'évaluation des risques :

- AI1. Prise en compte de l'incertitude dans la saisine Danger *Salmonella* en alimentation animale – Unité d'évaluation des risques liés à la santé, à l'alimentation et au bien-être des animaux (UERSABA) /GT Salmonella (SALAN)

Domaine Santé des végétaux – Direction de l'évaluation des risques/Laboratoire de santé des végétaux :

- AI2. Amélioration de la prise en compte de l'incertitude dans le processus d'analyse de risque phytosanitaire – Application à la saisine *Ambrosia trifida* (méthode d'élicitation) – Unité expertise des risques biologiques/GT Ambroisie

Domaines Produits phytopharmaceutiques, biocides et fertilisants – Direction des produits réglementés

- AI3. Analyse des incertitudes dans le cadre de l'évaluation de dossiers de produits réglementés - Unité évaluation toxicologie des produits réglementés /CES Biocides

Domaine Santé environnement – Direction de l'évaluation des risques :

- AI4. Elaboration d'un tableau des incertitudes – Retour d'expérience de la saisine Jouets UERA/UESC/CES Evaluation des risques chimiques liés aux articles et produits de consommation

Le tableau 4 décrit les étapes de l'analyse d'incertitude couvertes par les études de cas. Comme pour le poids des preuves, certaines études de cas sont liées à la mise en œuvre d'une

recommandation spécifique (ex : utilisation d'un tableau des incertitudes) tandis que d'autres s'intéressent à une ou plusieurs étapes de la démarche.

Tableau 4 : Etudes de cas par étape de l'analyse d'incertitude

Etapes de l'analyse d'incertitude (AI)	Planification de l'évaluation							
	Cadrage	Formulation de la/des question(s)	Dévelop-pement de la méthode d'évaluation	1- Identification et description des sources d'incertitude	2- Evaluation individuelle des sources d'incertitude retenues	3- Evaluation combinée de l'impact des sources d'incertitude	4 Hiérarchisation des sources d'incertitude	5- Communication des incertitudes
Titre de l'étude de cas								
AI1 : Danger *Salmonella* GT SALAN	X	X	X	(X)	(X)	(X)	(X)	(X)
AI2 : Amélioration de la prise en compte de l'incertitude dans l'ARP - GT Ambroisie				X	X			
AI3 : Analyse des incertitudes dans le cadre de l'évaluation de dossiers de produits réglementés DEPR /CES Biocides				X	X	X		
AI4 : Tableau des incertitudes - retour d'expérience de la saisine jouets				X				

Toutes les études de cas ont produit des résultats, présentés dans le chapitre suivant. L'étude de cas AI1 n'est pas encore terminée (X). Des travaux Anses concernant les étapes 3, 4, 5 réalisés en dehors des études de cas ont également été mobilisés pour illustrer certaines recommandations.

3.2 Organisation du séminaire

Un séminaire d'une journée a été organisé par l'Anses. Il comportait une séance plénière de présentation des recommandations, une session de discussion autour de posters et des ateliers. Cent quarante experts et scientifiques Anses ont participé à cette journée ainsi qu'aux ateliers conduits avec huit facilitateurs.

3.2.1 Des ateliers participatifs

Les ateliers ont été organisés autour des six thématiques suivantes :

- Thème 1 - Planification de l'évaluation concernant le poids des preuves et l'incertitude
- Thème 2 - Établissement des lignes de preuves
- Thème 3 - Identification et description des incertitudes
- Thème 4 - Intégration des lignes de preuve pour établir le poids des preuves
- Thème 5 - Evaluation et hiérarchisation des incertitudes
- Thème 6 - Expression des conclusions concernant le poids des preuves et l'incertitude

Pour chaque thématique, les recommandations du GT MER ont été présentées en plénière, et les études de cas ou des travaux antérieurs de l'Anses ont été exposées sous forme de posters pour illustrer ces recommandations (Annexe 1).

L'animation des ateliers a été confiée à des facilitateurs ayant suivi une formation d'intelligence collective de l'université de Cergy-Pontoise. Les principes fondamentaux de l'intelligence collective ont été appliqués : bienveillance, équilibre du temps de parole, écoute et disposition circulaire du groupe.

Dans chaque atelier constitué de 14 à 23 participants, des questions de clarification ont été posées après une présentation des recommandations par l'équipe référente du GT MER. Les demandes principales de clarification par thématique sont regroupées en annexe 2.

Chaque participant a ensuite pu s'exprimer sur la mise en application des recommandations : *j'appliquerai cette recommandation, car…* ; *j'appliquerai cette recommandation, si…* ; *je ne pourrais pas appliquer cette recommandation, car…* Ils ont justifié les raisons de leur choix.

3.2.2 Des posters pour illustrer les recommandations

Quinze posters ont illustré les recommandations du GT MER lors du séminaire. Les posters se répartissent dans les différents domaines de l'Anses, prioritairement en Santé environnement (Tableau 5).

Tableau 5 : liste des posters illustrant les recommandations par domaine de l'Anses

1 en Santé travail – Direction de l'évaluation des risques : Poster **4** (PDP1): Grille d'évaluation de la qualité des études épidémiologiques
4 en Santé alimentation – Direction de l'évaluation des risques Poster **3** : Revue systématique de la contamination du poisson par des microcystines Poster **13** : Prise en compte de l'évaluation de l'impact combiné des incertitudes dans l'évaluation du risque lié à Vibrio parahaemolyticus lors de la consommation de coquillages Poster **14** : Méthodes probabilistes pour évaluer l'impact combiné de l'incertitude lié au risque de réaction allergique à l'arachide Poster **15** : Analyse de sensibilité dans le cadre d'une appréciation quantitative des risques microbiologiques
2 en Santé animale – Direction de l'évaluation des risques Poster **1** (AI1): Schéma conceptuel relatif au processus d'introduction des salmonelles dans les filières de production des denrées alimentaires Poster **11** : Elicitation formalisée des connaissances d'experts pour l'évaluation individuelle des incertitudes avec la méthode Sheffield
6 en Santé environnement – Direction de l'évaluation des risques : Poster **2** (PDP2): Protocole de la méta-analyse des études de provocation aux champs électromagnétiques chez des sujets se déclarant électrohypersensibles Poster **5** (PDP6): Evaluation de la qualité des revues par la méthode R-Amstar Poster **6** (PDP4 & PDP5): Etablissement des lignes de preuve avec la méthode OHAT pour les effets sanitaires des particules de l'air ambiant Poster **8** (AI4): Elaboration d'une matrice d'incertitude – retour d'expérience sur la saisine « jouets » Poster **9** (PDP7): Etablissement d'une grille de cotation pour les effets perturbateurs endocriniens Poster **10** (PDP3): Sensibilité des conclusions à la méthode qualitative d'évaluation du poids des preuves
1 en Santé des végétaux – Direction de l'évaluation des risques/Laboratoire de santé des végétaux : Poster **12** (AI2): Elicitation des connaissances d'experts pour l'évaluation individuelle des incertitudes dans le cadre de l'analyse du risque phytosanitaire
1 pour les produits reglementés (phytopharmaceutiques, biocides et fertilisants) Poster **7** (AI3): Analyse des incertitudes dans le cadre de l'évaluation de dossiers de produits réglementés

3.2.3 Synthèse des méthodes mobilisées dans les études de cas et les posters

Les tableaux 6 et 7 synthétisent les méthodes d'évaluation du poids des preuves et d'analyse d'incertitude mises en oeuvre dans les études de cas et dans les posters. Ces méthodes ont mobilisé des ressources et parfois des compétences spécifiques décrites dans ces tableaux. Les contraintes en termes de délais sont également explicitées. Ces informations peuvent aider l'Anses à faire des arbitrages méthodologiques lors de l'étape de planification des futures saisines.

Tableau 6 : Applications de différentes méthodes relatives à l'évaluation du poids des preuves à l'Anses

Recommandations	Méthodes testées	Objectif de l'application	Modalités du choix de la méthode	Temps imparti	Ressources humaines	Moyens spécifiques (logiciel,…)	**Domaine**	Référence **Poster**
Planifier l'évaluation du poids des preuves Développer une méthode d'évaluation Etude de cas **PDP 2**	Méta-analyse	Rédiger le protocole de la méta-analyse des études de provocation aux champs électromagnétiques chez des sujets se déclarant électrohypersensibles	La démarche proposée pourrait permettre d'évaluer le niveau de preuve à partir d'études de provocation et, ainsi, de synthétiser des résultats apparemment discordants.	4 réunions	4 experts du GT Radiofréquences, 4 experts du GT MER, Anses : 2 agents de l'unité méthodologie et étude, 2 scientifiques de l'unité d'évaluation	Grille d'extraction des données pour l'étude de faisabilité	Santé environnement	**Poster 2** (Fite et al. 2017)
Rechercher et sélectionner les études	Revue systématique	Revue exhaustive des données de la littérature, méthode robuste et reproductible, analyse critique de la qualité des résultats	Méthode choisie en interne à l'unité. Choix validé par le CES.	6 mois pour la revue systématique	1 agent Anses + 2 experts de CES	Endnote pour la gestion des références bibliographiques et pour l'historique selon les étapes de sélection des articles. Le logiciel DISTILLER (utilisé par l'Efsa) aurait été utile.	Santé alimentation – Risque chimique	Rapport de l'Anses de juin 2016 (saisine n°2015-SA-0206). **Poster 3** (Arnich 2017)
Rechercher et sélectionner les études Evaluer la qualité des études **Etude de cas PDP 1**	OHAT	Définir les critères d'évaluation des études Analyser les risques de biais des études	GRADE trop complexe OHAT permet un premier tri de la qualité des études avec élimination (du palier 3)	2 heures par étude pour évaluer une publication avec la grille descriptive intégrant l'analyse des risques de biais	2 évaluateurs si possible de différentes spécialités GT mis en place pour l'étude de cas avec 7 experts du CES VLEP impliqués		Santé travail	Saisine 2015-SA-0140 **Poster 4** (Sissoko et al. 2017)
Evaluer la qualité des revues **Etude de cas PDP 6**	R-Amstar	Classer les revues en fonction de leur niveau de qualité et éliminer les revues de mauvaise qualité	Evaluer plusieurs dizaines de revues de synthèse (épidémiologiques, toxicologiques)	15 mn à 1 heure pour remplir la grille Amstar (en fonction si critère rédhibitoire) 3 demi-journées de GT pour tester et adapter la grille + 4 demi-journées pour évaluer 90 revues (passage en réunion si cotation différente après double lecture)	2 coordinateurs Experts du GT Silice	Grille R-Amstar ad 'hoc définie par le GT Silice	Santé environnement	**Poster 5** (Paillat, Fourneau, and Silice cristalline 2017)

Evaluer la qualité des études **Etudes de cas PDP 5**	OHAT	Analyse des risques de biais	Structuration et harmonisation des critères d'évaluation de la qualité des études	1 heure minimum par étude et relecteur → durée : 10 mois	2 experts 1 coordinateur	Développement d'outils Excell et d'un aide mémoire	Santé environnement	Saisine 2014-SA-0156 **Poster 6** (Redaelli et al. 2017)
	OHAT	Regrouper les études d'une ligne de preuve (dans un premier temps par niveau de confiance initial = sous-ligne de preuve).	Méthode directive	Très rapide après un premier temps d'appropriation : 30 minutes pour 1ère sous-ligne 10 minutes pour 2è sous-ligne 5 minutes pour les 2 suivantes	Nécessité d'expertise collective pour conclure sur la confiance finale d'une ligne de preuve	Développement d'outils Excell et définition de modalités communes de notation pour chaque facteur augmentant ou diminuant la confiance	Santé environnement	Saisine 2014-SA-0156 **Poster 6** (Redaelli et al. 2017)
Utiliser des modèles statistiques pour intégrer différentes lignes de preuve **Etude de cas PDP 7**	développement d'une méthodologie ad-hoc pour la combinaison quantitative des lignes de preuves (avec MOA et mise en œuvre d'une élicitation d'experts pour la cotation des lignes de preuves)	intégration des lignes de preuves (données épidémiologiques, expérimentales, *in silico*) pour identifier les substances susceptibles d'être des perturbateurs endocriniens (PE)	Cette méthodologie permet d'associer des scores de vraisemblance aux jugements d'experts. Elle permet de formaliser et d'expliciter les critères pris en compte par les experts.	Consolidation, phase de test : 1 an	compétences requises : modélisateur en (éco)toxicologie, (éco)toxicologue 1 coordinateur		Santé environnement	**Poster 9** (Beausoleil et al. 2017)
Décrire et expliciter les choix pour l'intégration des lignes de preuves selon une méthode qualitative **Etude de cas PDP 3**	Etablissement des lignes de preuves développé pour l'expertise BPA Etablissement des lignes de preuves développé par le GT RF	Evaluer la sensibilité des conclusions d'une expertise à la méthode d'établissement et d'intégration des lignes de preuve	Demande de comparaison de deux méthodes par les parties prenantes	7 jours	1 coordinateur	non	Santé environnement	**Poster 10** (Roth-Delgado et al. 2017)

Tableau 7 : Applications de différentes méthodes relatives à l'analyse d'incertitude à l'Anses

Recommandations	Méthodes testées	Objectif de l'application	Modalités du choix de la méthode	Temps imparti	Ressources humaines	Moyens spécifiques (logiciel,…)	**Domaine**	Référence (poster)
Planifier l'analyse d'incertitude **Étude de cas AI 1**	Établissement du schéma conceptuel	Planifier l'analyse des incertitudes relatives à la saisine SALAN	Choix d'un schéma conceptuel global, puis d'un schéma par question.	Phase d'analyse : clarification des questions posées, identification des principaux éléments et analyse de leurs liens : 1,5 jours répartis sur les 3-4 premières réunions (en séance plénière) Phase de formalisation : construction du schéma conceptuel (figure et description) : 1 jour Phase de validation du schéma conceptuel : 0,5 jour réparti sur 2 réunions (en séance plénière)	Phase d'analyse : l'ensemble du GT SALAN Phase de formalisation : 1 coordinatrice scientifique (unité d'expertise) + 1 expert du GT MER Phase de validation : l'ensemble du GT SALAN		Santé et bien être des animaux	**Poster 1** (Khamisse et al. 2017)
Identifier et décrire l'ensemble des incertitudes Evaluer les sources d'incertitude retenues **Etude de cas AI 3**	Grille d'évaluation des incertitudes et Tableau de présentation des incertitudes	Identification et évaluation qualitative des incertitudes	Choix d'une méthode qualitative du fait des contraintes de temps et ressources de l'évaluation des risques dans le cadre d'une demande de mise sur le marché d'un produit biocide.	Grille d'évaluation - Construction de l'outil : 5-10 jours - Utilisation de la grille d'évaluation 30 minutes pour un utilisateur expérimenté. Tableau de présentation : 30 minutes avec l'appui de la grille d'évaluation.	Grille d'évaluation • Développement : 2 agents • Utilisation : 1 agent Tableau de présentation : Utilisation : 1 agent	Grille sous format word	Produits reglementés - Biocides	**Poster 7** (De Saint-Jores et al. 2017)
Présenter le recensement des incertitudes dans un tableau **Etude de cas AI 4**	Typologie et descripteurs	Tester une proposition de tableau de recensement constituée de la typologie (lignes) et de descripteurs (colonnes).	Intérêt de l'approche : standardisation, évite des oublis	La durée de réalisation est compatible avec l'instruction normale d'une saisine (<5% en plus).	Pas de compétences supplémentaires nécessaires		Santé environnement	**Poster 8** (Bonvallot et al. 2017)
Evaluer individuellement les sources d'incertitude	Méthode Sheffield Quartile	Estimation de sensibilité et de spécificité d'indicateurs de	Choix méthode Sheffield : transparence et simplicité	Entre 6 mois et 1 an (pas en temps plein)	10 experts/filière 1 élicitateur 1 groupe pilote de 10	Développement d'un outil spécifique sur R	Bien-être animal	Rapport saisine 2015-SA-0087 **Poster 11** (Thebault, Chiron, et al.

		conscience	Choix méthode quartile : 20 distributions par filière animale	selon les élicitations et expérience acquise	personnes 2 coordinateurs scientifiques	shiny		2017)
Evaluer individuellement les sources d'incertitude **Etude de cas AI 2**	Elicitation des connaissances d'expert	Evaluer les incertitudes à l'aide d'une distribution de probabilité	Choix d'une méthode d'élicitation formelle pour gagner en transparence et en reproductibilité	Préparation des documents supports pour l'estimation d'une quantité : 1 jour Présentation de la démarche d'élicitation et mise en œuvre de l'éliciation pour la dite quantité : 1 réunion	Préparation des documents supports pour l'estimation d'une quantité : 1expert Présentation de la démarche d'élicitation : 1expert Mise en œuvre de l'éliciation pour la dite quantité : 3 experts	logiciel « The MATCH Uncertainty Elicitation Tool »	Santé des végétaux	**Poster 12** (Tassus et al. 2017)
Evaluer l'impact combiné des sources d'incertitude	Méthode séparant incertitude et variabilité : méthode simulation monte-carlo à deux dimensions	Estimation du nombre de cas humain de Vibrioses liés à la consommation de coquillages	Des sources d'incertitudes et de variabilité quantifiables : appproche quantitative probabiliste (méthode simulation monte-carlo à deux dimensions)	6 mois (pas en temps plein)	1 groupe de travail (5 personnes) avec 2 modélisateurs, 1 coordinateur scientifique, et 2 experts des données	Développement d'un programme spécifique sous R	Santé alimentation - Risque microbiologique	Rapport Saisine 2010-SA-0301 **Poster 13** (Thebault, Guillier, et al. 2017)
Evaluer l'impact combiné des sources d'incertitude	Méthode séparant incertitude et variabilité : méthode simulation monte-carlo à deux dimensions	Evaluation du risque de réaction allergique	Prise en compte de différentes sources de variabilité et d'incertitude et propagation séparée	6 mois	Equipe projet et 1 modélisateur	Logiciel R	Santé alimentation	**Poster 14** (Crépet 2017)
Hiérarchiser les incertitudes	Sobol	Hiérarchiser les sources d'incertitude	Méthode adaptée et efficace mais ne peut etre appliquée que sur un nombre restreint de paramètres en raison des délais de mise en œuvre	Réalisé dans le cadre des travaux de thèse de doctorat Pour le développement du modèle à 2d + analyse de sensibilité : 1 an.	1 agent	Logiciel de programmation et de simulation (SAS)	Evaluation des risques microbiologiques	**Poster 15** (Audiat-Perrin et al. 2017) Perrin, Frédérique. *Appréciation quantitative des risques liés à deux pathogènes d'intérêt pour la filière laitière: les Escherichia coli producteurs de shiga-toxines et Salmonella spp.* Thèse de doctorat: Sécurité sanitaire des aliments et risques microbiologiques. Ecole Doctorale ABIES. Maison du Lait, Paris. 2013

4 Mise à jour des recommandations du GT MER pour l'évaluation du poids des preuves et l'analyse d'incertitude

Pour mettre en œuvre l'évaluation du poids des preuves et l'analyse d'incertitude dans les expertises menées par l'Agence, le GT MER préconise de suivre les démarches décrites dans le chapitre 2 de ce rapport en se basant sur les recommandations présentées ci-dessous. Ces recommandations ont été actualisées à partir des résultats des études de cas et du séminaire (cf chapitre 3).

La majorité des participants du séminaire organisé en janvier 2017 considèrent que les recommandations seraient applicables (cf. Annexe 3) sous certaines conditions détaillées ci-dessous. Toutes les évaluations des risques ne nécessitent cependant pas le même niveau de détail. Le GT MER recommande de proportionner l'effort d'analyse aux besoins du contexte décisionnel.

Le chapitre 4.1 concerne la planification de l'évaluation qui est une étape commune à l'évaluation du poids des preuves et à l'analyse d'incertitude. Les chapitres 4.2 et 4.3 traitent des étapes « Etablissement des lignes de preuves » et « intégration des lignes de preuve » décrites sur la Figure 1. Les chapitres 4.4 à 4.7 portent sur les différentes étapes de l'analyse d'incertitude présentées sur la Figure 2. Enfin, les chapitres 4.8 et 4.9 concernent l'expression des conclusions et l'harmonisation des pratiques de l'Anses.

Chaque chapitre présente les recommandations initiales (entre « »), puis des éléments de clarification avec des exemples d'application, et enfin les recommandations mises à jour (dans un encadré avec en gras les modifications).

4.1 Planification de l'évaluation

4.1.1 Recommandations initiales

Recommandations du rapport sur l'évaluation du poids des preuves :

« Le GT MER recommande la mise en place d'un processus complet pour la planification de l'évaluation en 3 sous-étapes avec pour chacune un formulaire approprié. La sous-étape 1 (cadrage) est réalisée par les scientifiques de l'Anses. Les deux sous-étapes suivantes mobilisent les CES et les experts impliqués dans le traitement de la saisine.

Le GT MER recommande d'adopter une démarche rétroactive entre les sous-étapes 2 (formulation des questions) et 3 (choix de la méthode d'évaluation) de manière à affiner la formulation des questions en garantissant que celles-ci puissent être correctement traitées.

Le GT MER recommande **à l'aide d'une revue préliminaire et si besoin d'auditions** :

- d'identifier les enjeux (sanitaires, environnementaux, sociétaux, économiques avec les décideurs (et, dans l'idéal, les parties prenantes),
- d'évaluer l'étendue du corpus des données disponibles,
- **De formaliser la ou les question(s) au moyen d'une structure de description présentée dans la littérature (PICO, PECO, etc.)**,
- de formaliser les sous-questions à traiter par le collectif d'experts, si possible sous forme de **modèle conceptuel**,
- de **rédiger un plan de développement de la méthode d'évaluation** en préalable à sa conduite,
- de définir les modalités de communication de la conclusion dans le document du cadrage.

Dans le cadre de l'évaluation du poids des preuves, le GT MER recommande, pour la sous-étape 3, de spécifier au minimum :

- **le choix du type de revue de la littérature** : revue systématique ou revue approfondie pour traiter la ou les questions définies lors de la sous-étape 2. Ce choix doit tenir compte des éléments issus du document de cadrage, de l'analyse socio-économique de l'Anses et des critères proposés par l'EFSA : impact potentiel du résultat de la revue systématique, quantité et qualité des données, source et confidentialité des données, transparence nécessaire et controverse du sujet de la saisine, ressources à mettre en œuvre,
- les critères utilisés pour évaluer la qualité des études par type d'étude,
- la méthode retenue pour évaluer un ensemble d'études : méta-analyse, approche multicritères, ou approches qualitatives,
- la méthode retenue pour intégrer les lignes de preuves : modélisation statistique avec mise en place d'élicitation ou non de dires d'experts, analyse multicritères, ou approches qualitatives,
- les modalités d'expression des conclusions concernant le poids des preuves. »

Recommandations du rapport sur l'analyse d'incertitude :

« Les guides méthodologiques faisant actuellement référence dans le domaine de l'évaluation des risques préconisent, pour toute évaluation des risques, la mise en place d'une phase préliminaire de planification structurée et transparente permettant de définir le contexte décisionnel de l'évaluation et les ressources nécessaires, d'apprécier la ou les questions faisant objet de l'évaluation et de définir la méthode d'évaluation à mettre en œuvre. Ces guides préconisent également d'y inscrire la planification de l'analyse d'incertitude, en vue d'en maximiser l'utilité pour la gestion des risques.

Le GT MER recommande d'améliorer la planification existante, de la formaliser en trois étapes et d'inclure dans chacune d'elle les éléments de l'analyse d'incertitude correspondante :

- « Cadrage » : rédaction d'un document de cadrage de l'expertise précisant le contexte décisionnel et les ressources mobilisées *a priori.* En particulier, expliciter de manière transparente les éléments susceptibles de guider le choix de la méthode d'évaluation et d'analyse d'incertitude.
- « Formulation de la question » : développement d'un schéma conceptuel qui comprend une description graphique et une description narrative des principaux facteurs à prendre en compte dans l'évaluation ainsi que les incertitudes associées. Le schéma conceptuel devra également s'intéresser aux éléments qui pourraient ne pas être pris en compte dans l'évaluation mais essentiels à l'analyse d'incertitude.
- « Choix de la méthode d'évaluation » : rédaction d'un plan d'analyse (ou d'évaluation) décrivant les données, les informations, les méthodes et les modèles à utiliser pour l'évaluation des risques et l'analyse d'incertitude associée. L'expression des résultats et leurs communications – pour l'évaluation des risques comme pour l'analyse d'incertitude – doivent également être précisées. »

4.1.2 Clarification et exemples d'application

La planification est un processus iteratif en trois sous-etapes, « Cadrage », « Formulation de la question » et « Choix de la méthode d'évaluation ».

Le formulaire « Analyse de la saisine et cadrage interne de l'expertise » constitue un support au cadrage et à la formulation de la question de l'étape de planification de l'évaluation. Il permet notamment de définir le contexte sociétal, règlementaire et international, les questions scientifiques posées, ainsi que les compétences, les moyens et ressources nécessaires à leur traitement, les modalités d'organisation de l'expertise, la planification prévisionnelle et les modalités de communication. En complément, le formulaire « Grille de questionnement sociologique » permet d'aborder le contexte socio-économique, les pratiques des différents groupes sociaux, la construction du problème, les formes de savoir et les inégalités. Elle complète le cadrage.

Les sous-étapes 2 et 3 permettent de définir les modalités de mise en œuvre et les contraintes de faisabilité de l'évaluation. En l'absence de données permettant de répondre à la question posée, il peut être opportun de redéfinir le cadrage en lien avec les parties prenantes (sous-étape 1), voire de réviser le « contrat d'expertise ».

Lors de la sous-étape 2, la structure d'information (PICO-PECO-PECOTS) permet de préciser la population et les effets sanitaires d'intérêt, l'exposition à considérer, sa durée, sa localisation, les comparateurs evalués. La figure 3 présente un exemple de PECOTS appliqué à l'évaluation des effets sanitaires des particules de l'air ambiant (PDP5, (Redaelli et al. 2017)). Cette structure leur a permis de définir les contours de la revue de la littérature.

Populations	Population générale
Expositions	Catégories de composés et de sources des particules de l'air ambiant extérieur
Comparateurs	Absence d'exposition ou exposition plus faible
Evènements néfastes pour la santé (Outcomes)	Catégories d'évènements néfastes pour la santé associés à l'exposition aux particules de l'air ambiant extérieur
Durées d'exposition (Timings)	Expositions à court terme (un à plusieurs jours) et long terme (une à plusieurs années)
Localisation d'intérêt (Setting of interest)	France (mais inclusion des études quelle que soit leur localisation géographique)

EXPOSITIONS
SOURCES DES PARTICULES
- Trafic (carbone suie)
- Trafic (PM2,5)
- Diesel
- Nouveau Diesel[1]
- Essence
- Poussière de freinage
- Poussière de route / Matière crustale
- Combustion de charbon
- Combustion de pétrole
- Emission centrale électrique à charbon
- Industrie
- Combustion biomasse
- Poussière de désert
- Mer et océan
- Site de déchet dangereux
- Agriculture / Site rural

[1] inclut les véhicules diesel équipés de filtres à particules de type «wall-flow » (technologie introduite depuis 2000 en France)

OUTCOMES
- Décès toutes causes[1]
- Maladies cardiovasculaires
- Maladies respiratoires[2]
- Cancers broncho-pulmonaires
- Cancers autres
- Maladies neurologiques
- Troubles de la fertilité[3]
- Troubles périnataux[4]
- Diabète/Troubles du métabolisme

[1] Décès toutes causes hors causes traumatiques
[2] Inclut l'asthme chez l'enfant exposé *in utero* et les infections respiratoires infantiles
[3] Troubles de la préconception jusqu'à 20 semaines de grossesse chez l'Homme
[4] Troubles de 20 semaines de grossesse jusqu'au premier mois de vie après la naissance chez l'Homme

EXPOSITIONS
COMPOSÉS DES PARTICULES
- Carbone suie / carbone élémentaire
- Carbone organique
- Hydrocarbures aromatiques polycycliques
- Particules ultrafines
- Particules grossières
- CAP (concentrated ambient particles)
- Particules ultrafines de carbone
- Aérosols organiques secondaires
- Aérosols inorganiques secondaires
- NH_4^+
- NO_3^- / N
- $SO4^{2-}$ / S
- Ni
- Zn
- Cu
- V
- Si / SiO_2
- Fe
- Autres oligo-éléments
- Endotoxines
- Potentiel oxydant

Figure 3 *:* PECOTS évaluation des effets sanitaires des particules de l'air ambiant (Etude de cas PDP5)

Cette structure doit être appliquée pour chaque sous-question identifiée dans le schéma conceptuel.

Dans le cadre de la sous étape 2 « formulation de la question », le GT MER insiste sur la nécessité d'établir un schéma conceptuel pour permettre de bien comprendre l'ensemble de la question. Un schéma conceptuel est une représentation graphique accompagnée d'un texte explicatif, comportant les objets de la question posée, à savoir les agents d'intérêt, les populations exposées, les effets sanitaires et les relations entre ces différents éléments. Le schéma conceptuel peut intégrer également d'autres déterminants comme les facteurs sociaux-économiques.

Le schéma conceptuel peut permettre d'identifier les points critiques du système qui serviront ensuite de base pour définir les sous-questions de l'évaluation.

A titre d'exemple, dans l'étude de cas AI1 (Khamisse et al. 2017), un schéma conceptuel a été développé en deux temps : un premier schéma (figure 4) présente les principaux éléments du processus d'entrée potentielle de *Salmonella* via l'alimentation animale, correspondant à une vision globale du système étudié. Deux autres schémas détaillent chacun un sous-système du système global. L'usage d'éléments visuels caractéristiques (formes, couleurs, images) permet de différencier les séquences temporelles (Exemple figure 5).

Figure 4 : principaux éléments du processus d'entrée potentielle de *Salmonella* (étude de cas AI1)

Figure 5 : Schéma conceptuel du maillon portuaire (étude de cas AI 1)

Le développement de la méthode d'évaluation a été testé dans le cadre d'une étude de cas PDP2 relative aux travaux du GT « Radiofréquences et santé » sur des sujets se déclarant électrohypersensibles (EHS) aux champs électromagnétiques, afin d'élaborer un protocole pour la mise en œuvre d'une méta-analyse à l'Anses (Fite et al. 2017). L'objectif de la démarche est d'évaluer le niveau de preuve à partir d'études de provocation et, ainsi, de synthétiser des résultats apparemment discordants. Un protocole de méta-analyse a été défini. Il inclut les étapes suivantes : formulation de la question traitée dans la méta-analyse avec la structure d'information PECOTS, sélection des études, d'extraction des données et d'analyse statistique (Tableau 8). Plusieurs points critiques nécessitant des compétences spécifiques ont été identifiés : la réalisation d'une revue exhaustive de la littérature, l'évaluation de la qualité des articles, l'analyse des données extraites des articles, *etc*. Ainsi la réalisation d'une méta-analyse doit être précédée d'une étude de faisabilité. En outre, cette étude de cas a souligné la nécessité d'un temps d'appropriation de la méthode par les scientifiques et les experts de l'Anses.

Tableau 8 : Exemple de protocole de méta-analyse (étude de cas PDP2)

1.	Contexte	
2.	Question posée : structure d'information de type PECOTS	
3.	Méthode	
	3.1.	Revue systématique (restitution avec le diagramme PRISMA) *Table 1. Définition de la requête (cf. profil de recherche bibliographique Anses)* *Table 2. Critères d'éligibilité et d'exclusion des études* *Table 3. Grille de lecture* *Table 4. Grille d'extraction des données*
	3.2.	Analyse quantitative
	3.2.1.	Choix d'un logiciel pour estimer les paramètres et réaliser les tests statistiques
	3.2.2.	Choix de la méthode de calcul pour étudier l'association entre l'exposition et l'effet
	3.2.3.	Définition des modèles statistiques
	3.2.4.	Analyse de sensibilité des conclusions au modèle statistique sélectionné
	3.2.5.	Seuil de significativité statistique pour le test de l'effet
	3.2.6.	Recherche de biais de publication
	3.2.7.	Evaluation de la robustesse des résultats
4.	Ressources nécessaires	

En définitive, les recommandations du GT MER sont applicables dans différents domaines de compétences de l'Anses (santé animale, santé environnement, santé végétale) et l'accompagnement d'au moins un expert du GT MER a été apprécié dans les différents cas (GT SALAN, GT Particules, GT Radiofréquences).

4.1.3 Recommandations mises à jour

Que ce soit dans le cadre d'une évaluation du poids des preuves ou dans celui d'une analyse d'incertitude, le GT MER recommande la mise en place d'un processus complet pour la planification de l'évaluation des risques en 3 sous-étapes:

1. **Cadrage :** rédaction d'un document de cadrage de l'expertise précisant les enjeux, le contexte décisionnel et les ressources mobilisées a priori. En particulier, expliciter de manière transparente les éléments susceptibles de guider les choix méthodologiques.
2. **Formulation de la ou des question(s),** en se basant sur un schéma conceptuel qui comprend une description graphique et une description narrative des principaux facteurs à prendre en compte.
3. **Choix et description de la méthode d'évaluation :** rédaction d'un plan décrivant les données, les informations, les méthodes et les modèles à utiliser. L'expression des résultats et modalités de communication doivent également être précisées à ce stade.

Le GT MER recommande d'adopter une démarche rétroactive entre les trois sous-étapes pour pouvoir éventuellement modifier le cadrage (sous-étape 1) et la formulation des questions (sous-étape 2) de manière à s'assurer que les questions puissent être correctement traitées à l'issue de la sous-étape 3.

Dans le cadre d'une évaluation du poids des preuves :

Le GT MER recommande, pour la sous-étape 2, de formaliser la ou les question(s) au moyen d'une structure de description présentée dans la littérature (PICO, PECO, etc.) ;

Le GT MER recommande, pour la sous-étape 3, de spécifier au minimum (à partir d'une revue préliminaire de la littérature et si besoin d'auditions) :

- Le choix du type de revue de la littérature : revue systématique ou revue approfondie pour traiter la ou les questions définies lors de la sous-étape 2. Ce choix doit tenir compte des éléments issus du cadrage, de l'analyse socio-économique de l'Anses et des critères proposés par l'EFSA : impact potentiel du résultat de la revue systématique, quantité et qualité des données, source et confidentialité des données, transparence nécessaire et controverse du sujet de la saisine, ressources à mettre en œuvre.
- Les critères utilisés pour évaluer la qualité des études par type d'étude.
- Les méthodes retenues pour évaluer un ensemble d'études et pour intégrer les lignes de preuves (approches qualitatives ou quantitatives).
- Les modalités d'expression des conclusions concernant le poids des preuves.

Dans le cadre d'une analyse d'incertitude :

Le GT MER recommande, pour la sous-étape 2, d'identifier dans le schéma conceptuel les éléments qui pourraient ne pas être pris en compte dans l'évaluation mais essentiels à l'analyse d'incertitude.

Le GT MER recommande, pour la sous-étape 3, de décrire les méthodes d'analyse d'incertitude mobilisées et les modalités d'expression des conclusions concernant l'incertitude.

4.2 Établissement des lignes de preuve

4.2.1 Recherche et sélection des études

4.2.1.1 Recommandations initiales

- « **Développer une grille de lecture,** voire un tableau d'extraction des données, adaptable en fonction de la question de recherche et du type d'étude scientifique. Cette grille et ce tableau devront être élaborés et testés sur des études de cas en coordination avec les CES. Ces grilles pourront inclure trois types d'éléments concernant la pertinence des études par rapport à la question posée, des éléments descriptifs de l'étude et des critères d'évaluation de la qualité des études.
- **Le recours à au moins deux experts est souhaitable pour remplir les grilles de lecture.**
- **Lorsqu'une revue systématique est réalisée**, **certains grands principes doivent être appliqués :** utilisation d'au moins deux bases de données, la sélection des études par deux personnes indépendantes, et la définition des critères de sélection et d'exclusion des études en amont. Les collectifs d'experts pourront s'appuyer sur le document guide de l'EFSA (EFSA 2010).
- Dans le cas où une revue systématique n'est pas réalisée (en cas de ressources insuffisantes par exemple), le GT MER recommande que la **procédure** de recherche, de sélection et d'extraction soit décrite de manière aussi **précise** que possible dans le rapport d'expertise, en se rapprochant autant que possible des pratiques d'une revue systématique. »

4.2.1.2 Clarification et exemples d'application

Une grille descriptive des études épidémiologiques a été définie dans le cadre de l'étude de cas PDP1 avec le CES VLEP (Sissoko et al. 2017). L'ensemble des critères importants à considérer pour l'étude ont été repris afin que la grille puisse être complétée par les différents spécialistes (hygiéniste, épidémiologiste, métrologue). Des éléments sur les paramètres d'exposition et la stratégie d'échantillonnage ont été inclus dans la grille descriptive (cf. annexe 4).

La pertinence et la qualité d'une étude sont deux concepts différents. La pertinence porte sur la capacité de l'étude à répondre à la question posée. La qualité concerne le protocole de l'étude. La pertinence est évaluée avant la qualité, et conduit à la sélection d'un sous-ensemble d'études qui seront ensuite évaluées en détail avec des critères de qualité.

Lorsque deux experts sont mobilisés, il est préférable de choisir deux personnes ayant des compétences complémentaires.

Le terme « revue approfondie » est parfois utilisé mais il ne doit pas être confondu avec la revue systématique. La revue approfondie ne vérifie pas tous les critères de la revue systématique. Par exemple, dans une revue approfondie, la lecture détaillée des articles n'est réalisée que par un lecteur. La revue approfondie est généralement mise en oeuvre en cas de délais contraints.

La revue systématique a déjà été utilisée à l'Anses, par exemple sur les contaminations des poissons par les microcystines en s'appuyant sur le guide EFSA 2010 (Arnich 2017). La demande portait sur un état des connaissances et cette méthodologie a été choisie afin de recueillir les données de la littérature selon une méthode exhaustive, robuste et reproductible. Le principal intérêt est la traçabilité de la procédure de sélection des articles, selon des critères de pertinence (espèces comestibles pour l'Homme par exemple) et de qualité des résultats de contamination (obtenus selon une méthode analytique jugée suffisamment robuste). La principale difficulté de mise en œuvre a été le temps nécessaire à la mise en oeuvre de la revue systématique (6 mois), en particulier pour renseigner les grilles de lecture, par deux relecteurs, en raison d'un nombre élevé d'articles retenus (n = 69).
En santé environnement, certains collectifs peuvent utiliser d'autres guides de revue systématique par exemple ceux de la Collaboration Cochrane pour la phytopharmacovigilance.

La revue systématique est un élément central dans l'évaluation du poids des preuves. Cette méthode peut également être utile dans une analyse d'incertitude pour caractériser les connaissances disponibles et les connaissances manquantes.

Le comité de déontologie de l'Anses a rendu un avis relatif à la crédibilité de l'expertise (ANSES 2016a) dans l'objectif d'établir un cadre déontologique pour l'exploitation de la littérature scientifique. Cet avis recense les différentes sources d'erreurs et de biais dans les publications scientifiques et émet des recommandations sur l'usage des instruments bibliographiques aux fins d'expertise collective. Il souligne les difficultés soulevées par la sélection de publications scientifiques et de la littérature grise utilisées en cours d'expertise et préconise l'usage de la revue systématique (voire de la méta-analyse).

4.2.1.3 Recommandations mises à jour

Le GT MER formule les recommandations suivantes :

- **Développer une fiche de lecture adaptée à la question posée et au type d'étude scientifique. Un exemple est présenté dans l'annexe 4. La fiche inclura des éléments concernant la pertinence des études par rapport à la question posée et des éléments descriptifs de l'étude.**
- Le recours à au moins deux experts est souhaitable pour remplir les **fiches** de lecture.
- Lorsqu'une revue systématique est réalisée, certains grands principes doivent être appliqués : utilisation d'au moins deux bases de données, la sélection des études par deux personnes indépendantes, et la définition des critères de sélection et d'exclusion des études en amont. Les collectifs d'experts pourront s'appuyer sur le document guide EFSA (2010).
- Dans le cas où une revue systématique n'est pas réalisée (en cas de ressources insuffisantes par exemple), le GT MER recommande **de réaliser une revue approfondie de la littérature et de décrire de manière aussi précise que possible dans le rapport d'expertise la procédure de recherche, de sélection et d'extraction, en se rapprochant des pratiques d'une revue systématique.**

4.2.2 Évaluation de la qualité des études

4.2.2.1 Recommandations initiales

- Le GT MER recommande l'utilisation de **listes de critères formalisés** par type d'étude (épidémiologiques, toxicologiques, *etc*.) de manière à assurer la **transparence** du processus d'évaluation. Ces listes devront être élaborées en coordination avec les CES, en exploitant par exemple les critères mentionnés dans la méthode GRADE. Lorsque des études sont exclues, les **critères d'exclusion** doivent être explicites et définis par le collectif d'experts.
- Lorsque des études de synthèse sont considérées dans le processus d'expertise, le GT MER recommande l'utilisation de la méthode **AMSTAR** et **R-AMSTAR** (Kung et al. 2010) pour évaluer leur qualité.

4.2.2.2 Clarification et exemples d'application

L'objectif est d'évaluer la qualité des études sélectionnées lors de la sous-étape 1.

Dans le cadre de l'étude de cas PDP1 (Sissoko et al. 2017), la grille descriptive des études épidémiologiques a été complétée en intégrant des critères d'évaluation de la qualité (cf. annexe 4). La démarche OHAT a été privilégiée par rapport au système Grade, jugé plus complexe à mettre en œuvre. La méthode d'analyse des risques de biais des études individuelles proposée par la grille OHAT a été testée par le CES VLEP (Sissoko et al. 2017) pour évaluer des études

épidémiologiques portant sur les effets osseux du cadmium. Cette méthode a été jugée applicable par le groupe qui considère qu'elle n'augmente pas le temps d'expertise. L'approche OHAT a permis de répartir les études épidémiologiques selon trois paliers, dont un offrant la possibilité de rejeter l'étude.

Cette approche a été testée par le GT Particules dans l'étude de cas PDP5 (Redaelli et al. 2017) pour évaluer des études portant sur les effets sanitaires des composés et sources des particules de l'air ambiant. Ce groupe considère que l'approche OHAT est prometteuse car elle propose une structuration et une harmonisation des critères d'évaluation de la qualité des études. Des outils spécifiques ont été développés par le GT Particules à partir de la méthode OHAT : fiche descriptive, grille d'évaluation de la qualité d'une étude, grilles de synthèse des études sur l'Homme et sur l'animal, grille d'évaluation du niveau de preuve. Un aide mémoire, adapté au domaine, a été développé pour définir les facteurs influençant les biais des études individuelles. L'ensemble des grilles sont incrémentées et automatisées dans Excel. L'approche OHAT peut éventuellement n'être utilisée que pour l'analyse de la qualité interne des études individuelles et permettre un traitement indépendant des étapes suivantes d'établissement et d'intégration des lignes de preuve. La méthode OHAT a conduit les experts à modifier leur façon de lire les articles, à relever toutes les sources de biais, et à rendre tous les critères de qualité explicites. Ce groupe a estimé les durées moyennes d'évaluation de la qualité avec la méthode OHAT à une heure minimum pour une étude humaine ou une étude d'expérimentation animale.

La méthode R-AMSTAR a été testée par le GT Silice qui a évalué plusieurs dizaines de revues épidémiologiques (PDP6, Paillat, Fourneau, and Silice cristalline 2017). Le GT Silice considère que la méthode est pertinente, mais que certaines questions devaient être adaptées. Le collectif d'experts a utilisé R-AMSTAR pour évaluer la qualité de chaque revue épidémiologique avant de la décrire en détail dans une fiche de lecture. A partir du score obtenu, un taux de qualité de la revue est calculé. Un intérêt de R-AMSTAR est que cette méthode permet d'évaluer le niveau de qualité de chaque revue et d'éliminer celles qui n'atteignent pas un niveau suffisant, soit sur l'ensemble des critères, soit pour un critère clé. Le GT Silice considère qu'il serait utile d'adapter la méthode R-AMSTAR aux revues toxicologiques.

4.2.2.3 Recommandations mises à jour

Le GT MER recommande l'utilisation de listes de critères formalisés par type d'étude (épidémiologiques, toxicologiques, etc.) de manière à assurer la transparence du processus d'évaluation. **Un exemple est présenté en annexe 4.** Ces listes devront être élaborées en coordination avec les CES, en exploitant par exemple les critères mentionnés dans **les méthodes OHAT et** GRADE. Lorsque des études sont exclues, les critères d'exclusion doivent être explicites et définis par le collectif d'experts.

Lorsque des études de synthèse sont considérées dans le processus d'expertise, le GT MER recommande l'utilisation de la méthode AMSTAR et R-AMSTAR (Kung et al. 2010) pour évaluer leur qualité.

4.2.3 Analyse d'un ensemble d'études par type d'étude en vue d'établir des lignes de preuve

4.2.3.1 Recommandations initiales

- ➔ « Lors de l'utilisation de méthodes qualitatives avec ou sans notation (ex : critères de causalité Bradford Hill, méthode du CIRC, méthode dérivée de l'approche de Lewandovski et Rhomberg (Lewandowski and Rhomberg 2005) et dans le but d'assurer un niveau de transparence élevé au processus d'expertise, le GT MER recommande de **définir aussi précisément** que possible **les critères utilisés et la signification des notes attribuées** à ces critères par les experts.
- ➔ Le GT MER estime que **la méta-analyse est une méthode à privilégier** et recommande à l'Anses d'en réaliser sur des sujets stratégiques, en tenant compte des enjeux en termes de risques sanitaires et des ressources

disponibles (ressources humaines, disponibilité des données). Ces méta-analyses devraient être menées en s'inspirant des documents guides disponibles, notamment du document guide Cochrane.

- ➔ Lors de l'utilisation de méthodes quantitative, le GT MER recommande **d'analyser la sensibilité des résultats** aux paramètres d'entrée.
- ➔ Les **approches multicritères** semblent potentiellement intéressantes pour formaliser l'évaluation d'un ensemble d'études par un groupe d'experts. Toutefois, nous ne disposons pas d'un recul suffisant pour juger précisément de leur utilité dans le contexte de l'Anses. Le GT MER recommande de les tester sur des cas concrets variés en coordination avec les CES. »

4.2.3.2 Clarification et exemples d'application

Deux méthodes qualitatives d'établissement des lignes de preuve ont été comparées, celle développée par le GT Radiofréquences et Santé (RF), inspirée du CIRC, et celle développée par le GT Perturbateurs Endocriniens (PE) pour l'évaluation des effets sanitaires du bisphénol A (BPA). Ces méthodes ont été mises en œuvre dans des contextes spécifiques à l'Anses utilisant des schémas de détermination du niveau de preuve et des terminologies différentes. Cette étude de cas a nécessité d'établir des équivalences (cotation de la qualité des études et terminologie utilisée pour qualifier les lignes de preuve). Par ailleurs, l'application de l'arbre de décision BPA dans le cadre des RF s'est heurtée à la spécificité du corpus d'études RF. En effet, les deux expertises présentent initialement une bibliographie scientifique abondante, mais à la différence du BPA, les études sur les RF sont hétérogènes (réalisées sur de très nombreux signaux, fréquences, *etc.)*. Après discrimination sur la qualité de l'étude, le corpus d'études de bonne qualité sur le BPA concernant le même effet est plus fourni et plus homogène que le corpus d'études sur les RF. Cette étude de cas PDP3 (Roth-Delgado et al. 2017) montre qu'il est difficile d'appliquer directement une méthode qualitative d'établissement de ligne de preuve initialement développée dans un domaine donné à un autre domaine. En effet, les caractéristiques du corpus d'études conditionnent **la définition des lignes de preuves** et d'une façon générale la méthodologie applicable.

Le GT Particules a mobilisé la méthode OHAT pour établir les lignes de preuve dans l'étude de cas PDP5. Les résultats montrent que cette méthode OHAT permet d'utiliser des critères explicites. L'approche originale a été développée par ce GT pour harmoniser les critères de cotation des risques de biais.

Les approches multicritères semblent potentiellement intéressantes pour formaliser l'évaluation d'un ensemble d'études par un groupe d'experts. Toutefois, nous ne disposons pas d'un recul suffisant pour juger précisément de leur utilité dans le contexte de l'Anses. Le GT MER propose de les tester sur des cas concrets variés en coordination avec les CES.

4.2.3.3 Recommandations mises à jour

Lors de l'utilisation de méthodes qualitatives avec ou sans notation (ex : arguments en faveur de la causalité de Bradford Hill, méthode du CIRC, méthode dérivée de l'approche de Lewandovski et Rhomberg 2005, **OHAT**) et dans le but d'assurer un niveau de transparence élevé au processus d'expertise, le GT MER recommande de définir aussi précisément que possible les critères utilisés et la signification des notes attribuées à ces critères par les experts.

Le GT MER estime que la méta-analyse est une méthode à privilégier et recommande à l'Anses d'en réaliser sur des sujets stratégiques, en tenant compte des enjeux en termes de risques sanitaires et des ressources disponibles (ressources humaines, disponibilité des données). Ces méta-analyses devraient être menées en s'inspirant des documents guides disponibles, notamment du document guide Cochrane.

Lors de l'utilisation de méthodes quantitatives, le GT MER recommande d'analyser la sensibilité des résultats aux paramètres (ou variables) d'entrée.

4.3 Intégration des lignes de preuves

4.3.1 Recommandations initiales

➔ « Il est important d'encourager les experts à **décrire et expliciter les choix réalisés à chaque étape** afin d'assurer un niveau de transparence aussi élevé que possible. Certaines méthodes permettant de rendre le processus d'expertise transparent pourraient être utilisées dans ce but (ex : méthodes d'élicitation d'experts, intégration des avis d'experts dans les modèles prédictifs de type QSAR).

➔ Le GT MER recommande l'utilisation des méthodes **qualitatives proposées par le CIRC ou le WCRF** pour la combinaison des lignes de preuves, en incluant explicitement une liste permettant de vérifier la prise en compte des critères de Hill.

➔ Le GT MER recommande l'utilisation de la **modélisation statistique** pour combiner différentes lignes de preuve (notamment de l'approche Bayésienne). Ces approches nécessitant des compétences spécifiques et un investissement en temps relativement important, la faisabilité de leur utilisation devra cependant être évaluée sur des exemples concrets.

➔ Lorsque des méthodes quantitatives sont mobilisées, le GT MER recommande **d'analyser la sensibilité des résultats** aux paramètres d'entrée.

➔ Les méthodes **d'analyse multicritères** telles que celle proposée par Linkov et collaborateurs (Linkov et al. 2011) sont potentiellement intéressantes, mais le GT MER ne dispose pas d'un recul suffisant pour déterminer si elle pourrait être utile dans le contexte de l'Anses. Le GT MER recommande de la tester sur des cas concrets en coordination avec les CES.

➔ Dans le domaine spécifique de l'évaluation du mode d'action, la **méthode d'évaluation comparative du poids des preuves** pourrait être utilisée.

Le groupe de travail a recensé les méthodes existantes et a identifié pour cette étape l'opportunité de développer des méthodes spécifiques (ou des combinaisons de ces méthodes) pour l'intégration des lignes de preuves à l'Anses dans le cadre d'une convention de recherche et développement. »

4.3.2 Clarification et exemples d'application

Lors du séminaire, cinq participants sur vingt (25%) estimaient qu'ils ne pourraient pas appliquer la recommandation « Utiliser des modèles statistiques pour intégrer différentes lignes de preuve » et deux participants sur quinze (15%) estimaient qu'ils ne pourraient pas appliquer la recommandation « Décrire et expliciter les choix pour l'intégration des lignes de preuves selon une méthode qualitative » (Annexe 3). Les participants s'interrogent sur la compatibilité des méthodes statistiques avec l'expertise collective.

Des méthodes qualitatives d'intégration des lignes de preuve ont été développées à l'Anses, notamment pour les radiofréquences (arbre de décision) ou pour les perturbateurs endocriniens (grille de cotation). Ainsi plusieurs variantes ont été testées dans trois études de cas :

- L'étude de cas PDP3 (Roth-Delgado et al. 2017) a montré que la conclusion d'une expertise pouvait dépendre de la méthode utilisée pour établir et intégrer les lignes de preuve. Il est donc important de décrire en détail la méthode qualitative utilisée pour intégrer des lignes de preuve et de justifier le choix de la méthode en fonction du corpus d'études. Cette étude de cas a nécessité la construction d'une matrice d'intégration des lignes de preuves non prévue dans la méthode développée pour le BPA.
- L'étude de cas PDP7 (Beausoleil et al. 2017) du GT Perturbateurs endocriniens a conduit à la définition d'une méthode permettant de fixer des scores exprimant la probabilité de l'existence d'une association entre l'exposition à une substance et un effet néfaste. La méthode permet également de combiner cette probabilité avec un score reflétant le mode d'action de la substance étudiée. L'intérêt de cette approche est qu'elle permet de combiner différents types d'études (épidémiologiques, expérimentales *in vivo* et *in vitro* et études *in silico*) et de classer des perturbateurs endocriniens. Cependant, les résultats de cette méthode dépendent de plusieurs choix réalisés par les experts, et il est important de rendre ces choix aussi explicites que possible lors de sa mise en oeuvre. Il est également souhaitable de mobiliser au moins deux experts différents par type d'études.

- L'étude de cas PDP5 (Redaelli et al. 2017) du GT Particules a réalisé une étude de faisabilité pour deux méthodes qualitatives, OHAT (2015) et CIRC « adapté » (cette dernière ayant déjà été utilisée par le GT Radiofréquence dans la saisine 2012-SA-0091). L'étude conclut que la méthode OHAT est plus directive, plus structurante, transparente et reproductible que celle du CIRC « adapté ». Dans cette étude de cas, il semblait préférable d'utiliser une seule méthode pour les différentes étapes d'évaluation du poids des preuves.

Les méthodes d'analyse multicritères telles que celle proposée par Linkov et collaborateurs (Linkov et al. 2011) sont potentiellement intéressantes, mais le GT MER ne dispose pas d'un recul suffisant pour déterminer si elle pourrait être utile dans le contexte de l'Anses. Le GT MER propose aux CES de l'Anses de la tester sur des cas concrets.

4.3.3 Recommandations mises à jour

Il est important **que les experts décrivent en détails et justifient les méthodes utilisées pour rendre l'expertise scientifique transparente.**

Pour la combinaison des lignes de preuves, le GT MER recommande l'utilisation des méthodes qualitatives proposées **par le CIRC, le WCRF** (en incluant explicitement une liste permettant de vérifier la prise en compte des critères de Hill) **ou l'OHAT.**

Le GT MER **souligne l'intérêt** de la modélisation statistique pour combiner différentes lignes de preuve (notamment de l'approche bayésienne). **Cette approche** nécessitant des compétences spécifiques et un investissement en temps relativement important, la faisabilité de son utilisation devra cependant être évaluée sur des exemples concrets.

Lorsqu'**un modèle statistique est mobilisé**, le GT MER recommande d'analyser la sensibilité des résultats **aux hypothèses du modèle** et à ses paramètres d'entrée.

Dans le domaine spécifique de l'évaluation du mode d'action, la méthode d'évaluation comparative du poids des preuves pourrait être utilisée (Meek et al. 2014).

4.4 Identification et description des sources d'incertitude

4.4.1 Recommandations initiales

« Le GT MER propose d'adopter une typologie commune des sources d'incertitude composée des quatre classes suivantes :
i) sources d'incertitude liées au contexte, comprenant deux sous-classes (*cadrage* et *formulation de la question*),
ii) sources d'incertitude liées au *corpus* de connaissances, incluant celles liées à la variabilité (intrinsèque et liée à l'hétérogénéité) des caractéristiques d'intérêt. Cette classe comprend trois sous-classes (*état des connaissances*, *méthode de collecte de données* et *modèles existants*),
iii) sources d'incertitude liées à la méthode d'évaluation, comprenant trois sous-classes (*données sélectionnées*, *méthodes d'intégration des données* et *interprétation des résultats*)
iv) sources d'incertitude liées à la communication des résultats comprenant deux sous classes (*présentation des résultats* et *expression des résultats*). »

Le GT MER recommande :
- d'effectuer un recensement le plus complet possible des sources d'incertitude en se basant sur la typologie proposée,
- de présenter ce recensement par une matrice d'incertitude construite à partir de la typologie proposée (lignes) et de dimensions (colonnes) à définir en collaboration avec les collectifs d'experts à travers des études de cas,
- d'identifier les incertitudes pour lesquelles l'impact ne peut être évalué et indiquer les raisons de cette non-évaluation.

4.4.2 Clarification et exemple d'application

L'utilité et l'usage de la typologie des sources d'incertitude et du tableau des incertitudes ont soulevé plusieurs questions lors des études de cas et du séminaire. La typologie sert à recenser les sources d'incertitude de manière structurée. Les sources d'incertitude identifiées sont ensuite listées dans les lignes du tableau des incertitudes en respectant la structure de la typologie. Toutes les sources listées dans le tableau doivent ensuite être décrites. Cette description peut

s'effectuer de manière narrative et/ou à l'aide de descripteurs tels que la nature, l'amplitude et la direction des incertitudes. Lorsqu'une source d'incertitude n'est pas retenue pour les étapes ultérieures de l'analyse d'incertitude, elle doit néanmoins être listée dans le tableau des sources d'incertitude et la raison de sa non prise en compte, précisée, dans un but de tracabilité et de transparence de la démarche.

Lors du séminaire des demandes de clarification de la typologie ont été formulées. En conséquence, pour faciliter son applicabilité, le GT MER propose de définir par domaines de l'agence une liste des sources d'incertitude potentielles dans les différentes classes de la typologie présentée dans le rapport incertitude. Le GT MER juge en effet qu'il est difficile à l'heure actuelle de proposer un guide unique de construction du tableau des incertitudes qui soit applicable à l'ensemble les domaines de l'agence et préconise donc plutôt des mises en œuvre sur des cas variés par les différents collectifs d'experts afin de disposer d'exemples, qui constitueront des aides pour les expertises à venir.

S'agissant de l'étape de formulation de la question dans le tableau des incertitudes, il y a un besoin de clarification concernant l'identification des sources d'incertitude liées au contexte. Classiquement, une évaluation des risques sanitaires comprend une étape de reformulation de la question posée. Cette reformulation peut, le cas échéant, restreindre ou élargir l'exposition considérée ou la population cible, et dès lors restreindre ou élargir l'incertitude afférente. Dans un contexte réglementaire, il n'y a pas de reformulation de la question posée, le problème ne se pose donc pas. Toutefois, le GT MER considère que la typologie des sources d'incertitude doit être suffisamment générale pour être applicable à tous les domaines d'activité de l'agence ; quitte à ce que certaines classes de sources d'incertitude soient sans objet, comme cela est le cas de la classe *Contexte* dans le cadre d'une évaluation règlementaire.

Dans un contexte réglementaire, l'étude de cas AI3 sur l'« analyse des incertitudes dans le cadre de l'évaluation de dossiers de produits réglementés » (De Saint-Jores et al. 2017) a montré l'intérêt d'outils d'identification et d'analyse des incertitudes dans des domaines et contextes spécifiques, notamment pour garantir l'homogénéité des différentes évaluations (à travers des définitions et des exemples). En effet, la structure de l'évaluation des produits réglementés permet d'anticiper l'analyse et la prise en compte des incertitudes pour certaines composantes de ce type d'évaluation. Par exemple, pour les paramètres sensibles identifiables à l'avance, une caractérisation de l'incertitude associée est possible.
Dans le cadre de l'étude de cas AI3, un premier exemple de ce type d'outil a été élaboré pour un produit biocide (Tableaux 9 et 10).

Tableau 9 : exemple de présentation des incertitudes relatives à l'identification des dangers dans le cadre de l'évaluation d'un produit réglementé (étude de cas AI3)

Paramètres sensibles	**Incertitude/conservatisme Faibles** →			**Incertitude/conservatisme Elevés**
Modalités pour la caractérisation des dangers	Etude sur le produit	Etude sur produit similaire	Etude sur produit avec écarts de composition mais pire-cas	Classification par calcul
Toxicité aiguë orale				X
Toxicité aiguë cutanée				X
Toxicité aiguë par inhalation				X
Irritation cutanée				X
Sensibilisation				X
…				
Modalités absorption cutanée	Etude sur le produit ou sur un produit similaire et aux concentrations d'usage	Etude sur le produit ou sur produit similaire mais concentration plus faible que l'usage ou Etude sur produit des écarts de composition facilitant l'absorption	Etude sur le produit ou sur produit similaire mais application du ratio EFSA car concentrations supérieures à l'usage	Défaut EFSA
Absorption cutanée substance 1				Défaut 75%

Tableau 10 : exemple de présentation des incertitudes relatives à l'évaluation de l'exposition dans le cadre de l'évaluation d'un produit réglementé (étude de cas AI3)

Paramètres sensibles	**Incertitude/conservatisme Faibles** →			**Incertitude/conservatisme Elevés**
Usage / scénario	Usage clair et précis	Matériel non décrit précisément ou gamme de matériel large	Description générique du mode d'application	
Modèle	Modèle simple pour étude ou base de données de mesure (TNsG)	Modèle statistique (Riskofderm, EFSA, ART etc.)	Modèle physique (ConsExpo évaporation ou pulvérisation, SprayExpo, WEPM,…) Modèle spécifique créé pour l'évaluation (application d'appât en gel avec une seringue)	Scénario inverse
Domaine d'application des données d'exposition ou du modèle	Sur le produit représentatif	Usage similaire mais autre produit	Usage pire cas (ex : usage « pulvérisation » pour décrire un usage « injection » ou étude usage grand public pour un usage professionnel)	
Durée et fréquence journalière	Données d'enquête	Données d'enquête sur des usages similaires	Jugement d'expert en fonction de l'usage (ex : pour des traitements de surfaces, réalisme de traiter plusieurs sites dans une même journée,…)	Valeur par défaut guide ECHA (6-8h)

L'exemple traité dans l'étude de cas AI3 était un cas simple : une substance active et un usage unique avec un seul scénario d'exposition. La matrice est donc aisée à lire et la typologie applicable. Toutefois, ce cas n'est pas le plus fréquent : couramment un produit biocide a différentes applications, et/ou contient plusieurs substances actives ou co-formulants devant faire l'objet de l'évaluation. De plus, seul le champ de la santé humaine a été considéré dans cet exemple. Dans la plupart des cas, la multiplication des scénarios (plusieurs dizaines) et des champs concernés (efficacité, santé humaine et environnement) implique une multiplication des sources d'incertitude. Compte tenu du contexte (délais d'instruction et moyens à disposition), il est souhaitable d'adapter la liste des sources d'incertitude reportées au final en fonction de leur impact sur les conclusions de l'évaluation.

Dans un contexte non réglementaire, l'étude de cas portant sur la saisine « jouets » (AI4, Bonvallot et al. 2017) a montré l'applicabilité des deux outils (la typologie des sources d'incertitude et le tableau des incertitudes), leur utilité et leur plus-value. De plus, cette étude de cas a permis d'améliorer la proposition de structuration du tableau des incertitudes. En particulier, la colonne « source d'incertitude » peut être découpée en trois sous-colonnes : « objet », « choix effectué » et « origine : informations disponibles expliquant le choix » (cf. Tableau 11).

Tableau 11 : exemple de tableau d'incertitudes (étude de cas AI4)

Classes	**Sous-classes**	**Sources d'incertitude**			**Amplitude de l'impact sur le résultat** Faible, fort ou non qualifiable	**Direction** Sur/ sous estimation ou non qualifiable
		Objet	**Choix effectués**	**Origine : informations disponibles expliquant le choix**		
Contexte	**Cadrage** Ce qui est induit par le contexte /périmètre	Voie d'exposition	Sélection voie orale, exclusion respiratoire et cutanée	Voie orale supposée majoritaire	Supposé faible mais non quantifié	Sous-estimation
	Formulation de la question Ce qui entre dans le champ de l'expertise	Substances considérées	ATBC, DEHTP, DINCH, TXIB, DOIP	Bliographie et essais sur 21 jouets plastiques	Non qualifiable	Sous-estimation
Corpus de connaissances	**Etat des connaissances** Absence, incomplétude, inadéquation…	Substances évaluées	Exclusion du DOIP	Absence de données toxicologiques	Non qualifiable	Sous-estimation
	Méthodes de collecte des données Représentativité, protocole, puissance, méthode de mesure…	Essais de migration	Sur 21 jouets plastiques	Jugement d'experts non représentatif mais volonté de diversité	Non qualifiable	Non qualifiable
	Modèles existants Adéquation, validité, paramètres…	VTR	Sélection du modèle classique (cadrage)	Equation dose critisue animale/UF ou MOE	Non qualifiable	Non qualifiable
Méthode d'évaluation	**Données sélectionnées** Critères de sélection, jugement d'experts, extrapolation,…	Classe d'âge prise en compte	0-1 an	Mise en bouche à la naissance possible	Facteur 2 environ	Sur-estimation
	Méthodes d'intégration des données En lien avec le schéma conceptuel établi à la planification : choix des paramètres, extrapolation, logiciels utilisés, nombre de simulations…	Approche mélange	Exclusion de l'ERS cumulée	Effets critiques sélectionnés différents	Non qualifiable mais faisable	Sous-estimation
	Interprétation des résultats Peut générer des incertitudes en raison de biais cognitifs des experts, d'extrapolation d'un champ à l'autre ou de perception dans un contexte de forts enjeux économiques et politiques	Nous n'avons pas relevé d'incertitude particulière sur l'interprétation des résultats (approche classique QD > ou < 1)				
Communication des résultats	**Présentation des résultats**	Nous n'avons pas relevé d'incertitude particulière liée à la présentation ou l'expression des résultats				
	Expression des résultats					

4.4.3 Recommandations mises à jour

Le GT MER propose d'adopter une typologie commune des sources d'incertitude composée des quatre classes suivantes :
i) sources d'incertitude liées au contexte, comprenant deux sous-classes (*cadrage* et *formulation de la question*),
ii) sources d'incertitude liées au *corpus* de connaissances, incluant celles liées à la variabilité (intrinsèque et liée à l'hétérogénéité) des caractéristiques d'intérêt. Cette classe comprend trois sous-classes (*état des connaissances*, *méthode de collecte de données* et *modèles existants*),
iii) sources d'incertitude liées à la méthode d'évaluation, comprenant trois sous-classes (*données sélectionnées*, *méthodes d'intégration des données* et *interprétation des résultats*)
iv) sources d'incertitude liées à la communication des résultats comprenant deux sous classes (*présentation des résultats* et *expression des résultats*).

Le GT MER recommande à l'étape 1 « Identifier et décrire l'ensemble des incertitudes » :

- d'effectuer un recensement le plus complet possible des sources d'incertitude en se basant sur la typologie proposée. **Ce recensement inclura également les sources d'incertitude dont l'impact ne peut être évalué** (par exemple la restriction d'une évaluation à quelques substances dans le cadre d'une pollution complexe ou recueil de données insuffisant en raison d'un manque de temps ou de budget).
- de présenter ce recensement par un tableau d'incertitude construit à partir de la typologie proposée (lignes) et de dimensions (colonnes) à définir en collaboration avec les collectifs d'experts à travers des **exemples. En particulier, l'impact des sources d'incertitude évaluées doit être mentionné dans le tableau d'incertitude.**

4.5 Évaluation individuelle des sources d'incertitude

4.5.1 Recommandations initiales

« Pour estimer l'étendue des valeurs possibles, le GT MER recommande de distinguer les incertitudes liées à des hypothèses du modèle de celles liées à des paramètres ou variables d'entrée du modèle. Lorsque cela est possible, le GT MER recommande d'estimer les incertitudes individuelles à l'aide de méthodes statistiques.

Le GT MER recommande à l'Anses de veiller à la définition et à l'harmonisation du vocabulaire entre les différents collectifs d'experts sur cette question notamment lorsque des méthodes qualitatives, basées sur des descriptions nominales ou des échelles ordinales, sont appliquées. »

4.5.2 Clarification et exemples d'application

Quelle que soit l'évaluation des risques menée, le GT MER précise que cette étape peut être menée selon une approche qualitative ou quantitative. Chaque analyse des risques doit aborder explicitement la question de l'incertitude concernant chaque source. Ce travail contribuera à une meilleure transparence de l'évaluation des risques et donnera au décideur les éléments incertains ou/et les éléments variables aboutissants aux conclusions.

L'effort d'analyse (le choix de la méthode avec le niveau de détail des résultats) dépendra des objectifs de l'évaluation et du contexte décisionnel, des délais et des moyens mis à disposition du collectif d'experts. Le collectif doit chercher à dégager un consensus sur les incertitudes les plus marquantes.

De nombreuses méthodes existent pour évaluer les incertitudes, le choix de la méthode ou des méthodes retenues reposera en premier lieu sur le type d'informations traitées (cf. Tableau 2), ainsi que sur les compétences et les outils mis à disposition du collectif d'expert.

Lorsque les données permettant une estimation directe ne sont pas disponibles, les méthodes d'élicitation probabilistes peuvent parfois être utiles pour formaliser les connaissances d'un expert ou d'un groupe d'experts. L'objectif est de représenter les connaissances d'un ou de plusieurs experts sous la forme d'une distribution de probabilité décrivant leurs incertitudes sur la valeur du paramètre considéré. Cette approche a été mobilisée dans le cadre de l'autosaine 2015-SA-0087 qui concernait l'élicitation des sensibilités et des spécificités de différents indicateurs de conscience en abattoir pour différentes filières animales (Poster 11)(Thebault, Chiron, et al. 2017).
Le souhait de disposer d'une distribution de l'incertitude sur des valeurs de paramètres nous a conduits à choisir une élicitation formalisée. Les méthodes formalisées reposent sur quatre phases, la préparation et la sélection des experts, la formation des experts, l'élicitation individuelle et collective. Seul le résultat de l'élicitation collective, c'est-à-dire l'incertitude au niveau du groupe d'expert est pris en compte dans la modélisation. Parmi les trois méthodes formalisées décrites par l'EFSA (EFSA 2014), les contraintes de l'expertise nous ont amené à choisir la méthode Sheffield, pour des raisons de transparence de l'argumentation, de durée, de simplicité, du faible nombre d'experts à mobiliser (6-10) et de l'approche proche du fonctionnement des groupes de travail de l'ANSES. La phase d'élicitation collective différencie la méthode Sheffield des méthodes Delphi et Cooke, la méthode Sheffield étant la seule qui propose l'agrégation des opinions en face à face (behavioural agregation).

Cette possibilité d'échange direct entre experts oblige à un renforcement de la formation, un encadrement strict des phases et des échanges possibles entre experts, afin d'améliorer le partage des informations sans introduire de biais (heuristiques) entre experts.
Parmi les variantes de la méthode Sheffield, la méthode quartile a été choisie car elle était plus simple à mettre en oeuvre que la méthode roulette. Ces différents choix méthodologiques sont donc liés à notre contexte d'élicitation. Les différents choix possibles sont rapportés sur la figure 6 (en rouge l'exemple de l'autosaisine 2015-SA-0087).

Figure 6 : Choix d'une méthode d'élicitation pour l'autosaisine 2015-SA-0087 (poster 11)

La complexité du sujet pour les personnes élicitées a nécessité une phase de formation d'une journée. L'élicitation individuelle – conduite à distance – a été complété de manière systématique par une phase d'interview individuelle. Un outil rshiny *ad-hoc* a été développé pour faciliter la communication avec les experts au cours de l'élicitation. Cette organisation a permis une élicitation collective de qualité, garantissant robustesse, transparence et reproductibilité des valeurs obtenues. Les différents arguments ayant amené aux choix numériques de l'élicitation sont ainsi détaillés de façon anonyme dans le rapport lié à cette autosaisine 2015-SA-0087. Cependant, le coût en termes de temps et de ressources qu'un tel processus nécessite doit amener à réfléchir en amont sur la plus value pour l'expertise des valeurs à éliciter.

L'élicitation probabiliste a également été mise en oeuvre dans le cadre d'une analyse de risque phytosanitaire (ARP) relative à *Ambrosia trifida* L. Elle a fait l'objet de l'étude de cas AI2 (Tassus et al. 2017). Dans ce contexte, un collectif d'expert a testé l'utilisation d'une élicitation formelle des connaissances d'experts pour analyser quantitativement les incertitudes. La démarche d'élicitation choisie est également celle de Sheffield, selon la variante dite « Roulette ». Les résultats se présentent sous la forme de distribution de probabilité. L'exercice d'élicitation s'est déroulé en utilisant le logiciel « *The MATCH Uncertainty Elicitation Tool* ». L'intérêt de ce logiciel est qu'il est d'usage intuitif et disponible gratuitement en ligne (Morris, Oakley, and Crowe 2014). Cette étude de cas a montré qu'en l'absence de donnée observée, l'utilisation d'une méthode formelle d'élicitation sert de support de discussion. Elle demande aux experts d'expliciter les éléments pris en considération dans la formulation des connaissances. Elle contribue ainsi à améliorer la transparence de l'expertise. Cependant, le temps nécessaire à la mise en œuvre de cette approche devra être optimisé et pris en compte lors de la planification de l'expertise.

4.5.3 Recommandations mises à jour

Pour estimer l'étendue des valeurs possibles, le GT MER recommande de distinguer les incertitudes liées à des hypothèses du modèle de celles liées à des paramètres ou variables d'entrée du modèle. Lorsque cela est possible, le GT MER recommande d'estimer les incertitudes individuelles à l'aide de méthodes statistiques **et/ou probabilistes**.

Le GT MER recommande à l'Anses de veiller à la définition et à l'harmonisation du vocabulaire entre les différents collectifs d'experts **sur cette question** notamment lorsque des méthodes qualitatives, basées sur des descriptions nominales ou des échelles ordinales, sont appliquées.

4.6 Évaluation de l'impact combiné des sources d'incertitude

4.6.1 Recommandations initiales

« Le GT MER recommande :
- d'évaluer quantitativement l'impact combiné des incertitudes sur le résultat, lorsque les données le permettent ;
- de tenir compte des dépendances entre les variables dans l'évaluation de l'impact combiné de leurs incertitudes afin de se limiter aux combinaisons réalistes. Lorsque cela est possible, le GT MER recommande d'estimer les dépendances entre les variables à l'aide de méthodes statistiques. »

4.6.2 Clarification et exemples d'application

Quelle que soit l'évaluation des risques menée, la GT MER précise que l'évaluation de l'impact combiné des sources d'incertitude peut être qualitative ou quantitative, et qu'elle nécessite de combiner les incertitudes retenues à l'étape précédente afin d'aboutir à des conclusions exprimées avec leur incertitude associée. Cette combinaison des incertitudes pourra se faire quantitativement (comme elle se fait déjà dans de nombreux travaux de l'Anses, voir exemples ci-dessous) ou qualitativement par un consensus des experts du collectif sur l'incertitude finale résultant de la combinaison.

Les approches séparant variabilité et incertitude dans le cadre de saisines sont encore peu fréquentes. Un exemple est présenté dans la saisine 2010-SA-0301, portant sur l'évaluation des risques lié à *Vibrio parahaemolyticus* via la consommation de produits de la mer (Poster 13) (Thebault, Guillier, et al. 2017). L'approche choisie était une méthode quantitative probabiliste (simulations de Monte-Carlo à deux dimensions) séparant spécifiquement variabilité et incertitude. Cette approche à deux dimensions est comparable à celle abordée dans une saisine plus ancienne (2002) de l'Afssa portant sur *Cryptosporidum parvum* dans l'eau d'alimentation (Pouillot et al. 2004). Dans le cadre de la saisine sur *Vibrio parahaemolyticus*, dès le schéma conceptuel les principales sources d'incertitude et de variabilité prises en compte dans la modélisation ont été identifiées séparément (Figure 7).

Figure 7: schéma conceptuel séparant variabilité et incertitude (poster 13)

D'autres sources d'incertitude n'ont pas pû être quantifiées ou leur effet non évalué, même qualitativement ; elles ont toutefois été listées. D'autres enfin, ont été étudiées *via* des scénarios distincts (durée d'émersion chez des professionnels). Les sources d'incertitude non quantifiables jugées très critiques figurent dans le schéma conceptuel pour montrer leur rôle dans le mécanisme global, et des recommandations pour l'acquisition de données ont été formulées. Les résultats de la modélisation ont été exprimés sous la forme de la médiane et d'intervalles de crédibilité à 95% de la moyenne. Cependant, dans le rapport, la formulation « médiane de la moyenne » a été évitée car jugée trop complexe pour beaucoup de lecteurs. Les résultats obtenus par modélisation montraient de très fortes incertitudes (intervalles de crédibilité extrèmement larges, avec une estimation du nombre de cas pouvant être comprise entre 28 et 26 875 cas). Par contre, une approche relative, montrait l'impact potentiel d'une augmentation des températures (facteur 7 d'augmentation du nombre de cas) ou d'une dérive des pratiques zootechniques (facteur 50 d'augmentation du nombre de cas). En termes de communication face à ces grandes incertitudes, il paraît plus judicieux et utile d'exprimer le risque en risque relatif et non en risque absolu, pour mieux montrer l'impact de facteurs éventuels sur le risque. Une hiérarchisation des incertitudes aurait été utile, des recommandations plus ciblées en matière d'efforts de recherche.

Les simulations de Monte-Carlo à deux dimensions ont aussi été uilisée pour quantifier l'impact combiné de l'incertitude en évaluation du risque lié aux allergènes alimentaires (Poster 14) (Crépet 2017). Dans cet exemple, les variables d'entrée ont été modélisées par inférence bayésienne ce qui facilite la séparation de la variabilité et de l'incertitude au cours des différentes étapes d'évaluation du risque.

L'impact du choix du modèle (choix du formalisme et des équations) sur les conclusions peut être analysé à l'aide d'approche de type « ensemble modelling » (Tinker et al. 2015).

4.6.3 Recommandations mises à jour

Le GT MER recommande :
- d'évaluer quantitativement l'impact combiné des incertitudes sur le résultat, lorsque les données le permettent ;
- de tenir compte des dépendances entre les variables dans l'évaluation de l'impact combiné de leurs incertitudes afin de se limiter aux combinaisons réalistes. Lorsque cela est possible, le GT MER recommande d'estimer les dépendances entre les variables à l'aide de méthodes statistiques.

4.7 Hiérarchisation des sources d'incertitude selon leur contribution à l'incertitude globale

4.7.1 Recommandations initiales

« Le GT MER recommande de réaliser, dans la mesure du possible et en fonction des besoins de l'évaluation, une analyse de sensibilité pour classer les sources d'incertitudes. »

4.7.2 Clarification et exemples d'application

Le GT MER rappelle que la hiérarchisation des sources d'incertitude est utile, en particulier, pour formuler des recommandations d'actions visant à réduire l'incertitude.

Le GT MER note que l'analyse de sensibilité nécessite des compétences spéciques. Ce type d'analyse a déjà été utilisé à l'ANSES (Audiat-Perrin et al. 2017). Selon les besoins, des analyses de sensibilité qualitative peuvent également être utilisées pour aboutir à une hiérarchisation des sources d'incertitude faisant consensus.

A titre d'exemple, (Audiat-Perrin et al. 2017) a évalué l'impact sanitaire de mesures de gestion du risque en tenant compte de la variabilité et de l'incertitude des paramètres d'entrée du modèle. Cette analyse a été réalisée dans le cadre d'une évaluation quantitative du risque lié à la présence

de salmonelles dans des fromages. Pour cela, un modèle de simulation de Monte-Carlo à deux dimensions a été développé et une analyse de sensibilité a été réalisée en calculant les indices de Sobol des sept paramètres d'entrée pour lesquels l'incertitude a été prise en compte (Saltelli et al. 2008). La hiérarchisation des paramètres en fonction de leur contribution à l'incertitude globale, permet de cibler les données à collecter pour diminuer cette incertitude. D'autre part, l'utilisation d'un modèle à deux dimensions permet d'évaluer et comparer les mesures de gestion en tenant compte de l'incertitude liée aux paramètres d'entrée.

4.7.3 Recommandation mise à jour

Le GT MER recommande de réaliser, dans la mesure du possible et en fonction des besoins de l'évaluation, une analyse de sensibilité **pour classer les sources d'incertitude.**

4.8 Expression des conclusions

4.8.1 Recommandations initiales

Recommandations du rapport sur le poids des preuves :

« La conclusion de l'analyse du poids des preuves doit toujours être associée à un texte explicatif répondant à la question posée, notamment lorsque le poids des preuves est exprimé de manière quantitative.

Lorsque le poids des preuves est analysé avec une méthode qualitative, le GT MER recommande d'exprimer les conclusions concernant le poids des preuves selon une classification en 4 niveaux correspondant à des niveaux de preuve croissants. Une classe supplémentaire « évaluation impossible » peut également être considérée. Le GT MER recommande que chaque niveau soit défini précisément dans les rapports d'expertise.

Le GT MER recommande de structurer le contenu de la conclusion en adaptant les travaux de la collaboration Cochrane et de la méthode GRADE aux domaines de l'agence. Un rapprochement avec la Collaboration Cochrane pourrait être envisagé afin de bénéficier de son expérience.

Le GT MER recommande de caractériser l'incertitude dans la conclusion soit sous forme qualitative, soit sous forme quantitative, selon la méthode d'analyse du poids des preuves utilisée. Cette recommandation est également formulée dans le rapport « Traitement de l'incertitude dans le processus d'évaluation des risques sanitaires des substances chimiques » rédigé par le GT Perturbateurs Endocriniens et reprotoxiques de catégorie 2 ».

Recommandations du rapport sur l'analyse d'incertitude :

« Le GT MER recommande :

- d'inclure systématiquement dans les évaluations un chapitre conclusif sur l'incertitude présentant l'impact combiné des incertitudes et les principales sources évaluées, et listant les sources d'incertitude qu'il n'a pas été possible d'analyser. Ce chapitre doit résumer ces points essentiels afin d'éclairer efficacement les parties prenantes ;
- de présenter le recensement et la description des incertitudes sous la forme d'une matrice ;
- lorsque cela est possible et important pour l'évaluation, de présenter les résultats sous forme probabiliste telle que la probabilité de dépassement, l'intervalle de crédibilité, l'intervalle de confiance et à défaut d'utiliser une échelle ordinale uniformisée. Des modalités de présentation précises seront définies en collaboration avec les collectifs d'experts à travers des études de cas. »

4.8.2 Clarification et exemples d'application

Pour exprimer le niveau de preuve, le GT MER recommande d'utiliser une échelle avec un nombre pair de niveaux pour éviter l'écueil d'un positionnement neutre. Plusieurs options ont été envisagées : une échelle à 2 niveaux semblait insuffisante pour nuancer les différentes valeurs possibles de preuves, une échelle à 6 niveaux semblait au contraire excessive. Ainsi, une échelle à 4 niveaux apparaît plus pertinente. A ces 4 niveaux, le GT MER a estimé nécessaire d'ajouter un niveau supplémentaire correspondant à la situation où aucune réponse ne peut être formulée. A terme, il serait souhaitable que les définitions des quatre niveaux soient harmonisées à l'Anses. Une échelle numérique présente l'avantage d'être plus neutre qu'une échelle basée sur des termes exprimant une position tranchée (« effet avéré », « absence d'effet »...).

Un exemple de classification en quatre niveaux est présenté dans l'étude de cas (Beausoleil et al. 2017) sur les perturbateurs endocriniens.

L'étude de cas (Roth-Delgado et al. 2017) confirme que la terminologie utilisée (« effet possible », « effet avéré » *etc*.) dans les méthodes qualitatives peut avoir un impact fort sur l'interprétation des conclusions et que cette terminologie doit être modulée de manière narrative dans les conclusions.

La conclusion sur l'analyse d'incertitude doit inclure un texte explicatif utilisable par le gestionnaire, comme cela est le cas, par exemple, dans les rapports d'analyse de risque phytosanitaire ou encore dans le rapport sur l'étude de l'alimentation totale infantile publiés par l'Anses et ne doit être restreinte à des résultats numériques.

Sauf en cas de contraintes réglementaires, l'analyse d'incertitude doit être faite systématiquement. En cas de contraintes de temps importantes, une méthode simple peut être utilisée. L'avis 2017-SA-0071 relatif au risque pour la santé humaine lié à la contamination par l'Herpès B d'un groupe de Macaques de Java traité par un groupe d'expertise collective d'urgence a listé simplement les différentes sources d'incertitude. Le rapport d'expertise collective de la saisine « Jouets » (2013-SA-0176) présente les résultats d'une analyse qualitative de l'incertitude de manière simple et synthétique, à l'aide d'un tableau des incertitudes incluant une description des principales sources et une évaluation qualitative des impacts.

Il est possible d'exprimer les conclusions sous forme quantitative même lorsque la disponibilité des données est limitée, par exemple en utilisant des méthodes d'élicitation probabiliste des experts ou des méthodes bayésiennes. Les résultats de l'évaluation des incertitudes s'expriment alors à l'aide de distribution de probabilités. Cette approche est mise en œuvre à l'Anses dans le cadre d'une expertise collective sur les méthodes d'échantillonnage en cours (Poster 11, Thebault, Chiron et al., 2017) ou de l'étude de cas AI2 ARP *Ambrosia trifida* (Tassus et al. 2017), par exemple. Dans l'exemple de l'ARP, l'expression probabiliste présente l'avantage d'exprimer plus précisément les incertitudes (plutôt qu'une échelle ordinale à trois classes dont les termes ne sont pas clairement définis) et contribue à une meilleure robustesse des résultats.

Si une échelle ordinale est utilisée pour qualifier le niveau d'incertitude, il serait souhaitable que les définitions des niveaux soient harmonisées à l'Anses. Le schéma d'aide à la décision pour les ARP utilisé à l'Anses mobilise une échelle ordinale à trois modalités (niveau d'incertitude : faible, modéré ou élevé) pour qualifier l'incertitude. Son adaptation à d'autres domaines pourrait être étudiée.

L'incertitude est inhérente à l'évaluation des risques, elle se présente à chacune de ses étapes et impacte les résultats de l'évaluation. Elle doit donc être prise en compte tout le long de l'évaluation des risques. Ainsi, il est essentiel que l'analyse d'incertitude fasse partie intégrante de l'évaluation des risques. D'autre part, la démarche de l'analyse d'incertitude et ses résultats doivent être présentés dans leur totalité, dans un chapitre dédié du rapport d'expertise collective et dans l'avis, tel que cela est illustré dans le rapport d'expertise collective sur l'étude de l'alimentation totale infantile (EAT infantile, Saisine n°2010-SA-0317). Le rapport sur l'EAT infantile dédie, en effet, une partie à la présentation de la méthode utilisée, aux limites de l'étude et aux incertitudes recensées. Des recommandations pour pallier aux incertitudes sont énoncées. L'ensemble de ces éléments sont également rapportés en conclusion du rapport et dans l'avis de l'Anses. Une communication aussi complète sur l'incertitude entourant les résultats de l'étude répond à un enjeu de transparence.

4.8.3 Recommandations mises à jour

Le GT MER recommande de présenter systématiquement les conclusions d'une évaluation du poids des preuves et les conclusions d'une analyse d'incertitude dans des chapitres spécifiques, non seulement dans l'avis de l'Anses, mais aussi dans le rapport d'expertise collective associé.

Dans le cadre d'une évaluation du poids des preuves :

Le GT MER recommande :

- d'exprimer les conclusions concernant le poids des preuves soit de manière quantitative, soit de manière qualitative selon une classification en 4 niveaux correspondant à des niveaux de preuve

croissants. Une classe supplémentaire « évaluation impossible » peut également être considérée. Le GT MER recommande que chaque niveau soit défini précisément dans les rapports d'expertise,

- **de rédiger un texte explicatif répondant à la question posée, notamment lorsque le poids des preuves est exprimé de manière quantitative,**

- de structurer le contenu de la conclusion en adaptant les travaux de la collaboration Cochrane et de la méthode GRADE aux domaines de l'agence. Un rapprochement avec la Collaboration Cochrane pourrait être envisagé afin de bénéficier de son expérience,

- de caractériser l'incertitude dans la conclusion concernant le poids des preuves, soit sous forme qualitative, soit sous forme quantitative, selon la méthode utilisée. Cette recommandation est également formulée dans le rapport « Traitement de l'incertitude dans le processus d'évaluation des risques sanitaires des substances chimiques » rédigé par le GT « Perturbateurs Endocriniens et reprotoxiques de catégorie 2 ».

Dans le cadre d'une analyse d'incertitude :

Le GT MER recommande

- de présenter dans la conclusion l'impact combiné des incertitudes ainsi que les principales sources évaluées, en listant les sources d'incertitude qu'il n'a pas été possible d'analyser. Les points essentiels doivent être résumés à chaque étape de l'évaluation afin d'éclairer efficacement les parties prenantes,

- de présenter le recensement et la description des incertitudes **dans un tableau**,

- lorsque cela est possible et important pour l'évaluation, de présenter les résultats sous forme probabiliste telle que la probabilité de dépassement, l'intervalle de crédibilité, l'intervalle de confiance et à défaut d'utiliser une échelle ordinale uniformisée.

4.9 Harmonisation des pratiques de l'Anses et partage d'expériences

4.9.1 Harmonisation du vocabulaire

Le GT MER recommande à l'Anses de veiller à la définition et à l'harmonisation du vocabulaire entre les différents collectifs. Un glossaire a été initié.

4.9.2 Partage d'expériences

L'objectif recherché par l'évaluation du poids des preuves et l'analyse d'incertitude est de rendre transparents les résultats de l'expertise pour en assurer la crédibilité et la confiance de la part de la communauté scientifique comme des parties prenantes. Afin de faciliter l'appropriation des concepts et méthodes par les collectifs d'experts :

- ➔ Le GT MER recommande de mettre à disposition des collectifs d'experts un **soutien méthodologique** pour faciliter la mise en œuvre des méthodes d'évaluation du poids des preuves et d'analyse d'incertitude. Des référents méthodologiques pourraient être identifiés pour aider les collectifs d'experts sur la base de leurs expériences acquises notamment lors des études de cas menées dans le cadre du GT MER (tableaux 6 et 7).
- ➔ Le GT MER recommande de mettre en place **un système d'information** pour capitaliser les travaux d'expertise. Ce système d'information inclurait :
  - un descriptif de la méthode d'évaluation du poids des preuves et d'analyse d'incertitude mise en œuvre,
  - des fiches descriptives de méthodes
  - des grilles de lecture par type d'études,
  - le contenu des grilles de lecture des revues bibliographiques,…

4.9.3 Relations entre les recommandations du GT MER et le système qualité

Les documents du système qualité Anses ont été développés en regard d'une norme (NF X-50-110) publiée en 2003.

Pour l'étape de planification de l'évaluation :

- ➔ Le GT MER recommande, pour la sous étape de cadrage, d'utiliser le formulaire **« grille de questionnement sociologique »,** mis à disposition par l'unité risques et société, pour identifier les enjeux importants dans le contexte de la saisine.
- ➔ Le GT MER recommande, pour la sous-étape de formulation de la question, d'utiliser **le formulaire de profil de recherche bibliographique** qui prend en compte les structures de description de la question posée (PICO, PECO, etc.). [ANSES/PR1/9/06-01]

Le GT MER recommande d'inscrire l'évaluation du poids des preuves et l'analyse d'incertitude au sein du processus d'expertise de l'Anses. Dans ce but, le GT MER recommande d'intégrer explicitement les éléments relatifs à l'évaluation du poids des preuves et l'analyse d'incertitude dans le **document de cadrage interne** [ANSES/PR1/9/01-01], dans le rapport d'expertise collective [ANSES/PR1/9/01-04], et dans l'avis de l'Anses [ANSES/PR1/9/01-06], de façon à accroître la transparence des travaux d'expertise.

Le GT MER recommande également de développer **un formulaire de description du plan de développement de la méthode d'évaluation** par le collectif d'experts. Pour cela, le GT MER recommande d'adapter le plan détaillé proposé par l'OHAT aux besoins de l'agence.

5 Conclusions du groupe de travail

Les études de cas et les résultats du séminaire présentés dans ce rapport montrent que la démarche générale d'évaluation du poids des preuves et d'analyse d'incertitude proposée par le GT MER est globalement bien comprise par les collectifs Anses. En parallèle à ce travail, certains collectifs Anses se sont emparés de la démarche et des méthodes proposées par le GT MER, à l'exemple de la saisine 2016-SA-0120 pour évaluer le risque d'introduction de la dermatose nodulaire contagieuse ou plus récemment le GT *Citrus greening* pour la conduite d'une analyse de risque phytosanitaire.

Le séminaire a montré que certaines recommandations présentées dans les rapports du GT MER publiés en 2016 devaient être clarifiées. Les études de cas ont confirmé que les démarches générales d'évaluation du poids des preuves et d'analyse d'incertitude étaient applicables dans les différents domaines couverts par l'Anses. Les résultats du séminaire et des études de cas ont permis au GT MER d'illustrer et d'actualiser ses recommandations.

Plus précisément, les études de cas ont révélé que certaines méthodes étaient à la fois pertinentes et applicables par les collectifs Anses :

- les structures d'information PICO/PECO et les schémas conceptuels pour la planification de l'évaluation
- les fiches de lecture et de critères de qualité pour sélectionner et évaluer les études pertinentes
- la méthode OHAT pour analyser la qualité et les risques de biais des études individuelles, établir des lignes de preuves, et intégrer des lignes de preuves (même si cette méthode est parfois perçue chronophage)
- la méthode R-AMSTAR pour évaluer la qualité des études de synthèse bibliographique, sous réserve d'adaptation de certains aspects de la méthode aux contextes des saisines
- les tableaux d'incertitude pour identifier, lister et décrire les incertitudes de manière transparente
- l'élicitation probabiliste, à l'exemple de la méthode Sheffield, pour évaluer les incertitudes relatives aux connaissances d'experts.

D'autres méthodes ont suscité un intérêt de la part des collectifs d'experts mais leur application pose des difficultés pratiques :

- la revue systématique demande des moyens humains importants ; la solution alternative est de conduire une revue approfondie de la littérature.
- la méta-analyse demande de concilier des données pas toujours accessibles et des protocoles expérimentaux pouvant être hétérogènes
- l'analyse de sensibilité nécessite des compétences spécifiques.

Enfin, des méthodes n'ont pas encore été testées par les collectifs Anses, notamment la modélisation statistique et l'analyse multicritères pour intégrer des lignes de preuve.

Dans la perspective du déploiement des recommandations au sein de l'agence, la mise en place de formations et de soutiens méthodologiques faciliterait l'appropriation de la démarche et des méthodes par les collectifs. La mise à disposition d'un système d'information permettrait le partage de l'expertise entre les collectifs.

Date de validation du rapport d'expertise collective par le groupe de travail : 26 juin 2017

6 Bibliographie

AFNOR. 2003. NF X 50-110. Qualité en expertise - Prescriptions générales de compétence pour une expertise.

ANSES. 2016a. Avis n° 2016-1 relatif à la crédibilité de l'expertise : Un cadre déontologique pour l'exploitation de la littérature scientifique [Saisine 13]. Maisons-Alfort: Anses.

ANSES. 2016b. Évaluation du poids des preuves à l'Anses : revue critique de la littérature et recommandations à l'étape d'identification des dangers. Maisons-Alfort: Anses.

ANSES. 2016c. Prise en compte de l'incertitude en évaluation des risques : revue de la littérature et recommandations pour l'Anses. Maisons-Alfort: Anses.

ANSES. 2016d. Traitement de l'incertitude dans le processus d'évaluation des risques sanitaires des substances chimiques - Réflexion sur la caractérisation et la prise en compte de l'incertitude en évaluation des risques sanitaires au sein de l'Anses. Maisons-Alfort: Anses.

Arnich, N. 2017. "Revue systématique de la contamination du poisson par des microcystines." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Audiat-Perrin, F., F. Tenenhaus-Aziza, V. Michel, and M. Sanaa. 2017. "Quantification de l'impact de mesures de gestion du risque de salmonellose lié à la consommation de fromage au lait cru avec un modèle AQR à 2 dimensions." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Beausoleil, C. , R. Beaudouin, C Minier, S. Fraize-Frontier, D. Makowski, B. Meek, C. Bladier, and C. Rousselle. 2017. "Établissement d'une grille de cotation pour les effets perturbateurs endocriniens." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Bernier, J., E. Parent, and J.J. Boreux. 2000. *Statistique pour l'environnement: traitement bayésien de l'incertitude*: Tec & Doc.

Bonvallot, N., C. Brochot, P. Glorennec, A. Mathieu, and C. Leroux. 2017. "Elaboration d'une matrice d'incertitude – retour d'expérience sur la saisine « jouets »." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Cox, D. R. 2006. *Principles of Statistical Inference*: Cambridge University Press.

Crépet, A. 2017. "Méthodes probabilistes pour évaluer l'impact combiné de l'incertitude lié au risque de réaction allergique à l'arachide." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Cullen, A. C., and H. C. Frey. 1999. *Probabilistic techniques in exposure assessment. A handbook for dealing with variability and uncertainty in models and inputs*. New York: Plenum Press.

De Saint-Jores, J., Y. Bayona, S. Fraize-Frontier, and S. Pierlot. 2017. "Analyse des incertitudes dans le cadre de l'évaluation de dossiers de produits réglementés." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Efron, B., and R.J. Tibshirani. 1993. *An introduction to the bootstrap*: Chapman & Hall/CRC.

EFSA. 2007. "Guidance of the Scientific Committee on a request from EFSA related to Uncertainties in Dietary Exposure Assessment (EFSA-Q-2004-019)." *EFSA Journal* 5 (1):438, 54 pp. doi: 10.2903/j.efsa.2007.438.

EFSA. 2010. "Application of systematic review methodology to food and feed safety assessments to support decision making (EFSA-Q-2008-717)." *EFSA Journal* 6:1637, 90 pp. doi: 10.2903/j.efsa.2010.1637.

EFSA. 2014. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. European Food Safety Authority (EFSA).

EFSA. 2016. Revised draft for internal testing - Guidance on Uncertainty in EFSA Scientific Assessment. In *EFSA Journal*. Parma, Italy: European Food Safety Authority (EFSA).

Fite, J. , O. Merckel, Sabrina Havard, C. Roth, J-F. Dore, J-P. Marc-Vergnes, G. Ledoigt, A. Massardier-Pilonchery, M. Touvier, O. Bruyère, D. Makowski, L. Watier, and C. Bladier. 2017. "Sensibilité des conclusions à la méthode qualitative d'évaluation du poids des preuves." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

IPCC. 2005. Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties. WMO/UNEP.

IPCS. 2004. IPCS Risk Assessment Terminology. Harmonization Project Document No. 1. Geneva: WHO Press.

IPCS. 2008. Uncertainty and Data Quality in Exposure Assessment. Harmonization Project Document No. 6. In *IPCS Harmonization Project Document*. Geneva: WHO Press.

Khamisse, E., I. Albert, C. Saegerman, S. Fraize-Frontier, and P. Colin. 2017. "Apport du schéma conceptuel dans l'évaluation des risques d'introduction de Salmonella spp. en alimentation animale." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Knol, A. B., A. C. Petersen, J. P. van Der Sluijs, and E. Lebret. 2009. "Dealing with uncertainties in environmental burden of disease assessment." *Environmental Health: A Global Access Science Source* 8 (1).

Kung, Jason, Francesco Chiappelli, Olivia O. Cajulis, Raisa Avezova, George Kossan, Laura Chew, and Carl A. Maida. 2010. "From Systematic Reviews to Clinical Recommendations for Evidence-Based Health Care: Validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for Grading of Clinical Relevance." *The Open Dentistry Journal* 4:84-91. doi: 10.2174/1874210601004020084.

Lewandowski, T. A., and L. R. Rhomberg. 2005. "A proposed methodology for selecting a trichloroethylene inhalation unit risk value for use in risk assessment." *Regul Toxicol Pharmacol* 41 (1):39-54. doi: 10.1016/j.yrtph.2004.09.003.

Linkov, I., P. Welle, D. Loney, A. Tkachuk, L. Canis, J. B. Kim, and T. Bridges. 2011. "Use of multicriteria decision analysis to support weight of evidence evaluation." *Risk Analysis* 31 (8):1211-1225.

Maxim, L. 2014. "A systematic review of methods of uncertainty analysis and their applications in the assessment of chemical exposures, effects, and risks." *International Journal of Environmental Health Research*. doi: 10.1080/09603123.2014.980782.

Meek, M. E., A. Boobis, I. Cote, V. Dellarco, G. Fotakis, S. Munn, J. Seed, and C. Vickers. 2014. "New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis." *Journal of Applied Toxicology* 34 (1):1-18.

Morris, D. E., J. E. Oakley, and J. A. Crowe. 2014. "A web-based tool for eliciting probability distributions from experts." *Environmental Modelling and Software* 52:1-4. doi: 10.1016/j.envsoft.2013.10.010.

O'Hagan, A., C.E. Buck, A. Daneshkhah, J.R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakley, and T. Rakow. 2006. *Uncertain judgements: Eliciting Experts' Probabilities*: John Wiley & Sons, Ltd.

Paillat, A. , C. Fourneau, and GT Silice cristalline. 2017. "Étude de cas « Silice cristalline » - Applicabilité de la grille R-Amstar en vue de l'analyse de revues épidémiologiques et toxicologiques." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Parent, E., and J. Bernier. 2007. *Le raisonnement bayésien - Modélisation et inférence*. Paris: Springer-Verlag France.

Pouillot, R., P. Beaudeau, J. B. Denis, and F. Derouin. 2004. "A Quantitative Risk Assessment of Waterborne Cryptosporidiosis in France Using Second-Order Monte Carlo Simulation." *Risk Analysis* 24 (1):1-17. doi: 10.1111/j.0272-4332.2004.00407.x.

Redaelli, M., A. Brasseur, A. Baeza-Squiban, M. Blanchard, G. Garçon, J-U. Mullot, and B. Jacquemin. 2017. "Effets sanitaires des particules de l'air ambiant selon leurs composés, sources et granulométrie." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Roth-Delgado, O. , C. Bladier, C. Beausoleil, E. Feinblatt, O. Merckel, and L. Watier. 2017. "Sensibilité des conclusions à la méthode qualitative d'évaluation du poids des preuves." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Saltelli, A. , M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. Chichester, UK: John Wiley & Sons, Ltd.

Saltelli, A. , S. Tarantola, F. Campolongo, and M. Ratto. 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Ltd, Chichester, UK: John Wiley & Sons, Ltd.

Saltelli, A., K. Chan, and E. M. Scott. 2000. *Sensitivity analysis*, *Wiley series in probability and statistics*. England: John Wiley & Sons Ltd. Original edition, 2000. Reprint, 2001.

Sissoko, F., I. Manière, S. Leconte, C. Rousselle, Anne Chevalier, Yuriko Iwatsubo, I. Guseva-Canu, D. Makowski, O. Bruyère, M. Touvier, and C. Bladier. 2017. "Elaboration d'une nouvelle grille d'évaluation de la qualité des études épidémiologiques - Application au cas du Cadmium." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Tassus, X., D. Makowski, P. Martin, S. Fraize-Frontier, C. Roth, and les membres du GT Ambroisie. 2017. "Élicitation probabiliste de connaissances d'experts pour l'évaluation de sources d'incertitude dans le cadre d'une analyse de risque phytosanitaire." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Thebault, A., J. Chiron, F. Etore, M. Sanaa, and C. Dunoyer. 2017. "Elicitation formalisée des connaissances d'experts : Méthode Sheffield." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Thebault, A., L. Guillier, D. Hervio-Heath, C. Bultel, and O. Cerf. 2017. "Évaluation du risque lié à Vibrio parahaemolyticus lors de la consommation de coquillages : prise en compte de l'évaluation de l'impact combiné des incertitudes." Séminaire scientifique sur l'évaluation du poids des preuves et l'analyse d'incertitude en évaluation des risques, Maisons-Alfort.

Tinker, J., J. Lowe, J. Holt, A. Pardaens, and A. Wiltshire. 2015. "Validation of an ensemble modelling system for climate projections for the northwest European shelf seas." *Progress in Oceanography* 138:211-237. doi: 10.1016/j.pocean.2015.07.002.

Tudur Smith, C., M. Marcucci, S. J. Nolan, A. Iorio, M. Sudell, R. Riley, M. M. Rovers, and P. R. Williamson. 2016. "Individual participant data meta-analyses compared with meta-analyses based on aggregate data." *Cochrane Database Syst Rev* 9:Mr000007. doi: 10.1002/14651858.MR000007.pub3.

ANNEXES

Annexe 1 : Liste des recommandations (résumées) et des posters associés

Thème 1 - Planification de l'évaluation concernant le poids des preuves et l'incertitude

R1.1 Formaliser le cadrage et définir les résultats attendus

R1.2 Formuler les questions et réaliser un schéma conceptuel

*Poster **1** : Schéma conceptuel relatif au processus d'introduction des salmonelles dans les filières de production des denrées alimentaires (responsable : Elissa KHAMISSE DER/UERSABA)*

R1.3 Développer la méthode d'évaluation

*Poster **2** : protocole de la méta-analyse des études de provocation aux champs électromagnétiques chez des sujets se déclarant électrohypersensibles (responsable : Sabrina HAVARD DER/UME)*

Thème 2 - Établissement des lignes de preuves

R2.1 Réaliser une revue systématique ou une revue approfondie de la littérature

*Poster **3** : revue systématique de la contamination du poisson par des microcystines (responsable : Nathalie ARNICH DER/UERALIM)*

R2.2 Utiliser des grilles pour évaluer la qualité des études par type d'étude

*Poster **4** : grille d'évaluation de la qualité des études épidémiologiques (responsable : Fatoumata SISSOKO DER/UESC)*

*Poster **5** : évaluation de la qualité des revues par la méthode R-Amstar (responsable : Amandine PAILLAT DER/UERA)*

R.2.3 Etablir les lignes de preuves en utilisant soit une approche qualitative précisant les critères utilisés soit une approche quantitative en privilégiant la méta-analyse sur les sujets stratégiques

*Poster **6** : établissement des lignes de preuve avec la méthode OHAT pour les effets sanitaires des particules de l'air ambiant (responsable : Matteo REDAELLI DER/UERA)*

Thème 3 - Identification et description des incertitudes

R3.1 identifier et décrire les incertitudes en suivant la typologie proposée dans le rapport

*Poster **7** : Analyse des incertitudes dans le cadre de l'évaluation de dossiers de produits réglementés (responsable : Jérémy DE SAINT JORES DEPR/UETPR)*

R3.2 Présenter le recensement par une matrice d'incertitudes

*Poster **8** : Elaboration d'une matrice d'incertitude – retour d'expérience sur la saisine « jouets » (responsables : Aurélie MATTHIEU DER/UESC et Carole LEROUX DER/UERA)*

R3.3 Sélectionner les sources d'incertitude qui seront évaluées et celles dont l'impact ne peut pas être évalué

Thème 4 : Intégration des lignes de preuve pour établir le poids des preuves

R4.1 Utiliser des modèles statistiques pour intégrer différentes lignes de preuve

*Poster **9** : établissement d'une grille de cotation pour les effets perturbateurs endocriniens (responsable : Claire BEAUSOLEIL DER/UESC)*

R4.2 Décrire et expliciter les choix pour l'intégration des lignes de preuves selon une méthode qualitative

*Poster **10** : Sensibilité des conclusions à la méthode qualitative d'évaluation du poids des preuves (responsable : Olivia ROTH-DELGADO DER/UERAP)*

Thème 5 - Evaluation et hiérarchisation des incertitudes

R5.1 Evaluer individuellement les incertitudes

*Poster **11** : élicitation formalisée des connaissances d'experts pour l'évaluation individuelle des incertitudes avec la méthode Sheffield (responsable : Anne THEBAULT DER/UME)*

*Poster **12** : élicitation des connaissances d'experts pour l'évaluation individuelle des incertitudes dans le cadre de l'analyse du risque phytosanitaire (responsable : Xavier TASSUS LSV/UERB)*

R5.2 Evaluer l'impact combiné des sources d'incertitude

*Poster **13** : prise en compte de l'évaluation de l'impact combiné des incertitudes dans l'évaluation du risque lié à Vibrio parahaemolyticus lors de la consommation de coquillages (responsable : Anne THEBAULT DER/UME)*

*Poster **14** : méthodes probabilistes pour évaluer l'impact combiné de l'incertitude lié au risque de réaction allergique à l'arachide (responsable : Amélie CREPET DER/UME)*

R5.3 Hiérarchiser les incertitudes

*Poster **15** : Analyse de sensibilité dans le cadre d'une appréciation quantitative des risques microbiologiques (responsable : Frédérique AUDIAT-PERRIN DER/UERALIM)*

Thème 6 - Expression des conclusions concernant le poids des preuves et l'incertitude

R6.1 : Exprimer dans un chapitre conclusif structuré la classification du poids des preuves en 4 niveaux

R6.2 : Exprimer dans un chapitre conclusif structuré l'incertitude (en termes de probabilité lorsque cela est possible) et présenter les conclusions dans un format spécifique (donner une matrice des incertitudes,…)

Annexe 2 : Principales demandes de clarification par thématique (séminaire du 20/01/2017)

Thématique 1 - Planification de l'évaluation (21 participants)

R1.1 Formaliser le cadrage et définir les résultats attendus

1) Comment définir le(s) schéma(s) conceptuel(s) ? Quelle forme pour le schéma ? Liste d'étapes,… ?

2) Est-ce que ce ne sont pas plutôt les objectifs plutôt que les « résultats » qui sont à définir *(cf. intitulé de la recommandation)* **?**

3) Périmètres en lien avec missions définies pour l'Anses ?

4) Qui participe (pétitionnaire, Anses, collectifs,…) ? Jusqu'ici, c'était les chercheurs de l'Anses qui réalisaient cette étape : est-ce que cela va changer ? Enjeux économiques : qui cernent les enjeux : experts ou coordinateurs ?

5) Consultation des parties prenantes pour le cadrage (ex. EATi, allaitement, PNNS,…) ?

6) Prise en compte des enjeux sociétaux ? Place des SHS ?

7) Si R1.1 aboutit à un contrat avec le décideur, est-il prévu de réviser R1.1 et ses attendus en cours d'expertise ?

8) Je ne comprends pas pourquoi le décideur ne s'exprime pas directement en PICO/PECO/PECOTS

9) Quelle(s) limite(s) aux résultats attendus (au-delà de la question de la saisine) ?

10) Où arrêter une revue préliminaire des données bibliographiques (temporalité, zone géographique) ?

11) PICO ? PECO ? PECOTS ?

12) Schéma conceptuel complet ? Je ne comprends pas le champ de la problématique : il manque un schéma large conceptuel pour « dompter » les enjeux sanitaires, socio-économiques, sociétaux *(voir domaine Santé environnement)*.

13) Comment compléter le cadrage actuel avec des éléments permettant d'établir un schéma conceptuel et de prendre en compte PICO, PECO(T) ? Comment intégrer les éléments relatifs au poids des preuves et à l'incertitude dans la première phase de cadrage ?

R1.2 Formuler les questions et réaliser un schéma conceptuel

1) Comment valide-t-on une question qui a été reformulée ? Cas d'une reformulation tardive, validation par les différentes parties prenantes (tutelles, Anses, experts) ?

2) Quelles sont les consignes/conseils pour élaborer le schéma conceptuel ? Est-ce que le schéma porte sur la méthode ou sur le sujet en lui-même ? Est-ce qu'il y a une phrase de conclusion pour un schéma conceptuel?

3) A quoi sert un schéma conceptuel ?

4) Selon quelle méthode ? Existe-t-il des méthodes de référence ?

R1.3 Développer la méthode d'évaluation

1) Comment choisir les bonnes méthodes ? En s'appuyant sur quoi (boite à outils) ?

2) Comment définir les ressources en fonction de la méthode retenue et des délais ?

3) Risque lié à la disponibilité des ressources (notamment documentaires). Comment le formaliser (dès le cadrage)?

Thème 2 - Établissement des lignes de preuves (23 participants)

R2.1 Réaliser une revue systématique ou une revue approfondie de la littérature

1) Sélection avant évaluation ? Déterminants de choix ? A quel stade ? Uniquement pour les dangers ?
2) Quand utiliser l'une ou l'autre ? Possibilité de faire approfondie puis systématique pour une sélection (contexte de la saisine) ? Signification approfondie ? Pourquoi pas exhaustive ?

R2.2 Utiliser des grilles pour évaluer la qualité des études par type d'étude

1) Est-ce que l'on exclut les études non sélectionnées ?
2) Ajouter « sélectionnées » dans le titre de la recommandation : « Utiliser des grilles pour évaluer la qualité des études sélectionnées par type d'étude » et préciser les types d'études épidémiologiques, *in vivo*, *in vitro*
3) Pourquoi parle-t-on de qualité et non de pertinence ?
4) Comment formaliser les grilles d'évaluation : utiliser quels critères ?
5) Comment la grille permet-elle d'évaluer la qualité ? *in vivo* et *in vitro* ?
6) Je ne comprends pas si les grilles s'adaptent à tous les types d'études ?
7) On parle d'études publiées : comment est pris en compte le biais de publication (résultats négatifs) ?
8) Je ne comprends pas que l'on puisse évacuer même en ayant des critères très stricts, le côté « humain » : avec la même grille, l'un va exclure et l'autre va garder.
9) Comment marche la méthode R-Amstar ? principe ? choix et validation des critères de sélection ? Application à des produits réglementés ? Quid des **BPL (bonnes pratiques de laboratoire)** ?

R.2.3 Etablir les lignes de preuves en utilisant soit une approche qualitative précisant les critères utilisés soit une approche quantitative en privilégiant la méta-analyse sur les sujets stratégiques

1) Traduction de « body of evidence »
2) Je ne comprends pas exactement ce que l'on entend par méta-analyse : est-ce que cela concerne les études où on poole les résultats et on refait des analyses statistiques ? est-ce que c'est une simple conclusion/synthèse des études de qualité ?
3) Les lignes de preuve peuvent-elles être indépendantes des critères de sélection ou des critères de la grille ?
4) Quel niveau de détail par exemple : effet sur un type donné de cellules ? effet sur un modèle animal ?
5) Combien de lignes de preuve pour une expertise en général ?
6) A quoi correspond en pratique une ligne de preuves ? En pratique, comment établir une ligne de preuves surtout si peu de publications ?

Thème 3 - Identification et description des incertitudes (20 participants)

R3.1 identifier et décrire les incertitudes en suivant la typologie proposée dans le rapport

1) Comment être exhaustif sur les incertitudes dans la partie cadrage dans le cas de saisines réglementaires ?
2) Pourquoi ne pas faire coïncider la typologie des sources d'incertitude avec la matrice (cf. études de cas) ?

R3.2 Présenter le recensement par une matrice d'incertitudes

1) Matrice très difficile à interpréter : faut-il formuler avec des phrases ? faut-il revoir les colonnes ? La question est donc : comment lire cette matrice (mode d'emploi en particulier pour les décideurs) ?
2) Comment garantir l'exhaustivité ?
3) Comment évaluer l'impact de l'incertitude a priori ?

R3.3 Sélectionner les sources d'incertitude qui seront évaluées et celles dont l'impact ne peut pas être évalué

1) Quelle est la question (« *sélectionner les sources d'incertitude qui seront évaluées et celles dont l'impact ne peut pas être évalué* » ou cf. rapport GT MER « *identifier les incertitudes pour lesquelles l'impact ne peut être évalué et indiquer les raisons de cette évaluation* ») ?
2) Selon quels critères (et quels moyens humains) peut-on identifier ces incertitudes dont les impacts ne peuvent être évalués ?
3) Quelle interprétation le gestionnaire peut-il faire de l'absence d'évaluation ? quelle interprétation le gestionnaire peut-il faire d'une incertitude qui serait énorme ?
4) L'évaluation de l'incertitude « a priori » ne doit-elle pas être considérée comme incertaine en soi ?

Thème 4 : Intégration des lignes de preuve (19 participants)

R4.1 Utiliser des modèles statistiques pour intégrer différentes lignes de preuve

1) **Est-on obligé d'utiliser des méthodes quantitatives ?** Reformuler la recommandation : « *il est possible* d'utiliser des *méthodes quantitatives*/des modèles statistiques » ? **Quels critères** pour faire du quantitatif ou du qualitatif ? Les méthodes quantitatives sont-elles considérées comme meilleures que les méthodes qualitatives ?
2) **Comment fait-on le lien entre poids des preuves et incertitudes ?** Est-ce que les résultats de l'évaluation du poids des preuves peuvent s'interpréter en termes d'incertitudes ? Possible problème d'interprétation en termes de probabilité ?
3) A partir de combien d'études par ligne de preuves peut-on mettre en œuvre une méthode d'intégration quantitative ?
4) Définition des poids : comment faire une pondération ?
5) Peut-on prendre en compte les conclusions d'articles de synthèse dans l'intégration des lignes de preuve ? Si oui, comment ?

R4.2 Décrire et expliciter les choix pour l'intégration des lignes de preuves selon une méthode qualitative

1) **Explicitation autour de la distinction de critères techniques ou scientifiques : ce qui relève des contraintes opérationnelles de contraintes de robustesse scientifique ?** Peut-on prendre en compte des paramètres d'ordre non scientifique mais plutôt « techniques » (temps, coût, compétences etc.) pour choisir telle ou telle méthode ?
2) Dans la description et l'explicitation, il faudrait définir une liste de critères scientifiques et techniques pour orienter le choix : **critères retenus pour établir les lignes de preuve ?**
3) Peut-on dissocier le choix de la méthode d'intégration des lignes de preuves du choix de la méthode d'établissement des lignes de preuves (si on a choisi une méthode donnée pour l'établissement des lignes de preuves, peut-on en choisir une autre pour l'étape d'intégration) ?
4) Peut-on ou doit-on dans certaines situations assortir le poids de la preuve d'une incidence de confiance/incertitude ?

Thème 5 - Evaluation et hiérarchisation des incertitudes (14 participants)

R5.1 Evaluer individuellement les incertitudes

1) Quel est l'objectif de l'évaluation individuelle ?
2) Quelle méthode appliquer en fonction de la nature de l'incertitude (et l'objectif de l'évaluation) ?

R5.2 Evaluer l'impact combiné des sources d'incertitude

1) Evaluation impact combiné des sources d'incertitude et application dans un contexte réglementaire
2) Evaluation individuelle hiérarchisation individuelle/évaluation combinée hiérarchisation combinée conditionnelle

R5.3 Hiérarchiser les incertitudes

1) Besoin de clarification entre incertitude et source d'incertitude (**glossaire**). Définir clairement « source d'incertitude » : est-ce la cause de l'incertitude s'appliquant à un paramètre, ou est-ce le paramètre lui-même ?
2) Proposer si possible une grille de hiérarchisation pour le qualitatif et préciser dans quel cas l'utiliser

Thème 6 - Expression des conclusions (17 participants)

R6.1 : Exprimer dans un chapitre conclusif structuré la classification du poids des preuves en 4 niveaux

1) Faut-il garder obligatoirement 4 niveaux ?
2) A quoi applique-t-on la classification des preuves selon les 4 niveaux : tout au long du processus ou sur le résultat ?
3) La définition des 4 niveaux doit-elle être homogène entre les avis de même type (groupe d'experts) ?
4) Quelle est la place de cette conclusion dans l'avis ?
5) Est-ce que la conclusion dans l'avis peut être réduite à une phrase ?
6) Faut-il établir une classification/formulation différente pour le gestionnaire et pour le grand public ?

R6.2 : Exprimer dans un chapitre conclusif structuré l'incertitude (en termes de probabilité lorsque cela est possible) et présenter les conclusions dans un format spécifique (donner une matrice des incertitudes,…)

1) Quelle traduction en termes d'impact(s) sur l'évaluation ? Sa crédibilité ?
2) Quelle(s) conséquence(s) potentielle(s) sur les décisions ?
3) Comment prioriser les évaluations pour lesquelles on fait l'exercice ?
4) De quel chapitre de conclusion parle-t-on (avis ou rapport) ?

Annexe 3 : Applicabilité des recommandations par thématique

Le nombre de participants aux six ateliers thématiques est variable. Chacun se positionne sur une seule thématique.

L'applicabilité des recommandations a été évaluée par les participants du séminaire comme indiquée ci-dessous :

Recommandation	*j'appliquerai cette recommandation, car…*	*j'appliquerai cette recommandation, si…*	*je ne pourrais pas appliquer cette recommandation, car…*
Thématique 1 : Planification de l'évaluation concernant le poids des preuves et l'incertitude (N*= 21)			
1.1 Formaliser le cadrage et définir les résultats attendus	1	8	2
1.2 Formuler les questions et réaliser un schéma conceptuel	7	5	0
1.3 Développer la méthode d'évaluation	1	11	1
Thématique 2 : Etablissement des lignes de preuve (N= 23)			
2.1 Réaliser une revue systématique ou une revue approfondie de la littérature	8	13	2
2.2 Utiliser des grilles pour évaluer la qualité des études par type d'étude	4	11	3
2.3 Etablir les lignes de preuves en utilisant soit une approche qualitative précisant les critères utilisés soit une approche quantitative en privilégiant la méta-analyse sur les sujets stratégiques	2	7	2
Thématique 3 : Identification et description des incertitudes (N= 20)			
3.1 identifier et décrire les incertitudes en suivant la typologie proposée dans le rapport	1	9	2
3.2 Présenter le recensement par une matrice d'incertitudes	4	4	3
3.3 Sélectionner les sources d'incertitude qui seront évaluées et celles dont l'impact ne peut pas être évalué	3	9	0
Thématique 4 : Intégration des lignes de preuve pour établir le poids des preuves (N= 19)			
4.1 Utiliser des modèles statistiques pour intégrer différentes lignes de preuve	3	12	5
4.2 Décrire et expliciter les choix pour l'intégration des lignes de preuves selon une méthode qualitative	6	7	2
Thématique 5 : Evaluation et hiérachisation des incertitudes (N= 14)			
5.1 Evaluer individuellement les incertitudes	7	10	1
5.2 Evaluer l'impact combiné des sources d'incertitude	3	6	1
5.3 Hiérarchiser les incertitudes	0	3	1
Thématique 6 : Expression des conclusions (N= 17)			
6.1 Exprimer dans un chapitre conclusif structuré la classification du poids des preuves en 4 niveaux	2	8	2
6.2 Exprimer dans un chapitre conclusif structuré l'incertitude et présenter les conclusions dans un format spécifique	5	16	1

*N : Nombre de participants.

Annexe 4 : Grille d'évaluation des études épidémiologiques pour l'élaboration de VLEP

Etude de cas PDP 1

La grille sera idéalement complétée par deux experts respectivement pour le volet "effet" et le volet "exposition". Il est laissé à l'appréciation de chaque expert les champs à remplir dans la grille selon leurs compétences. Cette grille fera l'objet d'une phase test qui pourra conduire à une version révisée.

									Analyse du risque de biais par l'approche OHAT
Logistique	Nom du relecteur								
	Date de lecture								
	Saisine associée								
Article	Réference de l'article (dont auteur, date publication, titre de l'étude, journal, DOI)	*Tous éléments permettant l'identification de l'étude afin de pouvoir s'y référer au besoin et de pourvoir la citer*							
	Journal								
	Objectif de l'étude	*Etude analytique, étude descriptive, etc.*							
		Objectif de l'étude et/ou hypothèses testées							
	Sources de Financement/lien d'intérêt potentiel	*Pas de conflit ; information disponible* *Renseigner les liens d'intérêt pour chaque auteur si information disponible*							
Design de l'étude	Type d'étude	*Cohorte, cas-témoins, étude transversale, méta-analyses, rapport de cas* *Caractéristique de l'étude : uni ou multicentrique etc.*							
	Date de l'étude								
	Type d'effets étudiés								
	Agent(s) d'exposition	*Caractérisation de l'agent : définition (nom, n°CAS, etc.),*							
Population d'étude et suivi	Pays								3,7
	Nombre de sujets								
	Description de la population ou des groupes (dont témoins)	*Age,*	*sexe ratio,*	*CSP,*	*ethnicité,*	*pays, effectif,*	*nombre de cas,*	*nom de la cohorte ou de l'étude en objet*	
	Sélection des individus => Population Profession (activité)	*Critères inclusion/ exclusion* *Mode de recrutement : volontariat, registres…*							
	Paramètres de suivi épidémiologique	*Médiane de la durée de suivi*					*Ou moyenne de la durée de suivi*		
		Date de début et fin de suivi					*Pourcentage de perdus de vue ou d'exclus*		
Effets ou pathologies étudiés	*Définition de l'/les effet(s) sanitaire(s) étudié(s)*								9
	Description de la mesure de l'effet sanitaire								
	Mode de recueil de l'effet sanitaire	*moment et lieux (CHU) ; nombre de personnes intervenantes, qualification des intervenants*							
	Référence bibliographique supplémentaire décrivant l'effet sanitaire								

Exposition : description générale	*Descriptif de l'exposition (profession, activité) ou environnementale*		8
	Voies d'exposition (*respiratoire, cutanée, orale)*		
	Utilisation d'EPI		
	Co-expositions possibles		
	Profil et fréquence d'exposition (exposition cumulée ; existence de pics d'exposition etc.)		
	Probabilité d'exposition[1]		
	Méthode d'évaluation quantitative, semi-quantitative ou qualitatives des expositions	*Prélèvement et analyse de l'air des lieux de travail, Biométrologie (urine, sang, air exhalé…), modélisation, matrice emploi-exposition, expertise individuelle des dossiers, déclaration des sujets ou des proches (téléphone, face-à-face, auto-questionnaire),*	
	Pertinence de l'évaluation de l'exposition par rapport à la voie principale de pénétration		
Exposition : Stratégie d'échantillonnage	Dimension de la campagne	*Internationale, nationale, régionale, d'un secteur d'activité*	
	Nombre et type de sujets exposés, GES (groupes d'exposition similaire)		
	Nombre de sujets ayant fait l'objet de mesures d'exposition		
	Nombre total de mesures		
	Type de prélèvement atmosphérique : (ambiance ou individuel,		
	Prélèvement biologique Moment de prélèvement (jour, heure, fin / début de poste)		
	Représentativité des mesures Adéquation entre le moment et la durée du prélèvement / cinétique du produit		
Exposition : Prélèvements et analyses des échantillons	Références des méthodes de prélèvement et d'analyse		
	Conditions de conservation et transport des échantillons		
	Nombre de laboratoires d'analyse		
	Préparation et traitement de l'échantillon		
	Techniques d'analyse		
	Limites de quantification, de détection		
	Contrôle de qualité (accréditation, CQE, CQI)		
Analyse statistique	Méthode d'analyse statistique	*Type de modèle (linéaire, logistique, Cox, etc.)*	

[1] Dans le cadre des matrices emplois-expositions, le terme probabilité d'exposition désigne la proportion de travailleurs exposés dans l'emploi concerné.

	Traitement des variables d'intérêt		
	Ajustement	*oui/non* *Facteurs de confusion sur quelles variables ?*	4
	Interactions testées	*Résultats bruts ou stratifiés ?*	
	Puissance	*Calcul a posteriori de la puissance ?*	
Résultats	Nombre < LQ ou LD		
	Détails des niveaux	*moyenne géométrique ou arithmétique, médiane, écart-type, intervalle de confiance, étendue – maximum, minimum*	
	Résultats / Force de l'association observée	*Estimation et intervalle de confiance, risques relatifs, Odds ratio… (avec références tableaux et figures)*	10
	Relation dose réponse	*Toute information en rapport à la relation dose réponse*	
	Autres résultats		
Discussion	Informations complémentaires	*Méthodes de traitement des données manquantes*	
	Autres éléments de la discussion	Surappariement, surajustement	
	Biais	*Mention de biais de sélection, de classement … Biais de confusion : mention, méthodes de prise en compte des facteurs*	
	Forces		
	Faiblesses (hors biais)		
Conclusions des auteurs	Relation dose réponse	*Toute information en rapport à la relation dose réponse*	
	Autres conclusion des auteurs		
Conclusions de l'expert	Risque de biais de confusion (cotation : ++, +, -, --)		
	Risque de biais dans l'évaluation de l'exposition (cotation : ++ + - --)		
	Risque de biais dans l'évaluation de l'effet (cotation : ++ + - --)		
	Palier selon la méthode OHAT		
	Références bibliographiques à récupérer		
	Intervention d'un autre expert pour une compétence particulière		
	Existence d'une relation dose réponse ?	oui/non	
	Etude à retenir pour l'expertise	oui/non	

Cotation des critères de biais :

++ = très faible probabilité d'un biais,

+ = biais peu probable,

- = biais possible,

-- biais très probable

Notes

Agence nationale de sécurité sanitaire
de l'alimentation, de l'environnement et du travail
14 rue Pierre et Marie Curie
94701 Maisons-Alfort Cedex
www.anses.fr / @Anses_fr

ISBN 979-10-286-0196-6 - Dépôt légal : novembre 2017 - © Anses Éditions : novembre 2017 - Date de publication : novembre 2017 - Couverture : Parimage (création), Anses (réalisation)