

anses

agence nationale de sécurité sanitaire
alimentation, environnement, travail



Connaître, évaluer, protéger

Traitement de l'incertitude dans le processus d'évaluation des risques sanitaires des substances chimiques

Rapport d'expertise collective

Mars 2016

Édition scientifique

anses

agence nationale de sécurité sanitaire
alimentation, environnement, travail



Connaître, évaluer, protéger

Traitement de l'incertitude dans le processus d'évaluation des risques sanitaires des substances chimiques

Rapport d'expertise collective

Mars 2016

Édition scientifique

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**

Saisine n°2009-SA-0331

RAPPORT d'expertise collective

**Groupe de travail
« Perturbateurs Endocriniens »**

Mars 2016

Mots clés

Incertitude, variabilité, sources d'incertitude, matrice des incertitudes, analyse de sensibilité, poids de la preuve.

Présentation des intervenants

PRÉAMBULE : Les experts externes, 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.

COMITE DE TRAVAIL « INCERTITUDE »

Président

M. Luc BELZUNCES – Directeur de Recherche - Laboratoire de Toxicologie Environnementale, UR 406 A&E, INRA

Membres

Mme. Elisabeth ELEFANT – Médecin spécialisé en tératologie humaine - Centre de référence sur les Agents tératogènes hôpital Trousseau APHP Paris

Mme Laura MAXIM – Chargée de recherche – CNRS

M. Alexandre PERY – Directeur de l'Ecole Doctorale ABIES – AgroParisTech

Anses

Mme Claire BEAUSOLEIL – Chef de projet scientifique, Unité « Évaluation des dangers et des risques des substances » – Anses

M. Sébastien DENYS – Chef d'unité, Unité « Méthodologie et études » – Anses

Mme Valérie PERNELET-JOLY – Chef d'unité, Unité « Évaluation des risques liés à l'air » – Anses

M. Christophe ROUSSELLE – Chef d'unité, Unité « Évaluation des dangers et des risques des substances » – Anses

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

GRUPE DE TRAVAIL « PERTURBATEURS ENDOCRINIENS »

Président

M. Claude EMOND – Université de Montréal, Canada

Vice-président

M. Jean-Pierre CRAVEDI - Directeur de Recherche - INRA

Membres

M. Jean-Philippe ANTIGNAC - Ingénieur analyste - ONIRIS, LABERCA

Mme Martine APPLANAT-Directeur de Recherche – INSERM.

M. Brice APPENZELLER - Responsable de laboratoire de biomonitoring - Centre de Recherche Public en Santé, Luxembourg

M. Rémy BEAUDOUIN-Chargé de Recherche - INERIS.

M. Luc BELZUNCES – Directeur de Recherche – Laboratoire de Toxicologie Environnementale, UR 406 A&E, INRA

Mme Marie-Chantal CANIVENC-LAVIER-Chargé de Recherche -INRA.

M. Nicolas CHEVALIER-Médecin endocrinologue-Praticien hospitalier- CHU de Nice.

Mme Cécile CHEVRIER –Chargé de Recherche -INSERM.
Mme Martine CLAUW - Toxicologue-vétérinaire - INPT/ENVT, Université de Toulouse
Mme Elisabeth ELEFANT - Médecin spécialisé en tératologie humaine - Centre de référence sur les Agents tératogènes - AP-HP hôpital Armand Trousseau, Paris
Mme Florence EUSTACHE - Médecin - CECOS, AP-HP, Hôpital Jean Verdier, Paris
M. René HABERT - Professeur des universités - Université Paris Diderot
Mme Brigitte LE MAGUERESSE-BATTISTONI - Directeur de Recherche – INSERM
Mme Sakina MHAOUTY- KODJA - Directeur de Recherche – CNRS.
M. Christophe MINIER - Ecotoxicologue - Université du Havre
M. Luc MULTIGNER - Médecin épidémiologiste – INSERM
M. Henri SCHROEDER – Enseignant chercheur à l'URAFPA, INRA USC 340.
M. Patrick THONNEAU - Médecin - INSERM
Mme Catherine VIGUIE – Vétérinaire – Directrice de Recherche INRA

COMITÉ D'EXPERTS SPÉCIALISÉ « SUBSTANCES »

Les travaux, objets du présent rapport ont été présentés et adoptés par le Comité d'experts spécialisé « Substances » :

Président

M. Michel GUERBET – Professeur de toxicologie à l'UFR médecine pharmacie de Rouen - Pharmacien toxicologue

Vice-président

M. Dominique LAFON – Médecin toxicologue chez Nexter Group – Médecine du travail, toxicologie, reprotoxicité

Membres

M. Marc BARIL - Professeur associé à l'Université de Montréal – Chimiste toxicologue, VLEP

M. Sylvain BILLET – Enseignant chercheur / maître de conférence en toxicologie à l'Université du Littoral Côte d'Opale – Toxicologie respiratoire, nanomatériaux

Mme Michèle BISSON – Responsable d'étude à l'INERIS – Pharmacien toxicologue, toxicologie générale - VTR

Mme Anne CHEVALIER – Epidémiologiste retraitée de l'Institut de Veille Sanitaire

M. François CLINARD – Epidémiologiste à l'Institut de Veille Sanitaire – Pharmacien toxicologue, épidémiologie, évaluation des risques sanitaires

Mme Fatiha EL-GHISSASSI – Scientifique, Section des Monographies de IARC (IMO) Centre International de Recherche sur le Cancer - Docteur es science en biochimie spécialiste en cancérogénèse et génotoxicité

Mme Mounia EL-YAMANI – Responsable d'unité à l'Institut de Veille sanitaire – Docteur es science en biochimie, toxicologie, VLEP

M. Claude EMOND – Professeur adjoint de clinique à l'Université de Montréal – Toxicologie, modèle PBPK, toxicocinétique, nanotoxicologie, perturbateurs endocriniens

M. Guillaume GARCON – Professeur de toxicologie à l'Université de Lille 2 – Toxicologie générale, cancérologie, modèles expérimentaux, toxicologie respiratoire, pollution atmosphérique

M. Ludovic LE HEGARAT – Chef d'unité adjoint Toxicologie des contaminants - Anses – Laboratoire de Fougères- Toxicologie, génotoxicité, nanomatériaux

M. Karim MAGHNI – Professeur sous octroi agrégé à l'Université de Montréal – Toxicologie, immunologie, asthme, allergies, nanomatériaux

Mme Véronique MALARD – Ingénieur chercheur en toxicologie au Commissariat à l'énergie atomique et aux énergies alternatives, Centre de Marcoule. – Toxicologie « *in vitro* », biologie cellulaire, nanotoxicologie, protéomique.

M. Fabrice MICHIELS – Médecin du travail / toxicologue à l'Association Interentreprises pour la Santé au Travail 19

M. Jean-Paul PAYAN – Chef du laboratoire Pénétration Cutanée, Cinétique et Métabolisme à l'INRS, Nancy – Pharmacien toxicologue, toxicocinétique

M. Henri SCHROEDER – Enseignant chercheur à l'URAFPA, INRA USC 340, Faculté des Sciences et Technologies, Université de Lorraine - Pharmacien biologiste - Neurotoxicité, comportement animal, développement cérébral, exposition périnatale

M. Alain SIMONNARD – Chef de département à l'INRS, Nancy - Pharmacien toxicologue, toxicologie générale et reprotoxicité, anatomopathologie

M. Olivier SORG – Chef de groupe de recherche à l'Université de Genève – Docteur es science en biochimie, toxicologie expérimentale, dermatotoxicologie

Mme Lydie SPARFEL – Professeur à l'Université de Rennes 1 / IRSET 'Institut de Recherche en Santé, Environnement et Travail' UMR INSERM 1085 – Pharmacien Toxicologue, immunotoxicologie, toxicogénomique, cancérologie, biologie cellulaire et moléculaire

M. Jérôme THIREAU – Chargé de recherche au CNRS – Docteur es science, physiologie animale, biologie cellulaire, cardiotoxicité

PARTICIPATION ANSES

Coordination scientifique

Mme Sandrine FRAIZE-FRONTIER – Chef de projet scientifique, Unité « Méthodologie et études » – Anses

Contribution scientifique

Mme Claire BEAUSOLEIL – Chef de projet scientifique, Unité « Évaluation des dangers et des risques des substances » – Anses

M. Sébastien DENYS – Chef d'unité, Unité « Méthodologie et études » – Anses

Mme Sandrine FRAIZE-FRONTIER – Chef de projet scientifique, Unité « Méthodologie et études » – Anses

Mme Valérie PERNELET-JOLY – Chef d'unité, Unité « Évaluation des risques liés à l'air » – Anses

M. Christophe ROUSSELLE – Chef d'unité, Unité « Évaluation des dangers et des risques des substances » – Anses

Secrétariat administratif

Mme Séverine BOIX – Anses

SOMMAIRE

Présentation des intervenants	3
Résumé.....	8
Sigles et abréviations	9
Glossaire	10
Liste des tableaux.....	13
Liste des figures	13
Préambule	14
1 Contexte, objet et modalités de traitement de la saisine.....	15
1.1 Contexte.....	15
1.2 Objet de la saisine.....	16
1.3 Modalités de traitement : moyens mis en œuvre et organisation.....	16
1.4 Prévention des risques de conflits d'intérêts.	16
2 Démarche de travail adoptée par le CT Incertitude.....	17
3 Incertitudes : Définitions et concepts	18
3.1 Incertitude <i>versus</i> Variabilité.....	18
3.2 L'incertitude : un concept multidimensionnel	19
3.3 Classification des sources d'incertitude	21
4 Approche par paliers	25
4.1 Description des étapes	26
4.1.1 Niveau 0 – <i>Screening</i> – Estimation ponctuelle sur la base d'hypothèses conservatoires, en utilisant le cas échéant des valeurs par défaut.....	26
4.1.2 Niveau 1 – Qualitatif – Identification systématique et caractérisation des sources d'incertitude	26
4.1.3 Niveau 2 – Déterministe – Approche déterministe de l'analyse de sensibilité	26
4.1.4 Niveau 3 – Probabiliste – Approche probabiliste de l'analyse de sensibilité et d'incertitude	27
4.1.4.1 Analyse probabiliste de l'incertitude	28
4.1.4.2 Analyse de sensibilité.....	30
4.2 Niveau d'analyse nécessaire à l'analyse d'incertitude	30
5 Revue critique des principales méthodes actuelles	32
5.1 Analyse d'incertitude dans l'évaluation de l'exposition – Démarche adoptée lors de l'ERS du BPA.....	32
5.1.1 Analyse qualitative de l'incertitude dans le calcul de l'exposition au BPA	32
5.1.1.1 Présentation de la démarche.....	32
5.1.1.2 Limites.....	35
5.1.2 Prise en compte quantitative de l'incertitude dans la caractérisation de l'exposition	36
5.1.2.1 Présentation de la démarche.....	36
5.1.2.2 Limites.....	37
5.2 Analyse d'incertitude dans l'évaluation du danger	37
5.2.1 Analyse des données de la littérature pour l'évaluation du danger – Poids de la preuve & Incertitude	38

5.2.2	Prise en compte quantitative de l'incertitude dans la caractérisation du danger.....	42
5.2.2.1	Principe général	43
5.2.2.2	Démarche proposée.....	44
5.2.2.3	Limites de la démarche	45
6	Développements méthodologiques pour le traitement de l'incertitude en ERS et applications	47
6.1	Démarche d'analyse des données pour l'identification du danger - Poids de la preuve et Analyse d'incertitude	47
6.2	Prise en compte quantitative de l'incertitude dans la caractérisation de relations dose-réponse.....	50
6.3	Démarche d'analyse qualitative de l'incertitude dans l'évaluation de l'exposition et son application au toluène	51
6.3.1	Étape 1 – Recensement des sources d'incertitude & classification utilisée	52
6.3.2	Étape 2 – Caractérisation systématique des sources d'incertitude.....	52
6.3.3	Étape 3 – Évaluation de l'influence des principales sources d'incertitude sur l'évaluation de l'exposition	53
6.3.4	Étape 4 – Détermination d'un niveau de confiance global	55
6.3.5	Discussion sur la démarche proposée.....	55
6.4	Proposition pour la prise en compte quantitative de l'incertitude dans le calcul de l'exposition	56
7	Recommandations.....	57
8	Conclusions	59
9	Bibliographie.....	60
	ANNEXES	66
	Annexe 1 : Caractérisation de l'incertitude dans le calcul de l'exposition – Glossaire	67
	Annexe 2 : Analyse qualitative de l'incertitude dans le calcul de l'exposition au toluène – Résultats.....	75
	Annexe 3 : Expert Profile.....	82

Résumé

Le présent rapport d'expertise collective se fonde sur une revue critique des principales démarches faisant référence dans le domaine de l'évaluation des risques.

Il propose, pour chacune des étapes de l'évaluation des risques, une démarche qualitative et/ou quantitative pour la prise en compte de l'incertitude, adaptée au contexte d'expertise collective en vigueur à l'Anses.

Pour prendre en compte l'incertitude dans l'étape d'identification du danger, il est proposé une démarche d'évaluation du poids des preuves structurée en 7 étapes :

- (1) (re)formulation de la (ou des) question(s) posée(s),
- (2) conduite de la recherche bibliographique,
- (3) analyse de la recevabilité des études et choix des critères d'inclusion (et par conséquent d'exclusion),
- (4) recueil et analyse des données issues des études prises de façon individuelle,
- (5) analyse globale de l'ensemble des données disponibles,
- (6) détermination du niveau de confiance associé à la base de données disponible, et
- (7) conclusion sur l'ensemble des niveaux de preuves.

Pour la prise en compte de l'incertitude dans l'étape de caractérisation de relations doses-réponses, il est proposé d'utiliser la démarche recommandée par l'IPCS (2014). Les points forts de cette démarche sont notamment d'identifier et de caractériser par des distributions statistiques les différentes sources d'incertitude pour proposer un calcul de l'incertitude globale.

Pour prendre en compte l'incertitude dans l'étape d'évaluation de l'exposition, il est proposé une approche par paliers afin proportionner le niveau d'analyse aux besoins de l'évaluation comprenant une approche qualitative systématique, à compléter, en fonction des besoins de l'expertise, par une approche déterministe et/ou probabiliste. L'approche qualitative proposée est structurée en quatre étapes :

- (1) Recensement des sources d'incertitude. Le recensement doit s'effectuer à l'aide d'une classification des sources d'incertitude structurée selon leur localisation dans le processus d'évaluation, en quatre classes (*sources d'incertitude contextuelle, sources d'incertitude des scénarios, sources d'incertitude des modèles et sources d'incertitude des données*)
- (2) Caractérisation systématique des sources d'incertitude. En particulier, la nature et le traitement dans le processus d'expertise de chacune des sources d'incertitude recensées doivent être rapportés.
- (3) Évaluation de l'influence des principales sources d'incertitude sur le résultat de l'évaluation de l'exposition. Cette étape se décline en trois sous étapes :
 - (3.1) Sélection des sources d'incertitude à prendre en compte dans la suite de l'analyse
 - (3.2) Caractérisation de l'influence des incertitudes en déterminant l'amplitude et la direction de l'influence de chaque source d'incertitude sélectionnée sur le résultat de l'évaluation de l'exposition
 - (3.3) Analyse critique des résultats
- (4) Détermination du niveau de confiance global des résultats

Comme pour l'étape de caractérisation de relation doses-réponses, l'approche quantitative proposée pour la prise en compte de l'incertitude dans l'évaluation de l'exposition est centrée sur une approche probabiliste. Elle traite de l'incertitude des facteurs d'entrée des modèles, et inclut de manière systématique une analyse de sensibilité en vue de classer les facteurs d'entrée du modèle considéré suivant l'importance de leur influence sur le résultat et de leur niveau d'incertitude.

Sigles et abréviations

AHRQ	<i>Agency for Healthcare Research and Quality</i>
ANSES	Agence Nationale de Sécurité Sanitaire de l'alimentation, de l'environnement et du travail
BMD	Benchmark Dose
BMDL	<i>Benchmark Dose Lower confidence limit</i>
BMR	<i>Benchmark Response</i>
BPA	Bisphénol A
EFH	<i>Exposure Factors Handbook</i>
ERS	Évaluation des risques sanitaires
EPHECT	<i>Project "Emission, Exposure Patterns and Health Effects of Consumer Products in EU"</i>
CT	Comité de travail
DJA	Dose journalière admissible
ECHA	<i>European Chemicals Agency</i>
EFSA	<i>European Food Safety Authority</i>
GRADE	<i>Grading of Recommendations Assessment, Development, and Evaluation</i>
GT PE	Groupe de travail « Perturbateurs Endocriniens »
IPCC	<i>Intergovernmental Panel on Climate Change</i>
IPCS	<i>International Programme on Chemical Safety</i>
JRC-EC	<i>Joint Research Center – European Commission</i>
MOA	<i>Mode Of Action</i>
NTP	<i>National Toxicology Program</i>
OCDE	Organisation de Coopération et de Développement Économiques
OMS	Organisation Mondiale de la Santé
PBTK	<i>Physiologically Based Toxicokinetics models</i>
PE	Perturbateur Endocrinien
PNRPE	Programme national de recherche sur les perturbateurs endocriniens
LOAEL	<i>Lowest-Observed-Adverse-Effect Level</i>
NOAEL	<i>No Observable Adverse Effect Level</i>
SCENIHR	<i>Scientific Committee on Emerging and Newly Identified Health Risks</i>
US EPA	<i>United States Environmental Protection Agency</i>
VTR	Valeur toxicologique de référence
WOE	<i>Weight of evidence</i>

Glossaire

Amplitude d'une source d'incertitude : en évaluation des risques, l'amplitude (*magnitude* en anglais) d'une source d'incertitude indique le degré d'incertitude qu'elle induit sur le résultat final (IPCC 2005).

Approche par paliers : une approche par paliers (*tiered approach* en anglais) se réfère à un processus progressif évoluant en complexité, de manière graduée, au fur et à mesure des étapes ou niveaux (IPCS 2008). Elle est recommandée par la majorité des guides méthodologiques qui traitent d'incertitude en ERS, car elle permet de proportionner l'effort d'analyse au besoin du problème.

Benchmark dose (BMD) : estimation ponctuelle de la dose correspondant à un faible niveau donné de risque (Gaylor et al. 1998).

Bootstrapping : technique d'inférence statistique basée sur une succession de ré-échantillonnage (Efron and Tibshirani 1993).

Classification des sources d'incertitude : classification définie comme une liste générique de sources d'incertitude. Cette liste est structurée selon la localisation des sources d'incertitude dans le processus d'expertise.

Dimension d'une source d'incertitude : les dimensions d'une source d'incertitude sont les aspects qui la caractérisent, par exemple, sa localisation dans le processus d'évaluation, sa nature (« variabilité » ou « limites de la connaissance »), son niveau sur le spectre des incertitudes, son traitement dans le processus d'évaluation, etc. (Walker et al. 2003).

Direction d'une source d'incertitude : en évaluation des risques, la direction d'une source d'incertitude indique le sens de son influence sur le résultat final (IPCS 2008). La direction d'une source d'incertitude est souvent exprimée en utilisant les symboles « + » et « - » ou les flèches « ↓ » et « ↑ » :

- « + » ou « ↑ » indique que le risque pourrait être surestimé en raison de l'incertitude,
- « - » ou « ↓ » indique que le risque pourrait être sous-estimé en raison de l'incertitude, et
- « +/- » ou « ↑↓ » indique que le risque a autant de chance d'être surestimé que sous-estimé en raison de l'incertitude.

Données de substitution : les données de substitution sont celles utilisées en l'absence de donnée spécifique à la situation d'étude.

Élément de preuve : en évaluation des risques, un élément de preuve (*line of evidence* en anglais) est un ensemble d'informations disponibles, idéalement publiées, issues de différentes études *in vivo*, *in vitro*, etc., sur lesquelles le processus d'évaluation fonde ses conclusions sur l'existence ou non d'une relation de cause à effet.

Erreur aléatoire de mesure : les erreurs aléatoires de mesure sont à rattacher à la notion de précision. Elles induisent une variation autour de la « vraie » valeur du paramètre considéré et sont souvent aggravées par la variabilité réelle de la quantité considérée (Zheng and Frey 2005).

Erreur systématique : les erreurs systématiques ou biais résultent d'imperfections du dispositif de mesure (exemple : calibrations incorrectes d'appareils de mesures) ou de recueil des données. La moyenne des mesures réalisées ou des données recueillies peut alors différer de la valeur réelle.

Extrapolation : extrapoler consiste à inférer à partir de données ou d'informations fragmentaires ou incomplètes (par exemple, extrapolation de l'animal à l'être humain, du passé au futur, de la voie orale à la voie inhalée, etc.).

Évaluation du poids de la preuve : l'évaluation du poids de la preuve (*Weight of evidence approach* en anglais) est une démarche d'appréciation structurée et systématique, dont l'objectif est d'explicitier les forces et les faiblesses d'éléments de preuve (SCENIHR 2012). Différentes démarches existent, elles comprennent généralement les étapes suivantes :

Étape 0 – Planification du processus d'évaluation du poids de la preuve

Étape 1 – Identification des sources possibles des données et du manque de données en relation avec l'objectif de l'ERS

Étape 2 – « *Screening* » initial de ces sources de données pour identifier celles qui sont pertinentes pour la question posée

Étape 3 – Appréciation individuelle des études

Étape 4 – Appréciation individuelle des éléments de preuve

Étape 5 – Appréciation globale des éléments de preuve

Étape 6 – Caractérisation des incertitudes d'intérêt

Incertain : l'incertitude fait référence à une situation où le corpus de connaissances disponibles sur un sujet particulier (exemple : les effets toxiques des PE) est perçu comme problématique – à savoir, peu connu ou inconnu, incomplet, incorrect, biaisé, pas assez convaincant, etc. – par divers individus (ANSES 2013b).

Localisation d'une source d'incertitude : en évaluation des risques, la localisation d'une source d'incertitude est là où elle se manifeste dans le processus d'expertise, depuis la définition du périmètre de l'évaluation, jusqu'à la communication des résultats, en passant par l'identification et la caractérisation du danger, le calcul de l'exposition et du risque (Walker et al. 2003).

Matrice des incertitudes : en évaluation des risques, une matrice des incertitudes est un tableau servant à recenser les sources d'incertitude et à rapporter leurs caractéristiques (van der Sluijs et al. 2003). Les lignes du tableau listent toutes les sources d'incertitudes recensées. Cette liste est structurée selon la localisation des sources d'incertitude dans le processus d'expertise. Les colonnes de la matrice présentent les autres dimensions utilisées pour caractériser chaque source d'incertitude listée.

Modèle mathématique : représentation mathématique d'un phénomène ou d'un système d'intérêt. Un modèle mathématique a pour but de traduire de manière opérationnelle l'ensemble des connaissances à des fins de déduction. Par traduction opérationnelle, on entend la possibilité de régénérer l'ensemble des résultats possibles conditionnellement à un nombre restreint de paramètres (Parent and Bernier 2007). Un modèle répond à un objectif, simplifie la réalité et donne une représentation quantifiée du problème en main (Bernier, Parent, and Boreux 2000).

Nature d'une source d'incertitude : la nature d'une source d'incertitude indique son origine ou sa cause (van Asselt and Rotmans 2002).

Niveau d'une source d'incertitude : le niveau d'une source d'incertitude est là où elle se situe sur le spectre des incertitudes (Walker et al. 2003).

Paramètre d'un modèle : le terme « paramètre » fait référence à une quantité utilisée pour calibrer ou définir un modèle mathématique (exemples : la pente et l'ordonnée à l'origine d'un modèle de régression univarié ; la moyenne et la variance d'une loi normale) (Cullen and Frey 1999). La valeur d'un paramètre d'un modèle est habituellement déterminée par ajustement sur un jeu de données de calibration. Le terme « paramètre » est à distinguer du terme « variable d'entrée » d'un modèle.

Qualification de la base de connaissances (*qualification of the knowledge base* en anglais) : en évaluation des risques, la qualification de la base de connaissances donne une appréciation de la fiabilité des connaissances sur lesquelles se fonde l'ERS (van der Sluijs et al. 2003).

Source d'incertitude : en évaluation des risques, une source d'incertitude est un élément du processus d'évaluation qui induit de l'incertitude dans le résultat final. À un niveau d'agrégation élevé, il y a deux grandes sources d'incertitude, la variabilité et les limites de la connaissance (van Asselt and Rotmans 2002).

Spectre des incertitudes : ensemble des formes possibles d'incertitude. Le spectre des incertitudes en ERS comprend les différents niveaux ci-après (van Asselt and Rotmans 2002) :

- l'**inexactitude**, synonyme de manque de précision ou d'imprécision, fait référence à une incertitude métrique (exemple : les erreurs de mesure)
- le **manque d'observations** ou de mesures fait référence à un manque de données pouvant théoriquement être recueillies mais qui ne l'ont pas été (exemple : nombre insuffisant de sujets d'une étude).
- le **caractère non mesurable** dans la pratique fait référence à un manque de données pouvant être recueillies en principe mais infaisable expérimentalement ou dans le temps imparti.
- la **contradiction des observations** fait référence à l'existence de différents jeux de données ou observations disponibles permettant des interprétations contradictoires.
- l'**ignorance réductible** fait référence à un processus qu'on n'observe pas, qu'on n'imagine pas théoriquement à un temps donné mais peut-être dans le futur.
- l'**indétermination** fait référence à un processus dont les principes et les lois sont compris mais qui ne peut être complètement prédit et déterminé.
- l'**ignorance irréductible** fait référence à des processus et des interactions entre processus qui ne peuvent pas être déterminé par les capacités humaines.

Traitement d'une source d'incertitude : le traitement d'une source d'incertitude indique la manière dont elle a été prise en compte dans le processus d'évaluation, sous la forme d'un argumentaire justifiant le choix d'une hypothèse, de la gestion de données censurées, etc. (ANSES 2013a).

Valeur attachée aux choix (*value-ladenness of choices* ou *value loadings of choices* en anglais) : terme faisant référence à l'importance, parfois subjective, attachée aux choix réalisés au cours de l'évaluation des risques. Elle se réfère à la présence de valeur ou de biais sous-jacents aux choix réalisés (choix concernant la formulation de la question, la sélection ou non des données, leurs interprétations, etc.) (van der Sluijs et al. 2003).

Valeur par défaut : Les valeurs par défaut sont des valeurs de référence recommandées dans des situations précises. Le niveau d'incertitude associé à l'utilisation de valeurs par défaut dépend de la procédure et de la qualité des données sur lesquelles elles se fondent.

Variable d'entrée d'un modèle : le terme « variable d'entrée » fait référence à une valeur utilisée pour alimenter un modèle mathématique calibré ou supposé l'être. Le terme « variable d'entrée » est à distinguer du terme « paramètre » d'un modèle.

Variabilité : la variabilité est définie comme l'étendue des valeurs réelles ; elle exprime l'hétérogénéité ou la diversité de la réalité et présente, en général, la propriété d'être non réductible (IPCS 2008).

Liste des tableaux

Tableau 1 : Exemple de structure matricielle utilisée pour recenser et caractériser les sources d'incertitude en ERS.....	21
Tableau 2 : Classification des sources d'incertitude relatives à l'évaluation de l'exposition, recommandée par l'IPCS – Traduit de l'IPCS (2008)	24
Tableau 3: Classification des sources d'incertitude relatives au contexte utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA	33
Tableau 4 : Classification des sources d'incertitude relatives aux scénarios d'exposition utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA	33
Tableau 5 : Classification des sources d'incertitude relatives aux modèles d'exposition utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA	33
Tableau 6 : Classification des sources d'incertitude relatives aux données d'entrée des modèles d'exposition utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA	34
Tableau 7 : Structure matricielle utilisée par le GT PE pour recenser et caractériser les sources d'incertitude relatives à l'évaluation de l'exposition au BPA (ANSES 2013a).....	35
Tableau 8 : Niveau de confiance quantitativement calibré – Traduit de (IPCC 2005)	42
Tableau 9 : Échelle de vraisemblance – Traduit de (IPCC 2005)	42
Tableau 10 : Classification des sources d'incertitude relatives à l'évaluation de l'exposition à une substance chimique, en vue d'une ERS, structurée selon leur localisation dans le processus d'expertise	52
Tableau 11 : Échelle de jugement utilisée pour déterminer l'amplitude de l'impact d'une source d'incertitude sur l'évaluation de l'exposition au toluène	53
Tableau 12 : Représentation schématique de l'analyse quantitative de l'incertitude. Dans l'exemple, le paramètre 3, avec une incertitude élevée et un impact fort sur les prédictions devrait faire l'objet d'études complémentaires.	56

Liste des figures

Figure 1 : Typologie des sources d'incertitude – Traduit de M.B.A. van Asselt and J. Rotman, Uncertainty in integrated assessment modelling. From Positivism to Pluralism, <i>Climatic Change</i> , 54: 75-105, 2002. La variabilité et une connaissance limitée sont présentées comme les deux grandes sources d'incertitude contribuant au large spectre des incertitudes allant de l'ignorance irréductible à l'inexactitude.	19
Figure 2 : Illustration graphique des différents niveaux d'analyse d'incertitude, allant progressivement du niveau le moins élevé (Niveau 0) au niveau le plus élevé (Niveau 3). Les ronds noirs représentent des estimations ponctuelles de l'exposition et les lignes en pointillées indiquent l'étendue de l'exposition – Adapté de l'EFSA (2007).	25
Figure 3 : Arbre décisionnel pour l'attribution d'un niveau de confiance aux corpus de données des différents niveaux de preuves	49
Figure 4 : Démarche proposée par le CT "Incertain" pour l'analyse qualitative des sources d'incertitude en évaluation de l'exposition aux substances chimiques, à l'Anses	51
Figure 5: Incertitudes contextuelles et dans les scénarios.....	54

Préambule

Le présent document rapporte un premier travail de réflexion sur la caractérisation et la prise en compte de l'incertitude en évaluation des risques sanitaires (ERS), au sein de l'Anses. Il se place dans la perspective du développement d'une démarche standard, adaptée aux procédures d'expertises collectives en vigueur à l'Agence.

Le travail a été réalisé dans le cadre de la saisine « Perturbateurs endocriniens et reprotoxiques de catégorie 2 » (saisine PE), par un comité de travail dédié (CT Incertitude). Il fait suite au traitement de l'incertitude réalisé lors de l'ERS du bisphénol A (BPA), par le groupe de travail « Perturbateurs endocriniens » (GT PE).

Le document s'intéresse à l'intégralité du processus d'ERS, et considère l'ensemble des sources d'incertitudes pouvant se présenter lors d'une ERS de substances chimiques plus spécifiquement.

Ce document amorce un recensement et une analyse critique des démarches et outils, qualitatifs et quantitatifs, faisant référence actuellement dans le domaine des risques chimiques. Il revient sur le traitement de l'incertitude réalisé lors de l'ERS du BPA, et propose une première démarche standardisée, en vue d'une application dans la suite de la saisine PE et plus largement dans les ERS des substances chimiques conduites au sein de la Direction de l'Évaluation des Risques (DER).

Ce document définit les éléments structurant d'une démarche d'analyse d'incertitude en ERS. Il dresse les différentes sources d'incertitude rencontrées en ERS et identifie leurs dimensions. Il souligne les avantages d'une démarche par paliers mettant en jeu aussi bien des approches qualitatives que quantitatives – déterministes comme probabilistes – pour les analyser. Enfin, pour chaque étape de l'ERS, le document propose une démarche d'analyse qualitative et une autre quantitative, adaptées aux pratiques d'expertise collective en place à l'Anses. Le document recouvre, de cette façon, l'ensemble du processus d'ERS. Les recommandations formulées dans le document visent à initier, à court terme, un déploiement des démarches proposées.

Le document souligne que le travail réalisé ne vise pas à répondre à toutes les questions qui se posent en matière d'analyse d'incertitude, mais s'inscrit plutôt dans un processus dynamique d'évolution méthodologique, d'acquisition de données et de communication, pour informer au mieux les parties prenantes, et optimiser la prise de décision en matière de gestion des risques sanitaires.

1 Contexte, objet et modalités de traitement de la saisine

1.1 Contexte

L'attention accordée à la question de l'incertitude en évaluation des risques sanitaires (ERS) et à sa prise en compte dans le processus d'expertise a pris, ces dernières années, une réelle ampleur (Verdonck et al. 2007). Le nombre croissant de publications scientifiques, la multiplication d'approches et les nombreux guides méthodologiques qui s'y intéressent sont là pour en témoigner (ECHA 2012, EFSA 2007, IPCS 2008, SCENIHR 2012).

Il existe plusieurs définitions de l'incertitude. Celle de l'IPCS (2004) fait référence à « une connaissance imparfaite concernant l'état présent ou futur d'un organisme, d'un système ou d'une (sous) population considérés » (IPCS 2004). La définition donnée par l'Anses, dans le cadre de ses travaux d'expertise sur les PE, va au-delà d'un manque de connaissance. Elle inclut le potentiel d'erreur ou de biais, les insuffisances techniques des instruments de recherche, les insuffisances méthodologiques des protocoles de recherche, les choix par jugement d'expert des hypothèses qui remplacent les connaissances scientifiques manquantes, les divergences d'interprétation des données brutes disponibles et le caractère incomplet, incorrect ou ambigu de la communication des résultats. » (ANSES 2013b).

La nécessité d'une prise en compte appropriée de l'incertitude dans le processus d'ERS est aujourd'hui pleinement reconnue (Greenberg et al. 2012). Cela tient notamment à la complexité et la médiatisation accrues des questions de sécurité sanitaire (Wynne 1992, Gee 2009), telles que celles soulevées par les perturbateurs endocriniens (PE) (Phillips and Foster 2008, Gee 2006). Nourries par les incertitudes auxquelles les scientifiques se heurtent et auxquelles la société fait écho, ces questions font bien souvent l'objet de controverses (Guimaraes Pereira, Guedes Vaz, and Tognetti 2006, Craye, Funtowicz, and van der Sluijs 2005). Dans ce contexte, les méthodes destinées à mieux expliciter, de manière qualitative et/ou quantitative, les niveaux des incertitudes relatives à l'ensemble de la démarche d'évaluation des risques apparaissent indispensables.

Les enjeux d'une analyse d'incertitude sont multiples ; les premiers sont de caractériser les sources d'incertitude et de mieux appréhender la manière dont elles influent sur l'estimation du risque (van der Sluijs et al. 2003). Le cas échéant, des actions peuvent être suggérées pour réduire ces incertitudes (stratégies de recherches, acquisition de données, etc.). L'analyse d'incertitude peut contribuer à apprécier le niveau de confiance des résultats. Dans une démarche itérative, elle permet de revenir sur les choix initiaux pour améliorer ou assurer la robustesse de résultats (Verdonck et al. 2007). Enfin, la communication de l'incertitude permet, à l'ensemble des acteurs et des parties prenantes, de saisir la portée et les limites des connaissances disponibles dans leur globalité (van der Sluijs et al. 2003). Il est donc indispensable de la considérer pour garantir de façon optimale la qualité de la décision en matière de gestion des risques (NRC 2009).

Pour estimer de façon précise les risques sanitaires associés aux expositions environnementales, la démarche classique de l'ERS se fait en quatre étapes interdépendantes – (1) l'identification des dangers ; (2) la caractérisation de relations dose-réponse ; (3) l'évaluation de l'exposition ; et (4) la caractérisation des risques – se fonde sur un ensemble d'informations et de données, de nature différente, qualitative ou quantitative, relatives aux effets sanitaires, aux teneurs dans divers produits ou matrices, aux concentrations émises, aux doses d'expositions, etc. (NRC 2009). Or ces informations et ces données sont rarement complètes, manquent souvent de précision et, en général, ne sont pas représentatives de la population et/ou de la situation d'exposition considérées (van Asselt and Rotmans 2002, Kalberlah, Schneider, and Schuhmacher-Wolz 2003). Ainsi, l'incertitude est inhérente à l'ERS. Elle est présente à chacune de ses étapes. Elle n'empêche pas,

en principe, ni l'estimation des risques, ni la prise de mesures de gestion efficaces. Cependant, elle doit être prise en compte lors de l'élaboration de ces politiques (Vermeire et al. 2001).

1.2 Objet de la saisine

Le présent document porte sur le développement d'une démarche de prise en compte de l'incertitude en ERS, dans le contexte d'une expertise collective, telle que définie selon la norme NFX-50-110 suivie par l'Anses.

Les travaux s'inscrivent dans la continuité de l'analyse d'incertitude menée lors de l'ERS du Bisphénol A (BPA) (ANSES 2013a), et se placent dans une perspective d'application à des substances, dites perturbateurs endocriniens et reprotoxiques de catégorie 2.

Le document s'efforce de couvrir l'intégralité de la démarche d'ERS en prenant en compte les spécificités de chacune de ses étapes. Son objectif est de contribuer à l'élaboration d'une méthode originale de traitement de l'incertitude articulant approches qualitatives et quantitatives. Il amorce un travail de réflexion, de recensement et de développement d'outils et de méthodes pour traiter la question de l'incertitude en ERS des substances chimiques, dans un contexte d'expertise collective, caractérisé par la diversité du jugement des experts de différentes disciplines et compétences distinctes.

1.3 Modalités de traitement : moyens mis en œuvre et organisation

L'Anses a confié au groupe de travail « Perturbateurs Endocriniens » (GT PE) l'instruction de cette saisine.

Les travaux ont été conduits au sein d'un comité de travail dédié (CT Incertitude), constitué par plusieurs experts membres du GT PE, et élargi aux agents de l'Anses impliqués dans le traitement de la saisine PE et/ou dans le traitement de l'incertitude en ERS à l'Anses.

Les travaux ont été soumis régulièrement au GT PE (tant sur les aspects méthodologiques que scientifiques). Le rapport produit par le comité de travail tient compte des observations et éléments complémentaires transmis par les membres du groupe de travail.

Les travaux ont été présentés et adoptés par le CES « Substances ».

Ces travaux sont ainsi 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) »

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 rendues publiques *via* le site internet de l'Anses (www.anses.fr).

2 Démarche de travail adoptée par le CT Incertitude

Ce projet a été mené selon une démarche pragmatique. Il s'est appuyé sur l'expérience acquise lors de l'ERS du BPA menée par l'Anses (ANSES 2013a), et s'est déroulé en trois étapes. La première étape a consisté en une mise à plat du traitement de l'incertitude réalisé dans le cadre de l'ERS du BPA, avec pour objectif d'identifier les points d'amélioration et de développement ; la deuxième étape a consisté en une analyse critique des principales démarches faisant référence actuellement, telles que celles recommandées par l'IPCS¹, l'EFSA², l'ECHA³, etc. L'objectif de cette seconde étape était de passer en revue les méthodes existantes. La troisième étape avait pour objet de proposer une démarche de traitement de l'incertitude en ERS, adaptée des différentes démarches de référence analysées et compatible avec les procédures d'expertise en vigueur à l'Anses, en l'inscrivant dans un cas d'étude concret, lorsque cela a été possible.

Dans les deux premières étapes, le CT Incertitude a procédé par type d'approche, qualitative et quantitative, en s'intéressant aussi bien aux sources d'incertitude relatives à l'évaluation du danger (identification du danger et caractérisation des relations dose-réponse) qu'à celles relatives à l'évaluation de l'exposition. Dans la littérature, les sources d'incertitude relevant de ces deux volets de l'ERS sont traitées séparément. L'originalité du présent travail est de les considérer conjointement, de manière à mettre plus facilement en exergue les similitudes et les complémentarités des différentes approches d'analyse qui en font le traitement.

La suite du document présente une synthèse des travaux menés. Les chapitres 3 et 4 posent les éléments structurants de la démarche d'analyse d'incertitude en ERS. Le chapitre 5 décrit les principales méthodes d'analyse actuelles. Le chapitre 6 formule des propositions de démarche de traitement de l'incertitude adaptées aux pratiques d'expertise collective menées à l'Anses. Enfin, les deux derniers chapitres (chapitres 7 et 8) terminent par des recommandations pour une mise en œuvre à court et moyen terme et une conclusion sur le travail réalisé.

¹ *International Programme on Chemical Safety (IPCS)*

² *European Food Safety Authority (EFSA)*

³ *European Chemicals Agency (ECHA)*

3 Incertitudes : Définitions et concepts

3.1 Incertitude *versus* Variabilité

L'incertitude a souvent été définie par opposition à la variabilité ; les travaux portant sur la définition des termes et des concepts se sont longtemps attachés à distinguer ces deux notions (Hoffman and Hammonds 1994). La variabilité est définie comme l'étendue des valeurs réelles. Elle exprime l'hétérogénéité ou la diversité de la réalité, elle présente, en général, la propriété de ne pas être réductible. À l'inverse, l'incertitude est renvoyée à un manque de connaissance qu'il suffirait de combler par plus de recherches et/ou d'études (IPCS 2008). Or, de toute évidence, l'incertitude peut prévaloir y compris dans des situations où il y a beaucoup d'informations. L'apport de nouvelles informations peut diminuer l'incertitude ou, au contraire, révéler des incertitudes initialement ignorées ou sous-estimées (Rotmans and van Asselt 2001). L'incertitude n'est donc pas simplement une question d'absence de connaissances ou de connaissances insuffisantes, et les typologies proposées ces dernières années tendent à en élargir le spectre et les modalités d'expression (van der Sluijs et al. 2003).

Comme représenté à la Figure 1, à un niveau d'agrégation élevé, on peut distinguer deux grandes sources d'incertitude : la variabilité⁴ et une connaissance limitée (*i.e.* incomplète et/ou imparfaite) (van Asselt and Rotmans 2002, Craye, Funtowicz, and van der Sluijs 2005). La variabilité accompagnée de ressources limitées pour recueillir des données empiriques participe aux limites de la connaissance et concourt au caractère incertain et imprédictible de la réalité. Ainsi, une connaissance limitée découle partiellement de la variabilité ; cependant, la connaissance utilisée dans le cadre d'un processus déterministe, par exemple, peut elle aussi être incomplète et incertaine. Au final, un continuum transparait, avec un spectre des formes d'incertitude allant de l'inexactitude à l'ignorance irréductible⁵ (van Asselt and Rotmans 2002).

⁴ D'après l'article de van Asselt et Rotmans (2002), les sources de variabilité sont :

- le caractère aléatoire de la nature,
- la diversité des valeurs (*i.e.* des visions du monde, des normes et des valeurs différentes qui conduisent à des perceptions et des définitions différentes des problèmes, autrement dit à des jugements et/ou désaccords subjectifs se référant à l'« incertitude morale »),
- la diversité du comportement humain,
- la dynamique sociale, économique et culturelle (en lien avec le caractère non-linéaire, chaotique et imprédictible du processus sociétal, qui peut se référer également au comportement humain à un niveau macro), et
- les surprises technologiques (*i.e.* les nouveaux développements, les percées technologiques ou encore les conséquences technologiques non prévues).

⁵ **Spectre des incertitudes** : ensemble des formes possibles d'incertitude. D'après l'article de van Asselt et Rotmans (2002), le spectre des incertitudes en ERS comprend différents niveaux :

- L'*inexactitude* (*manque de précision* ou *imprécision*) a rapport avec les données ; elle exprime une incertitude métrique (exemples : les erreurs de mesure, *etc.*).
- Le *manque d'observations* ou de mesures fait référence à un manque de données pouvant être recueillies mais qui ne l'ont pas été.
- Le *caractère non mesurable* dans la pratique fait référence à un manque de données pouvant être recueillies en principe mais pas en pratique en raison, par exemple, du coût, du temps nécessaire ou encore parce qu'expérimentalement infaisable.

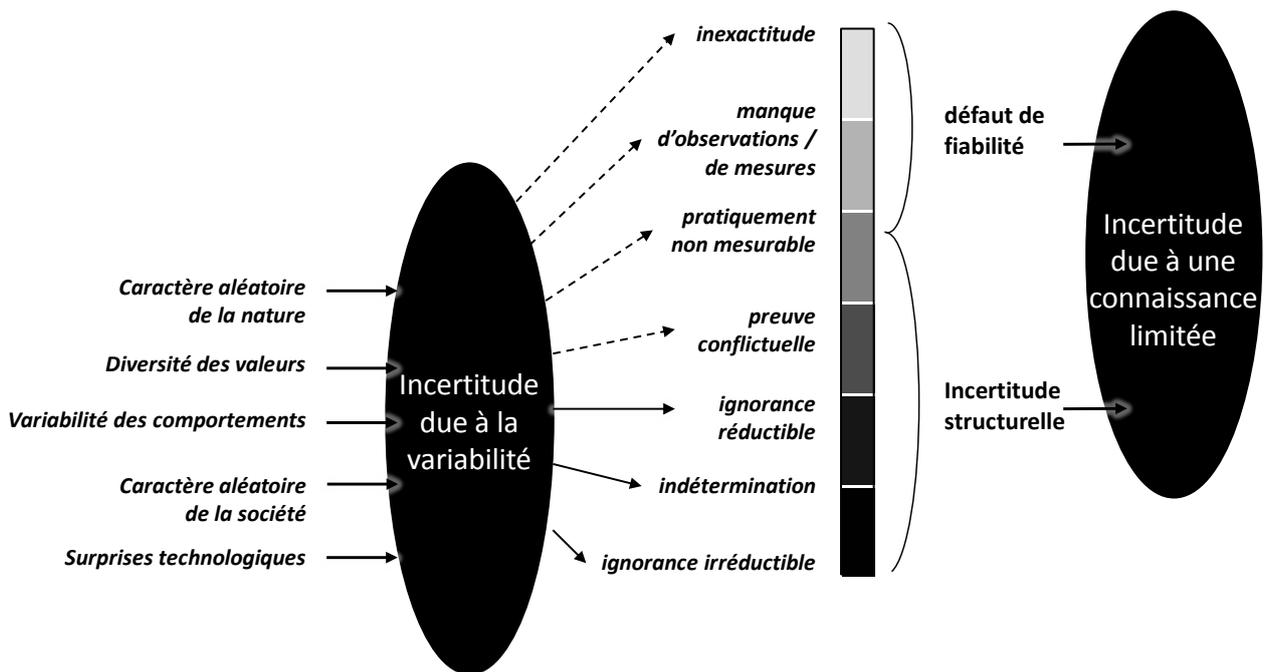


Figure 1 : Typologie des sources d'incertitude – Traduit de M.B.A. van Asselt and J. Rotman, *Uncertainty in integrated assessment modelling. From Positivism to Pluralism, Climatic Change*, 54: 75-105, 2002. La variabilité et une connaissance limitée sont présentées comme les deux grandes sources d'incertitude contribuant au large spectre des incertitudes allant de l'ignorance irréductible à l'inexactitude.

Sur la base de cette typologie, la définition de l'incertitude utilisée par l'Anses, dans le cadre de ses travaux d'expertise sur les PE, « fait référence à une situation où le corpus de connaissances disponibles sur un sujet particulier (exemple : les effets toxiques des PE) est perçu comme « problématique » (à savoir, peu connu ou inconnu, incomplet, incorrect, biaisé, pas assez convaincant, ...) par divers individus. » (ANSES 2013b).

3.2 L'incertitude : un concept multidimensionnel

La typologie présentée à la Figure 1 laisse percevoir l'incertitude comme un concept à plusieurs dimensions. En effet, une source d'incertitude se caractérise par différents aspects – appelés *dimensions* –, par exemple, là où elle se manifeste dans le processus d'expertise (sa

- La *contradiction des observations* se fonde sur différents jeux de données ou observations disponibles permettant des interprétations contradictoires.
- L'*ignorance réductible* fait référence à un processus qu'on n'observe pas, qu'on n'imagine pas théoriquement à un temps donné mais peut-être dans le futur.
- L'*indétermination* fait référence à un processus dont les principes et les lois sont comprises mais qui ne peut être complètement prédit et déterminé.
- L'*ignorance irréductible* fait référence à des processus et des interactions entre processus qui ne peuvent pas être déterminés par les capacités humaines.

*localisation*⁶), son origine selon qu'elle résulte de la variabilité ou des limites des connaissances du phénomène étudié (sa *nature*⁷), là où elle se situe sur le spectre des incertitudes (son *niveau*⁸) (Walker et al. 2003) ou encore la manière dont elle a été prise dans le processus d'évaluation (son *traitement*) (ANSES 2013a). Comme autres dimensions citées, il y a aussi la « qualification de la base de connaissances » (*qualification of the knowledge base*, en anglais), et la « valeur attachée aux choix » (*value-ladenness of choices* ou parfois *value loadings of choices*, en anglais). Ces deux dimensions font référence à deux formes d'incertitude très souvent masquées par des considérations quantitatives, à savoir l'*indétermination* et l'*ignorance* (Craye, Funtowicz, and van der Sluijs 2005). Elles visent respectivement à mettre en exergue la manière dont les hypothèses et le poids des « valeurs morales » (par opposition aux « valeurs numériques ») interviennent dans le processus de développement de la connaissance. La « qualification de la base de connaissances » correspond à appréciation de la fiabilité des connaissances sur lesquelles se fondent l'ERS. La « valeur attachée aux choix » correspond à l'importance, parfois subjective, attachée aux choix réalisés au cours de l'ERS. Cette dernière dimension se réfère à la présence de valeur ou de biais sous-jacents aux choix réalisés (*choix concernant la formulation de la question, la sélection ou non des données, leurs interprétations, etc.*) (van der Sluijs et al. 2003). Enfin, l'utilisation du couple *amplitude* et *direction* est souvent recommandée par les guides méthodologiques actuels (ECHA 2012, EFSA 2007, IPCC 2005, IPCS 2008) ; la *direction*⁹ d'une source d'incertitude indique le sens de son influence sur le résultat et son *amplitude*¹⁰, le degré d'incertitude qu'elle induit (IPCC 2005).

À titre d'exemple, pour caractériser les sources d'incertitude liées à l'évaluation de l'exposition au BPA, le GT PE a utilisé trois dimensions : la *localisation*, la *nature* et le *traitement* (ANSES 2013a).

La structure multidimensionnelle de l'incertitude peut se décliner sous une forme tabulée – ou matricielle – particulièrement commode pour recenser les sources d'incertitude et rapporter leurs caractéristiques (van der Sluijs et al. 2003). Comme présenté au Tableau 1, les lignes de la matrice recensent toutes les sources d'incertitudes identifiées. Cette liste est habituellement structurée selon la localisation des sources d'incertitude dans le processus d'expertise. Les colonnes de la matrice présentent les autres dimensions utilisées pour caractériser les sources d'incertitude rapportées. La *matrice des sources d'incertitude* ainsi définie est l'outil de prédilection pour l'analyse qualitative de l'incertitude (recensement et caractérisation systématique des sources d'incertitude). Par ailleurs, elle facilite la communication sur les sources d'incertitude en toute transparence.

⁶ La *localisation* d'une source d'incertitude est là où elle se manifeste dans le processus d'expertise, depuis la définition du périmètre de l'évaluation, jusqu'à la communication des résultats, en passant par l'identification et la caractérisation du danger, le calcul de l'exposition et du risque.

⁷ La *nature* d'une source d'incertitude indique son origine (ou cause) : « *incertitude découlant de la variabilité* » versus « *limites de la connaissance* ».

⁸ D'après l'article de Walker *et al* (2003), le *niveau* d'une source d'incertitude est là où elle se situe sur le *spectre* des sources d'incertitude allant de l'*ignorance irréductible* à l'*inexactitude*.

⁹ La *direction* d'une source d'incertitude est habituellement exprimée en utilisant les symboles « + » et « - » ou les flèches « ↓ » et « ↑ » :

« + » ou « ↑ » indique que le risque pourrait être surestimé en raison de l'incertitude ;

« - » ou « ↓ » indique que le risque pourrait être sous-estimé en raison de l'incertitude ; et

« +/- » ou « ↑↓ » indique que le risque a autant de chance d'être surestimé que sous-estimé en raison de l'incertitude.

¹⁰ L'IPCC (2005) recommande de définir l'*amplitude* en quatre modalités :

- *Pratiquement certain* : 99 % de certitude (incertitude considérée comme insignifiante) ;
- *Très probable* : environ 10 % d'incertitude (incertitude considérée comme limitée) ;
- *Probable* : environ 30 % d'incertitude (incertitude considérée comme significative) ;
- *Non probable* : 100 % d'incertitude (fort niveau d'incertitude).

Tableau 1 : Exemple de structure matricielle utilisée pour recenser et caractériser les sources d'incertitude en ERS

	Autres dimensions utilisées pour caractériser les sources d'incertitude			
	<i>dimension 1</i>	<i>dimension 2</i>	<i>...</i>	<i>dimension k</i>
Sources d'incertitude listées selon leur localisation dans le processus d'évaluation				
<i>source 1</i>				
<i>source 2</i>				
<i>...</i>				
<i>source n</i>				

3.3 Classification des sources d'incertitude

En règle générale, le recensement des sources d'incertitude s'effectue, à l'aide d'une classification. Cette dernière correspond à une liste générique de sources d'incertitude, structurée selon leur localisation dans le processus d'expertise. Différentes classifications de sources d'incertitude existent, toutefois, aucune d'entre elles ne fait réellement l'objet de consensus.

Les premiers travaux sur la classification des sources d'incertitude se sont généralement focalisés sur l'incertitude quantifiable (Bogen and Spear 1987, Frey 1992, Cullen and Frey 1999). Les sources d'incertitude quantifiable sont en majorité relatives aux données qui alimentent les modèles¹¹ (*données d'émission, de concentrations, de doses, etc.*) et à celles qui permettent d'estimer les paramètres des modèles. Par exemple, les méthodes analytiques utilisées pour déterminer les concentrations des substances chimiques dans l'air, l'eau, et l'alimentation présentent toutes des limites de détection et de quantification, qui leur confèrent plus ou moins de précision par rapport à la concentration réelle. Les valeurs toxicologiques de référence sont rarement dérivées d'études chez l'être humain et les études expérimentales à partir desquelles sont déterminées les LOAELs¹², NOAELs¹³ et BMDs¹⁴ présentent, généralement, soit un nombre limité de sujets, soit un nombre limité de doses étudiées, bien souvent les deux, *etc.* Toutes ces limites amènent à faire des extrapolations et autres hypothèses de travail qui sont, par définition, des sources d'incertitude à ne pas négliger.

Les sources d'incertitude non quantifiable ne sont pas à négliger non plus et doivent également être identifiées. Aujourd'hui, les hypothèses sous-jacentes à l'estimation du risque – par exemple, celles relatives aux situations « pire-cas » – sont bien souvent exprimées de manière insuffisante. Parmi les sources d'incertitude non quantifiable habituellement ignorées ou explicitées insuffisamment, il y a le périmètre de l'évaluation, défini au regard de la question posée aux

¹¹ Le terme « modèle » est considéré ici au sens mathématique du terme. Il s'agit d'une représentation mathématique d'un phénomène ou d'un système d'intérêt. Un modèle a pour but de traduire de manière opérationnelle ensemble de connaissances à des fins de déduction. Par traduction opérationnelle, on entend la possibilité de régénérer l'ensemble des résultats possibles conditionnellement à un nombre restreint de paramètres (Bernier et Parent, 2007). Un modèle répond à un objectif, simplifie la réalité et donne une représentation quantifiée du problème en main (Bernier et al, 2000).

¹² *Lowest-Observed-Adverse-Effect Level* (LOAEL) : Dose sans effet toxique observable

¹³ *No Observable Adverse Effect Level* (NOAELs) : Dose sans effet toxique observable

¹⁴ Benchmark Dose (BMD)

experts et du contexte réglementaire en vigueur. Il y a aussi la manière dont les connaissances ou les informations – aussi bien les publications, les modèles que les données – ont été sélectionnées et utilisées dans l'ERS. Il y a enfin la manière dont les résultats sont présentés et les conclusions, formulées. Négliger ces sources d'incertitude revient à n'explicitier que partiellement l'incertitude entourant les résultats de l'ERS ; cela peut induire, à tort, une impression de certitude, et conduire à surestimer la fiabilité des résultats de l'évaluation (Verdonck et al. 2007, van der Sluijs et al. 2005). Ainsi, une classification doit permettre le recensement de toutes les sources d'incertitude.

Conscient de cet enjeu, le RIVM a pris l'option de recommander une classification générique s'intéressant à l'ensemble du processus d'expertise depuis la définition du périmètre de l'évaluation, à la manière de rapporter les résultats (van der Sluijs et al. 2003). Cette classification se fonde sur celle développée par Walker *et al* (2003), et comprend cinq grandes sources d'incertitude incluant le *Contexte*, les *Données*, le *Modèle*, le *Jugement d'expert* et les *Résultats*. Ces cinq classes recouvrent l'ensemble du processus d'expertise (van der Sluijs et al. 2003) ; elles sont succinctement décrites ci-après :

- *Contexte* – La classe *Contexte* s'intéresse à la définition du cadre de l'ERS, *i.e.* aux choix réalisés concernant la population cible, la voie d'exposition, l'échelle de temps, l'échelle d'espace, *etc.*, et à ce que recouvre ces choix au regard de la question posée.
- *Données* – La classe *Données* s'intéresse à l'ensemble des données obtenues autrement que par jugement d'expert (*données de mesures, données de surveillance, données d'enquêtes, etc.*) utilisées dans l'évaluation, incluant celles utilisées pour la calibration des modèles. Il s'agit d'informations issues de recherches empiriques et/ou de système de recueil de données par opposition aux jugements d'experts.
- *Modèle* – La classe *Modèle* s'intéresse à la totalité des modèles utilisés pour l'évaluation, allant des modèles conceptuels aux modèles mathématiques (*modèles statistiques, modèles de processus causaux, etc.*). En particulier, pour la catégorie des modèles mathématiques, des sous-classes d'incertitude ont été introduites pour distinguer l'incertitude issue de la *structure du modèle* (relations), des *paramètres du modèle* (processus de paramétrage, conditions initiales, bornes des conditions), des *entrées du modèle* (données d'entrée, forces motrices externes), de même que des *aspects techniques du modèle* (implémentation informatique et logiciels utilisés).
- *Jugement d'expert* – La classe *Jugement d'expert* s'intéresse aux incertitudes induites par le recours au jugement d'expert lors de l'évaluation, cette source d'incertitude n'étant pas complètement couverte par les classes *Contexte*, *Modèle* et *Données*.
- *Résultats* – La classe *Résultats* s'intéresse aux sorties, indicateurs, propositions ou positions venant en réponse à la question posée.

La classification recommandée par le RIVM, contrairement à celle de l'IPCS, ne comprend pas de classe « Scénario », toutefois les sources d'incertitude s'y afférant sont bien considérées et se rattachent aux autres classes d'incertitude définies.

La classification recommandée par l'IPCS est spécifique à l'évaluation de l'exposition, elle

présente trois grandes classes – *Scénarios*, *Modèles* et *Paramètres*¹⁵ – définies ci-après, se déclinant elles-mêmes en 19 sous-classes présentées au Tableau 2.

Scénarios – La classe *Scénarios* se réfère aux faits, hypothèses et inférences considérés et utilisés au cours de l'évaluation mais qui, en réalité, ne sont pas représentatifs du champ ou de l'objet de l'expertise. Cette classe de sources d'incertitude inclut des spécifications incomplètes ou non adéquates au regard de l'agent, la population exposée, les informations spatio-temporelles, la situation d'exposition, les médias et voies d'exposition, etc.

Modèles – Comparée à la classification recommandée par le RIVM, la classe *Modèles* s'intéresse ici plus spécifiquement aux modèles mathématiques, le plus souvent rencontrés en évaluation de l'exposition. L'incertitude liée à un modèle découle du niveau de détail apporté dans sa définition. Ce niveau doit nécessairement être cohérent avec les objectifs de l'évaluation. Selon l'IPCS, l'incertitude liée aux modèles résulte principalement d'erreurs de modélisation (par exemple, de la non-considération de paramètres) et d'erreurs de relations (ou dépendances) (par exemple, tirer des conclusions erronées à partir de corrélations absentes ou incorrectes). Ces erreurs peuvent être dues à une définition inadéquate des bornes du modèle au regard du scénario adopté (par exemple, de la non considération d'éléments importants) ou à une mauvaise détermination des relations entre les paramètres (par exemple, par excès de simplification). Elles peuvent également survenir lorsque les hypothèses ne sont pas clairement explicitées ou lorsque le modèle développé ne tient pas compte de l'ensemble des hypothèses réalisées au regard du système étudié. L'utilisation d'un modèle au-delà du domaine pour lequel il a été conçu peut également induire une incertitude. Enfin, l'implémentation informatique du modèle peut souffrir d'erreurs de programmation.

Paramètres – La classe *Paramètres* recouvre les mêmes sources d'incertitude que la classe *Données* définie par le RIVM. Elle s'intéresse aux valeurs des paramètres et des variables d'entrée du modèle utilisées pour évaluer l'exposition. Ces données peuvent être issues de différentes sources : des référentiels tels que l'*Exposure Factors Handbook* (U.S.EPA 2011), des outils tels que *ConsExpo* (Delmaar, Park, and van Engelen 2005) ou *RefXP* (Mekel et al. 2007, Delmaar, Park, and van Engelen 2005), des enquêtes en population, des systèmes de surveillance, des études de mesures de concentrations, etc. Ces données sont le plus souvent entachées d'incertitude résultant d'erreurs : erreurs de mesure (aléatoires¹⁶ ou systématiques¹⁷), erreurs d'échantillonnage, erreurs liées au type de données utilisées (*données de mesure*,

¹⁵ D'après le guide de l'IPCS (2008), le terme « Paramètre » fait référence indifféremment à une variable d'entrée d'un modèle ou à un paramètre. Certains ouvrages préfèrent distinguer ces deux notions, cette dernière option est celle retenue pour le présent rapport. En particulier, dans l'ouvrage de (Cullen and Frey 1999), le terme « Paramètre » désigne une quantité utilisée pour calibrer ou définir un modèle tel que le(s) paramètre(s) d'un modèle probabilité (la moyenne et la variance d'une loi normale, la pente et l'ordonnée à l'origine d'un modèle de régression univarié, etc.). La valeur d'un paramètre d'un modèle est habituellement déterminée par ajustement sur un jeu de données de calibration. Le terme « entrées du modèle d'exposition (*exposure model inputs*) » est généralement préféré à « variable d'entrée » pour éviter toute confusion avec l'adjectif « variable ».

¹⁶ Les **erreurs aléatoires de mesure** sont à rattacher à la notion de précision ; elles induisent une variation autour de la « vraie » valeur du paramètre, sont souvent aggravées par la variabilité réelle de la quantité considérée (Zheng and Frey 2005). Par exemple, la concentration dans l'air d'un polluant, si la concentration varie dans le temps et que la méthode de mesure utilisée manque de précision et induit aléatoirement une erreur de mesure, alors, la variation observée de la concentration peut-être supérieure à la variation réelle. Si cette erreur n'est pas corrigée, elle peut conduire à une surestimation de la concentration.

¹⁷ Les **erreurs systématiques (ou biais)** résultent d'imperfections du dispositif de mesure (exemple : calibrations incorrectes d'appareils de mesures) ou de recueil des données. La moyenne des mesures réalisées ou des données recueillies peut alors différer de la valeur réelle.

modélisées ou de substitution¹⁸, valeurs par défaut¹⁹, jugement d'expert). L'extrapolation²⁰, ou le choix de distributions des paramètres peuvent également être des sources d'incertitude paramétriques.

Tableau 2 : Classification des sources d'incertitude relatives à l'évaluation de l'exposition, recommandée par l'IPCS – Traduit de l'IPCS (2008)

Classes d'incertitude	Sous-classes d'incertitude
Scénarios	<i>Agent</i> <i>Population exposée</i> <i>Microenvironnement</i> <i>Information temporelle et spatiale</i> <i>Activité de la population</i> <i>Mesure de gestion de risque</i> <i>Médias d'exposition (« Exposure pathway »)</i> <i>Situation d'exposition (« Exposure event »)</i> <i>Voie d'exposition (« Exposure route »)</i>
Modèles	<i>Hypothèses</i> <i>Dépendances</i> <i>Structure</i> <i>Équation</i> <i>Extrapolation</i> <i>Implémentation</i>
Paramètres	<i>Source de données</i> <i>Dépendance des données</i> <i>Valeur des données</i> <i>Unité des données</i>

La classification du RIVM et celle de l'IPCS présentent chacune leurs particularités, elles n'en restent pas moins cohérentes et complémentaires. L'Anses s'y est donc appuyée pour définir sa propre classification et mener à bien l'analyse qualitative de l'incertitude relative à l'évaluation de l'exposition au BPA. Sa démarche et la classification développées sont décrites au Chapitre 5 (p 32).

¹⁸ Les **données de substitution** sont celles utilisées en l'absence de données spécifiques à la situation d'étude.

¹⁹ Les **valeurs par défaut** sont des valeurs de référence recommandées dans des situations précises. Le niveau d'incertitude associé à l'utilisation de valeurs par défaut dépend de la procédure et de la qualité des données sur lesquelles elles se découlent.

²⁰ **Extrapoler** consiste à inférer à partir de données ou d'informations fragmentaires ou incomplètes (par exemple, extrapolation de l'animal à l'être humain, du passé au futur, de la voie orale à la voie inhalée, etc.). L'exactitude des données extrapolées pour la situation étudiée est une incertitude.

4 Approche par paliers

Comme souligné en introduction de ce rapport, une palette de méthodes et d'outils sont aujourd'hui disponibles pour réaliser une analyse d'incertitude en ERS. Ils vont du plus simple au plus sophistiqué. Leurs utilisations doivent être adaptées aux besoins de l'évaluation. En effet, les évaluations n'ont pas toutes le même niveau de complexité et ne requièrent pas toutes le même niveau d'analyse d'incertitude (une évaluation des risques sommaire n'appelle pas toujours à une analyse d'incertitude approfondie). Par ailleurs, toutes les incertitudes ne peuvent être traitées de la même manière ; par exemple, les incertitudes non quantifiables ne peuvent être traitées que par des approches qualitatives. Enfin, les données disponibles ne permettent pas toujours de mener une analyse approfondie ; or l'analyse d'incertitude ne suppose pas le recueil de données supplémentaires, elle a pour objectif de mettre en perspective les limites de l'information disponible et de son interprétation sur laquelle se base la quantification du risque (EFSA 2007).

Pour proportionner l'effort d'analyse au besoin du problème, la majorité des guides méthodologiques recommandent l'approche par paliers (*tiered or stepwise approach* en anglais). Ce dernier se réfère à un processus progressif évoluant en complexité, de manière graduée, au fur et à mesure des étapes (ou niveaux) (IPCS 2008). Le nombre de niveaux et leur contenu peuvent différer d'un guide à un autre. Cependant, les plus récents reprennent la proposition de l'IPCS, celle d'une approche en quatre niveaux : « Screening » (Niveau 0) », « Qualitatif » (Niveau 1), « Déterministe » (Niveau 2) et « Probabiliste » (Niveau 3) (cf. Figure 2) (EFSA 2007, IPCS 2008, NRC 2009).

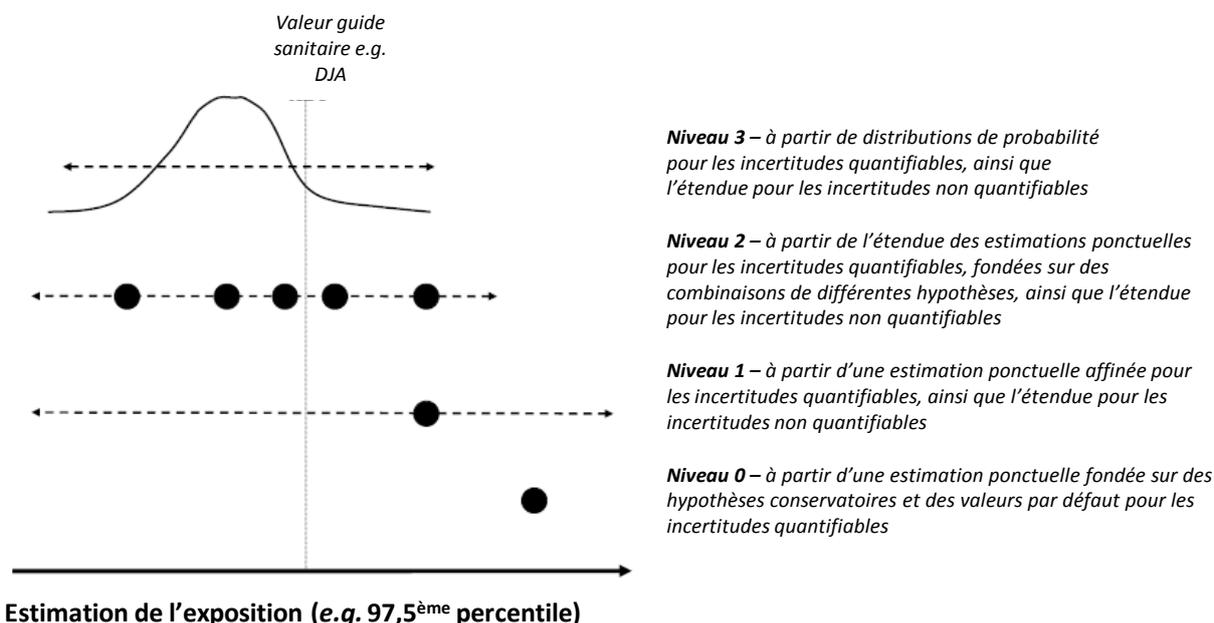


Figure 2: Illustration graphique des différents niveaux d'analyse d'incertitude, allant progressivement du niveau le moins élevé (Niveau 0) au niveau le plus élevé (Niveau 3). Les ronds noirs représentent des estimations ponctuelles de l'exposition et les lignes en pointillées indiquent l'étendue de l'exposition – Adapté de l'EFSA (2007).

4.1 Description des étapes

4.1.1 Niveau 0 – Screening – Estimation ponctuelle sur la base d'hypothèses conservatoires, en utilisant le cas échéant des valeurs par défaut

Ce niveau d'analyse est essentiellement utilisé dans le cadre d'une ERS relativement simple, menée en première intention lorsque les données disponibles ne permettent pas une analyse fine ou lorsque les expositions ou les risques considérés sont vraisemblablement bien en deçà ou bien au-dessus des valeurs sanitaires de référence (VTR²¹, DJA²², etc.). Dans ces situations, les incertitudes sont prises en compte, de façon relativement basique, à travers des facteurs d'incertitude par défaut, établis ou recommandés pour des problèmes similaires (IPCS 2008).

Bien qu'aisé à mettre en œuvre, ce niveau d'analyse est à réserver uniquement aux évaluations sommaires, reposant sur des hypothèses conservatoires. Pour des évaluations plus approfondies, mettant en jeu des hypothèses considérées plus réalistes, il est conseillé de passer à un niveau d'analyse supérieur (EFSA 2007).

4.1.2 Niveau 1 – Qualitatif – Identification systématique et caractérisation des sources d'incertitude

Ce niveau d'analyse comprend un recensement des sources d'incertitude puis une caractérisation qualitative systématique de chacune des sources identifiées. Pour une démarche systématique et structurée, le recensement s'appuie sur une classification (cf. paragraphe 3.3). S'agissant de l'étape de caractérisation, l'objectif principal est, selon l'IPCS (2008), d'évaluer l'influence individuelle de chaque source d'incertitude sur le résultat de l'évaluation. Les critères *amplitude*¹⁰ et *direction*⁹ y sont destinés, toutefois, bien d'autres critères existent et peuvent être utilisés en complément ou à leur place. L'ensemble de ces informations est recueilli et synthétisé sous une forme tabulée appelée matrice des incertitudes comme illustrée par le Tableau 1 (p. 21).

Dans la mesure du possible, une évaluation de l'effet combiné des différentes sources d'incertitude devrait compléter l'analyse des effets individuels. Pour cela, une approche de type *Weight-of-evidence* (WOE) avec expression du niveau de confiance ou de vraisemblance des résultats d'identification du danger, telle que décrite plus loin au Chapitre 5 pourrait être envisagée (IPCS 2008).

Enfin, des méthodes qualitatives telles que présentées au paragraphe 6.3.3 (p. 53) peuvent permettre la caractérisation de l'influence des sources d'incertitude en termes de direction⁹ et d'amplitude¹⁰.

Ce niveau d'analyse pourrait être mis en œuvre, de façon systématique, dans toute analyse d'incertitude. En effet, seule l'approche qualitative permet de traiter les incertitudes non quantifiables. En outre, en présence d'incertitudes quantifiables, l'approche qualitative constitue une bonne base de départ sur laquelle les analyses quantitatives de l'incertitude (Niveau 2 ou 3) devraient s'appuyer. Les analyses de niveaux 2 et 3, décrites ci-dessous, s'intéressent exclusivement aux incertitudes quantifiables.

4.1.3 Niveau 2 – Déterministe – Approche déterministe de l'analyse de sensibilité

L'approche déterministe correspond à une forme relativement sommaire d'une analyse de sensibilité dont l'objectif principal est d'identifier les éléments (*facteurs d'entrée, structure des modèles, scénarios d'exposition, etc.*) les plus influents sur les résultats de l'évaluation (ECHA 2012). Pour les facteurs d'entrée des modèles, par exemple, cela consiste à calculer différentes estimations ponctuelles de l'évaluation dérivées de différentes valeurs (e.g. faible, moyenne et

²¹ Valeur toxicologique de référence (VTR)

²² Dose journalière admissible (DJA)

forte) de ces facteurs. Ces différentes valeurs doivent traduire différentes hypothèses d'entrée, l'objectif étant d'en examiner l'impact sur les résultats (EFSA 2007). Ces calculs sont généralement réalisés en modulant les facteurs d'entrée potentiellement les plus influents au regard des résultats de l'analyse qualitative (Niveau 1) et en fixant la valeur des autres, à la moyenne par exemple. La modulation des facteurs d'entrée s'effectue, le plus souvent, à tour de rôle, mais une modulation conjointe est également possible. Les résultats de ces analyses sont souvent présentés en pourcentage de variation de l'estimation par rapport à une estimation considérée comme « valeur de référence » (IPCS 2008).

Dans certaines situations, notamment lorsque des contributeurs majeurs à l'incertitude sont identifiés, une caractérisation plus fine de leurs contributions – contributions propres (individuelles) et conjointes – à l'incertitude des résultats peut être nécessaire. Une analyse probabiliste de l'incertitude (Niveau 3) est alors à mettre en œuvre.

4.1.4 Niveau 3 – Probabiliste – Approche probabiliste de l'analyse de sensibilité et d'incertitude

Ce niveau d'analyse a pour objectif de caractériser quantitativement, à l'aide de distributions de probabilité, les effets individuels et conjoints de différentes sources d'incertitude. Ces analyses présentent plusieurs intérêts, par exemple, selon l'IPCS (2008) :

- elles aident à évaluer l'incertitude entourant les résultats d'un modèle en estimant la précision ;
- elles sont utiles au gestionnaire du risque, dans certains cas, en fonction du périmètre de l'évaluation ;
- elles peuvent aussi aider à identifier parmi les facteurs d'entrée du modèle, les sources clés d'incertitude, dans la perspective de futures actions de recherche ou de recueil de données notamment.

La variabilité et l'incertitude due aux limites des connaissances disponibles peuvent toutes deux être quantifiées à l'aide de distributions de probabilité (Morgan and Henrion 1990). Toutefois, parce qu'elles représentent des concepts différents, avec possiblement des implications différentes en matière de gestion de risque, il est nécessaire de les distinguer (Bogen et al. 2009). Il s'agit pour cela de bien les séparer et de les caractériser de façon conjointe dans l'estimation du risque (Cullen and Frey 1999, Bernier, Parent, and Boreux 2000).

Les sources d'incertitude faisant l'objet d'un traitement probabiliste sont généralement celles relatives aux données d'entrées d'un modèle. Cependant, d'autres sources d'incertitude peuvent aussi être d'importance considérable et devraient, de la même façon, être quantifiées. C'est le cas, par exemple, de l'incertitude des scénarios, à ceci près qu'à ce jour, il n'existe pas de méthode rigoureuse permettant de réaliser un tel traitement. Pour l'incertitude des modèles, le traitement se ramène souvent à celui de l'incertitude des facteurs d'entrée et met en œuvre des méthodes de propagation de l'incertitude (Rojas, Feyen, and Dassargues 2008).

Pour propager l'incertitude des facteurs d'entrée à travers un modèle, différentes approches existent : analytiques (e.g. transformation de variable), par approximation (e.g. approximation basée les séries de Taylor) ou encore numériques fondées sur des techniques de simulation (e.g. Monte Carlo, hypercube latin, etc.) (Parent and Bernier 2007, Bogen et al. 2009). Dans le cas d'un modèle simple, mettant en jeu peu de facteurs, une solution analytique peut être trouvée. Le plus souvent, la complexité des modèles conduit à utiliser des méthodes numériques, dont certaines permettent de séparer variabilité et incertitude liée au niveau des connaissances disponibles (e.g. simulation de Monte Carlo à deux dimensions, Monte Carlo par chaîne de Markov, etc.) (Frey 1992, Bernier, Parent, and Boreux 2000).

Pour les analyses mettant en œuvre des méthodes numériques, une formalisation de la démarche en quatre étapes peut être utilisée (Makowski 2013) :

Étape 1 – Définir les distributions de probabilité des facteurs d'entrée du modèle

Étape 2 – Générer des échantillons à partir des distributions définies à l'étape 1

Étape 3 – Calculer la sortie du modèle pour chaque série de facteurs d'entrée générée

Étape 4 – Estimer la distribution de la sortie du modèle

4.1.4.1 Analyse probabiliste de l'incertitude

Étape 1 – Définir les distributions de probabilité des facteurs d'entrée du modèle

Le point de départ d'une analyse d'incertitude traitant des facteurs d'entrée d'un modèle est la définition des distributions de probabilité. Habituellement, seuls les facteurs considérés comme déterminants sont concernés, *i.e.* ceux dont la variabilité et/ou l'incertitude relative au niveau de connaissance sont susceptibles d'impacter de façon non négligeable la ou les sorties du modèle. La valeur des autres facteurs est fixée, par exemple, à la moyenne. Une première liste de facteurs déterminants peut être établie au regard des analyses de niveaux 1 et 2, puis consolidée à l'aide d'une analyse de sensibilité comme précisée plus loin dans le présent chapitre.

Pour définir une distribution de probabilité, il existe trois grands types d'approches : les approches empiriques, les approches utilisant la connaissance des mécanismes (chimiques, biologiques et/ou physiques) sous-jacents, et les approches utilisant le jugement d'expert (Bogen et al. 2009). En réalité, ces trois types d'approches sont complémentaires, et dans la pratique, s'utilisent de façon combinée. L'approche purement empirique se sert des données de mesures ou d'observation existantes pour construire une fonction de répartition ou un histogramme de fréquence. Elle est particulièrement adaptée lorsque l'on dispose d'un échantillon aléatoire représentatif de la population et de la situation d'exposition considérée, ce qui est rarement le cas. En effet, pour beaucoup de facteurs d'intérêt, des données empiriques font cruellement défaut, et lorsqu'elles existent, elles ne sont pas toujours de qualité suffisante ou ne sont pas représentatives de la population et/ou de la situation d'exposition étudiées. Dans de telles situations, l'utilisation de données de substitution peut être appropriée à condition qu'elle n'introduise pas de biais importants. La pertinence ou non de l'utilisation de données de substitution est, en général, laissée à l'appréciation des experts (NRC 2009). Autrement, lorsque les situations où les mécanismes physiques, biologiques et/ou chimiques sous-jacents sont relativement bien connus et modélisés, un intervalle des valeurs possibles, voire même une forme de distribution peut être attribués aux facteurs d'intérêt ; par exemple, la connaissance actuelle du processus de dilution conduit très souvent à attribuer à la concentration d'un polluant dans l'environnement une distribution log-normale (Ott 1990). Une fois la famille de distribution sélectionnée, il ne reste alors plus qu'à estimer les paramètres de ces lois de probabilité en utilisant les données ou informations disponibles (données empiriques et/ou jugements d'expert).

Les méthodes d'inférence statistique classiques – dites fréquentistes – permettent d'estimer les statistiques de distribution habituelles (moyenne, variance, percentiles, *etc.*). Par ailleurs, en complément de ces différentes méthodes, l'analyse probabiliste de l'incertitude en ERS a de plus en plus recours à l'approche bayésienne (Ciffroy, Keller, and Pasanisi 2013, Ciffroy et al. 2011, Makowski 2013, Makowski and Mittinty 2010, Tressou 2008). La statistique bayésienne offre, en effet, une variété de méthodes et d'algorithmes pour calculer les distributions de probabilité en combinant de façon optimale données empiriques (données « objectives ») et connaissance d'expert (données « subjectives ») (Makowski 2013, Bogen et al. 2009).

De façon générale, quelle que soit la démarche adoptée, il est convenu de fonder le choix des distributions de probabilité des facteurs d'entrée des modèles, de préférence, sur des informations dites « objectives » – données empiriques (données de mesures ou d'observations) ou, à défaut, données modélisées – plutôt qu'exclusivement sur des informations dites « subjectives » (jugements d'expert) (ECHA 2012, EFSA 2007, IPCS 2008). Pour soutenir cette stratégie, l'US EPA²³ dispose d'une source de données de référence en l'*Exposure Factors Handbook* (EFH).

²³ *United States Environmental Protection Agency (US EPA)*

L'EFH est, en effet, une compilation des données clés sur les facteurs d'exposition, mise à jour régulièrement. Pour l'instant, il n'y a pas d'équivalence française ; toutefois, un projet sur les facteurs d'exposition pour l'ERS est en cours à l'Anses. Ce projet a pour objectif de constituer une base de données de référence sur les facteurs nécessaires à l'évaluation de l'exposition de la population française (population sensibles et travailleurs inclus). La base de données centralisera l'ensemble des données françaises disponibles, en particulier les données de distribution. Elle pourra notamment bénéficier des données produites par des organismes membres du R31²⁴ de l'Anses, ou issues de projets européens tel que le projet EPHECT²⁵ sur les habitudes d'utilisation et d'usage de produits de consommation recueillies auprès des consommateurs de 10 pays européens, notamment la France. Les travaux sont conduits dans le cadre d'un collectif national et pourraient également s'intégrer aux réflexions européennes, par exemple, par le biais du projet *Exposât*²⁶ piloté par le JRC-EC²⁷. À l'instar de l'EFH de l'US EPA, cette base de données devrait constituer une source importante de données sur laquelle s'appuyer pour spécifier la variabilité et estimer certains aspects de l'incertitude des connaissances rencontrées au cours de l'évaluation de l'exposition de la (des) population(s) française(s).

Étape 2 – Générer des échantillons à partir des distributions définies à l'étape 1

L'échantillonnage aléatoire simple, à l'aide de la technique de Monte Carlo, est la méthode la plus usuelle pour générer des échantillons représentatifs issus de distributions de probabilité. Cette stratégie d'échantillonnage permet d'obtenir des estimations sans biais de l'espérance et de la variance de variables aléatoires (Robert 2006). Cependant, des stratégies d'échantillonnage plus sophistiquées existent et peuvent être plus adaptées à l'étude de phénomènes plus complexes, mettant en jeu plusieurs facteurs interdépendants. Ces stratégies font appel à des méthodes de planification d'expériences numériques (plans factoriels, hypercube latin, surface de réponse, plan optimal, etc.) (Jourdan 2005) ; par exemple, un plan factoriel complet a été utilisé par l'EFSA pour combiner l'estimation du minimum, du maximum et des valeurs les plus fréquentes des différents facteurs considérés comme autant de sources d'incertitude (EFSA 2008). La méthodologie et les propriétés des plans d'expérience sont présentées dans de nombreux ouvrages, pour les approches classiques (Droesbeke, Fine, and Saporta 1997) comme pour les approches numériques plus récentes (Fang, Li, and Sudjianto 2006).

La taille de l'échantillon à simuler est également un élément essentiel du plan d'échantillonnage car la fiabilité des résultats de modélisation en dépend. Un échantillon de faible taille ne permet pas toujours de couvrir pleinement l'ensemble du domaine de définition des facteurs considérés ; cela peut conduire à un manque de précision de la variabilité (estimation de la moyenne, de la variance, des quantiles, etc.) des facteurs d'entrée, et aboutir *in fine* à une mauvaise

²⁴ Réseau des partenaires de l'Anses (R31) : en vue de l'accomplissement de ses missions, l'Agence a mis en place et anime un réseau d'organismes scientifiques qui interviennent dans son champ de compétences (article R1313-1 du Code de la santé publique). Ce réseau a été officiellement lancé en octobre 2010, et compte aujourd'hui 30 organismes partenaires de l'Anses. Une charte en établit le cadre de fonctionnement. Elle précise notamment que ce réseau a pour objectif de renforcer les coopérations aux fins :

- d'évaluation des risques sanitaires dans le domaine de l'alimentation, de l'environnement, du travail, et de la santé animale et végétale;
- de veille et d'alerte des pouvoirs publics en cas de risques pour la santé publique;
- d'amélioration de la connaissance des risques sanitaires dans le domaine de compétence de l'Anses.

²⁵ Project "Emission, Exposure Patterns and Health Effects of Consumer Products in EU" (Project EPHECT) <https://sites.vito.be/sites/ephect/Pages/home.aspx>

²⁶ <http://expofacts.jrc.ec.europa.eu/>

²⁷ Joint Research Center – European Commission (JRC-EC)

approximation de la distribution des sorties du modèle. *A contrario*, l'augmentation du nombre de simulation n'apporte pas nécessairement plus d'information, et la simulation d'un échantillon de très grande taille n'est pas toujours possible. Ainsi, le choix de la taille d'un échantillon résulte d'un compromis entre précision recherchée et temps de simulation raisonnable (Makowski 2013).

Étape 3 – Calculer la sortie du modèle pour chaque série de facteurs d'entrée générée

Il s'agit à ce stade-ci de faire tourner le modèle, *i.e.* de calculer la valeur de la (ou des) variable(s) de sortie du modèle, pour chaque série de valeurs des facteurs d'entrée, générée au cours de la précédente étape. Le temps de calcul nécessaire dépend de la complexité du modèle.

Étape 4 – Estimer la distribution de la sortie du modèle

La distribution des sorties du modèle peut être décrite et résumée selon une multitude de manières ; par exemple, graphiquement, à travers des histogrammes de fréquence absolue, relatives ou cumulées, de probabilités, des représentations graphiques de la fonction de répartition, de la densité de probabilité, *etc.* ; un résumé à l'aide des statistiques de distribution classiques (*e.g.* moyenne, variance, percentiles, *etc.*) peut également être très utile ; lorsque plusieurs variables de sortie sont en jeu, des projections de points dans le plan ou l'espace, le calcul de coefficients de corrélation, *etc.* peuvent venir en complément. L'ensemble de ces techniques sont d'usage courant en ERS (ANSES 2013a, Bruchou et al. 2013).

4.1.4.2 Analyse de sensibilité

En évaluation probabiliste des risques, en particulier lorsqu'il s'agit de traiter de l'incertitude des facteurs d'entrée des modèles, il est d'usage d'articuler analyse de sensibilité et analyse d'incertitude. Ces deux analyses ont des rôles complémentaires. L'analyse de sensibilité mesure quantitativement la contribution des facteurs d'entrée d'un modèle aux variations de ses sorties. Elle distingue ainsi les facteurs d'entrée qui ont une forte influence sur les sorties du modèle de ceux qui ont une moindre influence. En d'autres mots, l'analyse de sensibilité permet de hiérarchiser l'importance des différents facteurs d'un modèle. Les résultats d'une analyse de sensibilité peuvent donc aider à identifier les facteurs qu'il serait nécessaire de connaître de manière plus précise. Dans le cadre d'un processus itératif visant à simplifier ou parfaire un modèle, elle peut être une étape utile pour identifier ou restreindre la liste des facteurs d'entrée auxquels attribuer des distributions pour représenter la variabilité et/ou l'incertitude liées au niveau des connaissances disponibles. Au final, l'analyse de sensibilité permet de classer les facteurs d'entrée en termes de contribution à l'incertitude totale, sur laquelle pourront s'appuyer des recommandations d'actions visant à réduire l'incertitude (ECHA 2012).

L'analyse de sensibilité a fait l'objet de nombreux développements méthodologiques. Ces développements sont bien documentés dans la littérature ; on pourra notamment se référer aux travaux de Saltillo *et al.* pour un panorama relativement complet du sujet (Saltelli et al. 2008, Saltelli et al. 2012, Saltelli et al. 2004). L'ouvrage « Analyse de sensibilité et exploration des modèles – Application aux sciences de la nature et de l'environnement » présente également une synthèse concise sur le sujet (Bruchou et al. 2013). Il dresse une vision globale des différents modèles d'analyse de sensibilité actuels et propose une démarche pragmatique pour choisir la méthode la plus adaptée à la question posée.

4.2 Niveau d'analyse nécessaire à l'analyse d'incertitude

Les différentes possibilités d'analyse – allant de la plus basique (niveau 0) à la plus sophistiquée (niveau 3) – posent la question du niveau d'analyse d'incertitude nécessaire à une évaluation donnée. Une analyse probabiliste approfondie n'est ni nécessaire ni recommandée pour toutes les ERS :

- une analyse simple – mais informative et complète – peut suffire à révéler la robustesse des résultats d'une ERS (par exemple, une option de gestion au regard des incertitudes

clés) ; une analyse plus complexe n'est nécessaire que si elle apporte de l'information supplémentaire (NRC 2009) ;

- dans le cas d'une analyse quantitative d'incertitude, certaines variables bien caractérisées peuvent être fixées *a priori*. Ainsi, dans une seule et même analyse, les sources d'incertitude considérées peuvent être traitées à des niveaux d'analyse différents (IPCS 2008) ;
- le niveau d'analyse choisi pour traiter l'incertitude doit être cohérent avec l'évaluation qu'elle accompagne (par exemple, il n'est pas cohérent d'associer une analyse probabiliste de l'incertitude à une évaluation déterministe de l'exposition) (IPCS 2008) ;
- enfin, l'analyse d'incertitude doit être équilibrée entre les différents volets de l'ERS. Par exemple, si une analyse d'incertitude approfondie est menée pour le volet « évaluation de l'exposition », une analyse approfondie doit aussi être menée pour le volet « évaluation du danger ». Une analyse d'incertitude déséquilibrée entre deux volets peut induire une interprétation erronée du degré d'incertitude/niveau de confiance du résultat de l'ERS (IPCS 2008, NRC 2009).

Il est donc important de planifier l'analyse d'incertitude – ses objectifs et le niveau envisagé – dès la phase initiale de l'ERS. Cependant, si de nombreux guides sur les méthodes d'analyse existent, aucun ne traite du niveau d'analyse approprié pour traiter l'incertitude dans une ERS donnée, et cela fait défaut.

5 Revue critique des principales méthodes actuelles

Le présent chapitre fait brièvement état des approches qualitatives et quantitatives habituelles pour traiter l'incertitude en ERS. Il se décline en deux grandes parties, la première traite de l'incertitude dans l'évaluation de l'exposition et rapporte plus précisément la démarche adoptée lors de l'ERS du BPA, ses limites et les difficultés rencontrées. La deuxième partie s'intéresse à l'incertitude dans l'identification et la caractérisation du danger et présente une revue critique des méthodes actuelles spécifiques à ce volet. Cette différence d'approche vient du fait que le volet « Danger » de l'ERS du BPA, menée par le GT PE de l'Anses, n'a pas bénéficié d'analyse d'incertitude au même titre que le volet « Exposition », si ce n'est en filigrane dans le cadre du processus d'évaluation du poids de la preuve.

5.1 Analyse d'incertitude dans l'évaluation de l'exposition – Démarche adoptée lors de l'ERS du BPA

La démarche d'analyse d'incertitude dans l'évaluation de l'exposition au BPA a été définie sur la base des recommandations formulées par l'IPCS dans son guide méthodologique sur la caractérisation et la communication de l'incertitude dans l'évaluation de l'exposition (IPCS 2008). Elle a été adaptée aux procédures d'expertise collective en vigueur à l'Anses, et se décline en deux parties : une analyse qualitative (analyse d'incertitude de niveau 1), et une analyse probabiliste (analyse d'incertitude de niveau 3). Ces deux volets sont présentés ci-après.

5.1.1 Analyse qualitative de l'incertitude dans le calcul de l'exposition au BPA

5.1.1.1 Présentation de la démarche

L'analyse qualitative menée a consisté en un recensement et une caractérisation des sources d'incertitude auxquelles l'évaluation de l'exposition s'est heurtée. Elle s'est appuyée sur une classification des sources d'incertitude, développée spécifiquement pour les besoins de l'expertise Anses (ANSES 2013a). Toutefois, cette classification se veut générique, *i.e.* applicable non seulement au BPA mais aussi aux autres PE et même plus généralement aux substances chimiques. Elle a été définie sur la base des nombreux travaux publiés sur ce sujet et plus particulièrement sur les recommandations formulées par le RIVM (van der Sluijs et al. 2003) et l'IPCS (IPCS 2008). Les diverses sources d'incertitudes ont été réparties en quatre classes selon leur *localisation* dans le processus d'évaluation : *incertitudes contextuelles*, *incertitudes liées au scénario d'exposition*, *incertitudes liées aux modèles d'exposition* et *incertitudes liées aux données d'entrée qui alimentent les variables des modèles*. Ces quatre classes se déclinent elles-mêmes en sous-classes telles que définies dans les quatre tableaux ci-dessous (cf. Tableau 3 à Tableau 6).

Tableau 3: Classification des sources d'incertitude relatives au contexte utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA

Classe des sources d'incertitude contextuelle	
<p>Classe qui recouvre l'ensemble des sources d'incertitude relevant de la définition du périmètre de l'évaluation (Knol et al. 2009, Walker et al. 2003). Dans le contexte du GT PE, la définition du périmètre de l'expertise découle de choix qui ne sont pas dans le pouvoir du GT, mais d'autres instances (habituellement, le ou les ministères qui formulent une question, dans le cas des saisines, ou d'instances réglementaires) (ANSES 2013a). Cette classe d'incertitude se décline en deux sous-classes décrites ci-dessous.</p>	
- Choix des sources d'exposition	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude induites par le choix des sources d'exposition</i>
- Choix des populations cibles considérées	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude induites par le choix des populations cibles considérées.</i>

Tableau 4 : Classification des sources d'incertitude relatives aux scénarios d'exposition utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA

Classe des sources d'incertitude des scénarios	
<p>Classe recouvrant l'ensemble des sources d'incertitude induites par les hypothèses sur lesquelles reposent les scénarios (hypothèses qui concernent les populations cibles, les sources et voies d'exposition considérées, les échelles de temps considérées, etc.). Par définition, un scénario est une simplification de la réalité, il ne tient pas compte de tous les facteurs affectant le problème à analyser (Cullen and Frey 1999). Cependant, si les scénarios développés ne tiennent pas compte, par exemple, de toutes les populations cibles, ou des sources et des voies majeures d'exposition, ou des échelles de temps appropriées au périmètre de l'évaluation, alors l'estimation de l'exposition est susceptible d'être biaisée. Dans le contexte du GT PE, toutes les informations – erronées ou incomplètes, incluant approximations, extrapolations et autres valeurs par défaut – utilisées dans la définition des scénarios, ont été considérées (ANSES 2013a). Cette classe d'incertitude se décline en quatre sous- classes décrites ci-dessous.</p>	
- Populations cibles	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées aux populations cibles considérées</i>
- Sources d'exposition	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées aux sources d'exposition considérées</i>
- Voies d'exposition	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées aux voies d'exposition considérées</i>
- Échelle de temps	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées à l'échelle de temps d'exposition considérée</i>

Tableau 5 : Classification des sources d'incertitude relatives aux modèles d'exposition utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA

Classe des sources d'incertitude des modèles	
<p>Classe recouvrant l'ensemble des sources d'incertitude qui se réfèrent aux biais et à l'imprécision relatifs aux compromis réalisés ou aux connaissances insuffisantes lors de la définition de la structure des modèles et de leur calibration (<i>i.e.</i> l'estimation des paramètres). Ainsi, l'incertitude des modèles peut être liée à leur structure, leur validation, les extrapolations faites, leurs limites, etc. (IPCS 2008). Si les variables intervenant dans l'exposition à modéliser ne sont pas toutes considérées ou si les équations mathématiques choisies ne décrivent pas ou que partiellement les relations entre ces variables ou si les corrélations entre les variables ne sont pas ou que partiellement prises en compte, il peut en découler un biais d'estimation. Dans le contexte du GT PE, l'incertitude liée aux modèles concerne la capacité de ceux-ci à décrire de manière exacte et complète la réalité qu'ils sont censés caractériser, à travers la quantification des relations entre des données d'entrée et des réponses de l'ensemble ou d'une partie d'un système (ANSES 2013a). Cette classe d'incertitude se décline en trois sous- classes décrites ci-dessous.</p>	
- Choix des variables d'entrée	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées au choix des variables des modèles d'exposition</i>
- Choix des équations mathématiques	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées au choix des équations mathématiques qui caractérisent les relations entre les variables des modèles</i>
- Corrélation entre les variables d'entrée	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées aux corrélations entre les variables des modèles</i>

Tableau 6 : Classification des sources d'incertitude relatives aux données d'entrée des modèles d'exposition utilisée par le GT PE pour recenser les sources d'incertitude lors de l'évaluation de l'exposition au BPA

Classe des sources d'incertitude des données d'entrée des modèles	
<p>Classe recouvrant l'ensemble des sources d'incertitude qui se rapportent à une série de facteurs (manque de données, fiabilité et précision des mesures, extrapolation des données, etc.) induisant un manque de représentativité des données utilisées au regard de la population cible considérée (Cullen and Frey 1999). Ce manque de représentativité peut donner lieu à une estimation de l'exposition non généralisable à la population cible considérée. Il peut être dû, soit à la taille de l'échantillon, soit au nombre d'échantillons prélevés et analysés, soit à la méthode d'échantillonnage. Dans le contexte du GT PE, cette source d'incertitude concerne les valeurs numériques (valeurs ponctuelles ou distributions) utilisées pour alimenter le modèle (ANSES 2013a). Cette classe d'incertitude se décline en trois sous-classes :</p>	
- Représentativité des données	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées à la représentativité des données au regard de la population cible</i>
- Prise en compte de la variabilité	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées à la prise en compte de la variabilité des variables intervenant dans l'évaluation au regard de la population cible</i>
- Précision des mesures des variables	<i>Sous-classe qui recouvre l'ensemble des sources d'incertitude liées à la précision des mesures des variables intervenant dans l'évaluation</i>

La matrice ayant servi à recenser les sources d'incertitude est schématisée au Tableau 7 . Elle a été définie selon le modèle développé par Walker *et al* (Walker *et al.* 2003). Les lignes de la matrice listent les sources d'incertitude identifiées, de manière structurée, selon leur *localisation*⁶, et les deux colonnes rapportent respectivement leur *nature*⁷ et leur *traitement*. Dans le contexte du GT PE, ces deux dimensions ont été définies de la façon suivante :

- la *nature* d'une source d'incertitude « *indique l'origine (ou cause) de l'incertitude. L'incertitude peut-être due à un manque de connaissance scientifique, à un manque de données d'observation, à un manque de robustesse méthodologique, à des limites météorologiques, à des hypothèses reposant sur des connaissances disponibles et sur lesquelles se fondent les scénarios ou encore à un cadre administratif ou réglementaire imposé.* » (ANSES 2013a). Cette définition va au-delà de la simple déclinaison « *variabilité* » *versus* « *limites des connaissances* ». Elle pourra être affinée ultérieurement pour définir une typologie de la nature de l'incertitude.
- le *traitement* d'une source d'incertitude « *décrit la manière dont l'incertitude a été prise en compte dans le processus d'évaluation, par exemple, un argumentaire pour justifier une hypothèse, une description de la gestion de données censurées, etc.* » (ANSES 2013a). Dans le cas présent, le terme « *traitement* » ne doit pas être compris dans le sens de « *résolution de la source d'incertitude* ». Ici, le *traitement* fait référence à la gestion d'une source d'incertitude – une gestion nécessaire – pour pouvoir produire des résultats en dépit des connaissances manquantes ou limitées. Par exemple, pour combler l'absence de données ou de connaissances, le jugement d'expert peut être engagé pour obtenir une distribution des valeurs pour un facteur d'entrée d'un modèle. Le fait d'utiliser le jugement d'expert permet de pallier le manque de connaissances pour calculer l'exposition, mais ne résout pas l'incertitude, qui peut subsister, même si elle change de nature (il ne s'agit plus de manque de données, mais d'une hypothèse).

La matrice d'incertitude développée pour le BPA a préalablement été complétée par les membres du CT « Incertitude » du GT PE et par les membres de l'équipe projet de l'Anses ayant travaillé sur le calcul de l'exposition. Cette première version de la matrice a été soumise au sous-groupe en charge de l'évaluation de l'exposition, le sous-groupe « Exposition » du GT PE, pour être discutée collectivement en séance, et aboutir à la version finale dont la synthèse figure dans le rapport d'expertise (ANSES 2013a).

Tableau 7 : Structure matricielle utilisée par le GT PE pour recenser et caractériser les sources d'incertitude relatives à l'évaluation de l'exposition au BPA (ANSES 2013a)

Localisation des sources d'incertitude en évaluation de l'exposition en vue d'une ERS	Critères de caractérisation des sources d'incertitude en évaluation de l'exposition en vue d'une ERS	
	Nature (*)	Traitement (**)
Sources d'incertitude contextuelles		
- <i>Choix des sources d'exposition</i>		
- <i>Choix des populations exposées</i>		
Sources d'incertitude des scénarios		
- <i>Population cible considérée</i>		
- <i>Source d'exposition considérée</i>		
- <i>Voie d'exposition considérée</i>		
- <i>Echelle de temps d'exposition</i>		
Sources d'incertitude des modèles		
- <i>Choix des variables du modèle</i>		
- <i>Choix de l'équation mathématique</i>		
- <i>Corrélation entre les variables</i>		
Sources d'incertitude des données		
- <i>Représentativité des données</i>		
- <i>Prise en compte de la variabilité</i>		
- <i>Précision des mesures des variables</i>		

(*) Indiquer s'il s'agit d'un « manque de connaissance », d'un « manque de données », d'une « limite d'échantillonnage », d'une « hypothèse », d'un « cadre administratif réglementaire », ou si c'est « sans objet ».

(**) Décrire le traitement au sein de l'évaluation

5.1.1.2 Limites

L'étude de cas du BPA a permis d'esquisser une première classification des sources d'incertitude dans l'évaluation de l'exposition et de dresser les lignes générales de la méthode d'analyse qualitative de l'incertitude. L'acquis le plus important a été sans doute les échanges avec les membres du sous-groupe exposition qui ont permis de vérifier la faisabilité des propositions théoriques.

La matrice construite comprenait, à l'origine, deux dimensions supplémentaires : l'*amplitude* et la *direction* des sources d'incertitude. La matrice des incertitudes et notamment ces deux dimensions devaient notamment être renseignées par chacun des membres du sous-groupe « Exposition » de façon systématique, individuelle et indépendante, avant d'être discutée collectivement en séance ; le résultat consensuel du sous-groupe « Exposition » devait ensuite être présenté et discuté collectivement en séance avec l'ensemble du GT PE. Les contraintes de temps n'ont pas permis de mener la démarche envisagée jusqu'au bout.

Le caractère novateur de la démarche d'évaluation qualitative de l'incertitude proposée à l'Anses vient de son caractère systématique et générique. Les travaux relatifs au BPA ont servi de cas test pour calibrer et vérifier la robustesse de la démarche méthodologique en cours de construction, ceux relatifs au toluène ont permis de vérifier sa reproductibilité, ils ne doivent pas être considérés comme une fin en soi. Les sources d'incertitude à caractériser et à évaluer étant toujours les mêmes d'une ERS à une autre, une méthode générique est utile dans la mesure où la comparaison de l'évaluation de l'exposition entre plusieurs substances sera facilitée. Par ailleurs, documenter les mêmes classes d'incertitude permettra de sauvegarder explicitement des choix d'expertise faits pour une substance et, le cas échéant, de les utiliser pour les substances suivantes.

5.1.2 Prise en compte quantitative de l'incertitude dans la caractérisation de l'exposition

5.1.2.1 Présentation de la démarche

Lors de la caractérisation de l'exposition au BPA, les experts ont choisi de modéliser les doses d'exposition selon une approche probabiliste pour une prise en compte optimale de la variabilité et de l'incertitude associées aux variables d'entrée des modèles.

Différents médias – air, aliments, eau et poussières sédimentées – ont été considérés au regard de l'étendue relativement large d'utilisation et d'application de la substance, dans des consommables de la vie courante, les mettant potentiellement en jeu (ANSES 2013a). Les doses internes d'exposition ont été modélisées séparément pour chaque média, puis agrégée en une dose interne totale, somme des doses estimées pour chaque média.

Pour chaque modèle utilisé, une distribution de probabilité a été attribuée à chaque variable d'entrée. La variabilité et l'incertitude ont ensuite été propagées – sans séparation – selon la méthode de Monte Carlo 1D (10000 itérations). Les résultats ont été représentés sous la forme d'histogrammes ou de distribution cumulées des doses d'exposition, pour en déduire le pourcentage de dépassement des doses de référence toxicologiques. Une analyse de sensibilité a, par ailleurs, été réalisée, en utilisant l'indice des rangs de Spearman, en vue d'identifier et de hiérarchiser les variables influant le plus l'exposition. Ce travail a principalement été mené en utilisant le logiciel @Risk.

S'agissant de la définition des distributions des variables d'entrée, point critique de l'approche probabiliste, une stratégie a été adoptée. Cette stratégie a été définie en fonction du niveau d'information disponible. Les données dont on dispose pour documenter les variables d'entrée sont, en effet, de qualité et de quantité très hétérogènes d'une variable à une autre. Elles sont généralement issues de sources diverses : littérature, enquêtes de filière, données de mesures, etc. Dans certains cas, elles permettent de renseigner un intervalle de variation, dans d'autres cas, une valeur centrale voire un ensemble de percentiles sont précisés. Les jeux de données de mesures sont plus rarement disponibles. La stratégie adoptée pour définir les distributions de probabilité des variables tient compte de ces différentes situations et se décline en quatre options présentées ci-après :

- **Option 1** – Cas où seul un intervalle de variation est connu (ou estimé)

Pour le cas où, seul un minimum (m) et un maximum (M) sont documentés, une distribution uniforme sur [m ; M] est attribuée à la variable. L'intervalle [m ; M] est interprété comme l'ensemble des valeurs possibles (ou probables) de la variable.

- **Option 2** – Cas où un intervalle de variation et une valeur centrale sont connus (ou estimés)

Pour le cas où, le minimum (m), le maximum (M), et la valeur centrale (moyenne, médiane ou mode) d'une variable sont documentés, une distribution triangulaire sur [m ; M], de mode la valeur centrale, est attribuée à la variable.

- **Option 3** – Cas où un ensemble de percentiles est connu (ou estimé)

Pour le cas où plusieurs percentiles d'une variable sont renseignés, une fonction de répartition empirique est construite à partir des couples $(x_i; p_i)$ de données cumulatives disponibles, avec p_i la probabilité d'obtenir une valeur inférieure ou égale à x_i , où $i \in \{1, \dots, k\}$ et k le nombre de percentiles documentés.

- **Option 4** – Cas où un jeu de données de mesure est disponible

Il arrive que les données brutes d'une enquête soient disponibles. Dans cette situation, il est nécessaire d'avoir un pourcentage d'occurrence de chaque individu. Cette pondération est parfois spécifiée. Si ce n'est pas le cas, il est attribué, par défaut, un pourcentage d'apparition de $100/n$, n étant le nombre d'individu de l'enquête. Les différentes valeurs prises par la variable sont ensuite ordonnées, et si une même valeur est mesurée à plusieurs reprises, les probabilités d'apparition

sont sommées. Un ensemble de couples $(x_i ; p_i)$ – avec $\{x_1, x_2, \dots, x_m\}$ comme l'ensemble des modalités possibles du paramètre, et $\{p_1, p_2, \dots, p_m\}$ comme ensemble des probabilités d'occurrence respectives – est ainsi obtenu. Cet ensemble définit une distribution empirique discrète. Parfois, une étape supplémentaire de regroupement des issues possibles par classes est nécessaire afin d'éviter un trop grand nombre de modalités différentes avec un pourcentage d'occurrence associé faible.

5.1.2.2 Limites

La stratégie mise en œuvre pour définir les distributions de probabilité utilisées pour l'ERS du BPA souligne un effort de standardisation. Cette stratégie pourrait cependant être améliorée. Ainsi, s'agissant des options 1 et 2, d'autres lois que les lois uniforme ou triangulaire pourraient être proposées. Par exemple, si l'on s'intéresse à un ratio qui pourrait être compris entre 0,001 et 1000, une loi log-uniforme pourrait être plus pertinente qu'une loi uniforme, tel que cela est indiqué dans l'option 1, car de cette façon, la probabilité de se trouver dans l'intervalle entre 0,001 et 0,01 serait la même que de se trouver dans l'intervalle entre 100 et 1000. Alors que si la loi était uniforme, ce serait la même que de se retrouver dans l'intervalle 999,991 et 1000. S'agissant des options 3 et 4, si les lois de distribution pour un paramètre sont connues dans d'autres contextes (par exemple, construites à partir d'autres jeux de données plus fournis obtenus dans d'autres pays ou à d'autres époques), recourir à ces lois apparaît être une meilleure solution, si les données disponibles le permettent et s'ajustent à ces lois. L'avantage de recourir à des lois théoriques plutôt qu'à des distributions empiriques discrètes est que l'incertitude sur l'estimation des valeurs des paramètres de la loi théorique peut être évaluée à partir de l'ajustement aux données, et entrer en considération dans l'analyse quantitative de l'incertitude.

L'analyse de sensibilité peut aussi être améliorée. L'utilisation de l'indice des rangs de Spearman est simple, mais critiquable pour réaliser une analyse de sensibilité qui tienne compte des interactions entre paramètres. Le calcul des indices de Sobol est, quant à lui, adapté à tout type de modèle, de complexité variable, et son implémentation, qui nécessite un nombre important de calculs, ne pose pas de problème. En effet, des logiciels sont disponibles et le temps de calcul est raisonnable avec les moyens informatiques actuels.

5.2 Analyse d'incertitude dans l'évaluation du danger

L'IPCS définit l'évaluation du danger comme un processus dont l'objectif est de déterminer les effets adverses possibles d'un agent ou d'une situation auquel un organisme, un système ou une (sous)population est exposé. Elle comprend deux volets : l'identification du danger et la caractérisation de la relation dose-réponse (IPCS 2004). L'évaluation du danger implique une classification de la toxicité de l'agent (ou de la situation) considéré, sur la base de l'information disponible dans la littérature. Celle-ci amène à considérer toutes les incertitudes, qu'elle soit quantifiable ou non. Dans la caractérisation de la relation dose-réponse, les sources d'incertitude sont semblables à celles rencontrées au cours de l'évaluation de l'exposition, elles peuvent donc être traitées selon les mêmes méthodes. S'agissant de l'identification du danger, la question de l'incertitude se pose différemment. Les incertitudes auxquelles elle fait face portent sur la relation de cause à effet entre l'exposition humaine à un agent et la survenue d'un effet sanitaire particulier. L'enjeu de cette étape est de s'assurer de l'exactitude de la classification établie pour chaque effet sanitaire considéré, en s'appuyant sur une démarche cohérente et transparente, et non un traitement formel de l'incertitude ; l'approche habituellement utilisée consiste à apprécier les éléments de preuve (*lines of evidence*) (NRC 2009). L'incertitude y est considérée dans l'appréciation du risque de biais des études et l'expression du niveau de confiance final. Les approches qualitatives sont celles généralement utilisées bien que des développements quantitatifs existent (Linkov et al. 2009).

5.2.1 Analyse des données de la littérature pour l'évaluation du danger – Poids de la preuve & Incertitude

Dans le cadre d'une ERS, l'existence ou non d'une relation de cause à effet est établie sur la base de l'information disponible dans la littérature, laquelle provient de différentes études *in vivo*, *in vitro*, populationnelles, etc. L'ensemble de ces informations constitue des éléments de preuve²⁸ (*lines of evidence* en anglais) à la réponse finale. Aujourd'hui, l'évaluation de leur contribution à la réponse finale, *i.e.* l'évaluation du poids des preuves, repose souvent sur le jugement d'expert (SCENIHR 2012). Or la manière dont le jugement d'expert est établi puis utilisé n'est pas toujours clairement explicitée. L'objet d'une démarche d'appréciation des éléments de preuve (*WOE approach*) est d'apporter plus de transparence au processus d'expertise en explicitant les forces et les faiblesses des éléments considérés. Elle contribue, de cette façon, à la comparabilité et à la cohérence des résultats d'ERS entre eux (NTP 2015).

Plusieurs guides méthodologiques sur le sujet ont été publiés ces dernières années (Hart et al. 2010, NTP 2015, SCENIHR 2012). Les approches proposées sont relativement similaires et reprises, au moins en partie, du système d'évaluation GRADE²⁹ (*Grading of Recommendations Assessment, Development, and Evaluation*) (Guyatt et al. 2011) et de l'approche développée par l'AHRQ (*Agency for Healthcare Research and Quality*) (Schünemann et al. 2012) pour les besoins des études cliniques. Les approches *WOE* reposent sur une démarche d'appréciation systématique de différents éléments de preuve, structurée par étape. On retrouve généralement les sept étapes suivantes :

Étape 0 – Planification du processus d'évaluation du niveau de preuve

Étape 1 – Identification des sources possibles des données et du manque de données en relation avec l'objectif de l'ERS.

Étape 2 – Screening initial de ces sources de données pour identifier celles qui sont pertinentes pour la question posée.

Étape 3 – Appréciation individuelle des études.

Étape 4 – Appréciation individuelle des éléments de preuve.

Étape 5 – Appréciation globale des éléments de preuve.

Étape 6 – Caractérisation des incertitudes d'intérêt.

Étape 0 – Planification du processus d'évaluation du poids de la preuve

Le NTP conçoit cette étape comme une phase préliminaire visant à formaliser le cadre de l'évaluation et ses objectifs (NTP 2015). Il s'agit concrètement de rédiger un protocole détaillé précisant la stratégie d'évaluation choisie et les aspects d'ordre méthodologique, tels que la stratégie de recherche bibliographique, les critères de sélection des études, la méthode d'évaluation des données (humaines, animales, etc.), etc. Une modification du protocole au cours de l'évaluation est possible, elle doit cependant être documentée.

²⁸ En évaluation des risques, un **élément de preuve** (*line of evidence* en anglais) est un ensemble d'informations disponibles, idéalement publiées, issues de différentes études *in vivo*, *in vitro*, etc., sur lesquelles le processus d'évaluation fonde ses conclusions sur l'existence ou non d'une relation de cause à effet.

²⁹ L'approche GRADE propose un système de cotation du niveau de la preuve utilisable notamment dans le cadre d'une revue systématique. Il est toutefois particulièrement adapté à la recherche clinique. Une série d'articles publiés dans la revue « *Journal of Clinical Epidemiology* » en fait une description détaillée. Le NTP s'y réfère dans le cadre de sa propre approche de l'analyse du poids de la preuve.

Étape 1 – Identification des sources possibles des données et du manque de données en relation avec l'objectif de l'ERS

Pour identifier les sources possibles des données, le SCENIHR recommande de procéder à une recherche par élément de preuve (SCENIHR 2012). Il est demandé de privilégier les publications d'origine aux sources de données secondaires telles que les revues systématiques ou les méta-analyses. Le recours aux données confidentielles ou non publiées est possible, en complément de données publiées. Dans tous les cas, les bases de données bibliographiques et les mots clés utilisés doivent être précisés (NTP 2015).

Étape 2 – Screening initial de ces sources de données pour identifier celles qui sont pertinentes pour la question posée

L'objet de cette étape est de réaliser une première sélection des sources de données disponibles (publications). À ce jour, il n'y a pas de critère consensuel pour l'acceptation ou non d'une étude ; ceux utilisés doivent être précisés dans le protocole (NTP 2015). Le SCENIHR préconise, toutefois, de considérer les études épidémiologiques et les études expérimentales de façon similaire, et d'apprécier les études « positives » et « négatives » selon la même procédure, les mêmes critères, dès lors qu'elles présentent le même niveau de qualité (SCENIHR 2012).

Pour les publications retenues, le NTP demande à ce que l'extraction des données s'effectue selon la procédure d'assurance qualité spécifiée dans le protocole (NTP 2015). Les publications ne vérifiant pas les conditions de sélection doivent être référencées avec la mention « Publication relevée mais non considérée comme pertinente pour la question posée » (SCENIHR 2012).

Étape 3 – Appréciation individuelle des études

Cette étape s'intéresse essentiellement à la qualité des études sélectionnées à l'issue du screening (Étape 2). Le SCENIHR comme le NTP procèdent par type d'étude (étude de cohorte, étude Cas-Témoins, essai clinique randomisé, etc.) avec, cependant, des nuances entre les démarches.

L'approche du SCENIHR repose sur l'utilisation de deux critères : la qualité³⁰ scientifique globale et la pertinence³¹ de l'étude pour la question posée, définies selon les critères de Klimisch (Klimisch, Andreae, and Tillmann 1997). Les études « positives » et « négatives » doivent être évaluées selon la même procédure. La significativité statistique des résultats ne doit pas être le seul facteur considéré ; en particulier, les différents éléments clés suivants ne doivent pas être négligés : la force de l'association, la cohérence des résultats, la comparabilité des sous-groupes, la forme de la relation dose-réponse, etc. Au final, seules les publications de bonne qualité scientifique ou de qualité scientifique adéquate, ayant un intérêt direct ou indirect pour l'ERS sont retenues pour l'étape suivante (Étape 4). Les publications non retenues doivent cependant être

³⁰ Qualité d'une étude (Traduit du SCENIHR (2012)) :

- *Bonne qualité scientifique* – Le protocole de l'étude, sa conduite et la manière dont elle a été rapportée sont considérés comme appropriés, et la méthode utilisée, valide.
- *Qualité scientifique adéquate mais avec des limites significatives* – L'étude est acceptable scientifiquement avec des écueils importants en terme de protocole et/ou de conduite et/ou dans la manière de rapporter les résultats.
- *Qualité scientifique inadéquate* – Le protocole et/ou la conduite de l'étude présentent de sérieuses faiblesses
- *Non qualifiable* – Les informations disponibles ne permettent pas d'apprécier la qualité de l'étude

³¹ Pertinence d'une étude (Traduit du SCENIHR (2012)) :

- *Intérêt direct* – L'étude traite de l'agent, du modèle **ET** de la sortie considérés
- *Intérêt indirect* – L'étude traite de l'agent, du modèle **OU** de la sortie considérés
- *Intérêt insuffisant* – L'étude ne peut pas être utilisée pour répondre à la question posée

référéncées avec la mention « Publication relevée mais non considérée comme appropriée pour la question posée » (SCENIHR 2012).

La démarche du NTP s'appuie sur celle de l'AHRQ et s'attache à apprécier le risque de biais à l'aide d'une série de questions standards. Cette série de questions vise à évaluer le protocole de l'étude et sa performance ; elle couvre cinq aspects différents (sélection, performance, attrition, détection et signalement de biais), avec la particularité pour certaines questions d'être spécifiques à un type d'étude (AHRQ 2012a). Pour chaque résultat d'étude, les questions sont posées avec quatre réponses possibles en fonction du risque de biais : « certainement faible », « probablement faible », « probablement fort », « certainement fort » (Guyatt 2012). Contrairement au SCENIHR, le NTP ne définit pas de critère standard pour la sélection des études destinées à l'étape 4 ; ce ou ces critères doivent être *ad-hoc* à chaque évaluation et être clairement définis dans tout protocole d'évaluation de la preuve (NTP 2015).

Étape 4 – Appréciation individuelle des éléments de preuve

La présente étape s'appuie exclusivement sur les études sélectionnées à l'issue de l'étape 3. L'objectif est de réaliser, pour chaque élément de preuve, une appréciation intégrée des études sur lesquelles cet élément se fonde. Cette appréciation s'effectue par type d'étude (humaine, animale, *in vitro*, etc.). Diverses démarches existent, celle choisie par le NTP consiste à évaluer le niveau de confiance des éléments de preuve sur lesquelles repose l'ERS. Cette démarche reprend la démarche GRADE (Balshem et al. 2011) et la démarche de l'AHRQ (AHRQ 2012b), conceptuellement similaires. Ces approches utilisent une échelle de confiance à quatre niveaux³² : « confiance élevée », « confiance modérée », « confiance faible » et « confiance très faible ». Pour définir un niveau de confiance dans un élément de preuve, un niveau de confiance initial est tout d'abord attribué, en fonction de la capacité des études sur lesquelles il se fonde à mettre en évidence la relation de causalité. Cette capacité est établie d'après les critères de Bradford-Hill notamment (Hill 1965, Schünemann et al. 2012). Le niveau de confiance initial est ensuite augmenté ou, au contraire, diminué en présence de facteurs qui accroissent ou réduisent le niveau de confiance des résultats (NTP 2015). Au final, le niveau de preuve est obtenu en combinant la confiance globale dans l'existence (ou non) d'un effet sanitaire, et la direction de l'effet (toxique ou non toxique)³³.

³² Échelle de confiance (Traduit du (NTP 2015)) :

- **Confiance élevée (++++)** dans l'association entre l'exposition à une substance et le résultat. Il est hautement vraisemblable que l'effet réel soit reflété par la relation apparente.
- **Confiance modérée (+++)** dans l'association entre l'exposition à une substance et le résultat. L'effet réel peut être reflété par la relation apparente.
- **Confiance faible (++)** dans l'association entre l'exposition à une substance et le résultat. L'effet réel est vraisemblablement différent de la relation apparente.
- **Confiance très faible (+)** dans l'association entre l'exposition à une substance et le résultat. Il est hautement vraisemblable que l'effet réel soit différent de la relation apparente.

³³ Échelle du niveau de preuve (Traduit du NTP (2015))

- **Niveau de preuve élevé** : il y a une confiance élevée dans l'élément de preuve pour une association entre l'exposition à la substance et le(s) résultat(s) sanitaire(s).
- **Niveau de preuve modéré** : il y a une confiance modérée dans l'élément de preuve pour une association entre l'exposition à la substance et le(s) résultat(s) sanitaire(s)
- **Niveau de preuve faible** : il y a une confiance faible dans l'élément de preuve pour une association entre l'exposition à la substance et le(s) résultat(s) sanitaire(s) ou il n'y a pas de donnée disponible.
- **Preuve d'une absence d'effet sanitaire** : il y a une confiance élevée dans l'élément de preuve pour une absence d'association entre l'exposition à la substance et le(s) résultat(s) sanitaire(s).

À la différence du NTP, le SCENIHR commence par classer chaque étude dans l'une des trois classes suivantes : « Indique la présence d'un effet », « Indique l'absence d'un effet », « Est cohérent avec la présence comme avec l'absence d'un effet » (SCENIHR 2012). Cette classification s'effectue à l'aide d'une grille d'analyse *ad-hoc* au type d'étude. L'appréciation intégrée s'effectue ensuite selon deux critères : la *cohérence*³⁴ et l'*utilité*³⁵. Cette analyse conduit *in fine* à classer les études dans l'une des deux catégories suivantes :

- « Publication d'intérêt, de qualité suffisante (adéquate) et importante pour répondre à la question posée »
- « Publication d'intérêt, de qualité suffisante (adéquate) mais non jugée nécessaire pour répondre à la question posée »

Seules les études jugées importantes pour la question posée sont retenues pour la dernière étape.

Étape 5 – Appréciation globale des éléments de preuve

Le but de cette étape est de conclure sur l'identification ou non d'un danger, en s'appuyant sur les appréciations individuelles des éléments de preuve établis à l'étape précédent (NTP 2015, SCENIHR 2012). Cette conclusion est formulée par effet ou par groupe d'effets sanitaires ou biologiques, en fonction des objectifs de l'ERS et/ou des données disponibles. Elle s'appuie habituellement sur une classification qui intègre clairement les niveaux de preuve pour les différents types d'études considérés :

- **Reconnu dangereux pour l'être humain** : niveau de preuve élevé d'après les études chez l'être humain (quelles que soient les conclusions issues des autres types d'études)
- **Présumé dangereux pour l'être humain** : (niveau de preuve modéré d'après les études chez l'être humain ET niveau de preuve élevé ou modéré d'après les autres types d'études) OU (niveau de preuve faible d'après les études chez l'être humain ET niveau de preuve élevé d'après les autres types d'études)
- **Suspecté dangereux pour l'être humain** : (niveau de preuve modéré d'après les études chez l'être humain ET niveau de preuve faible d'après les autres types d'études) OU (niveau de preuve faible d'après les études chez l'être humain ET niveau de preuve modéré d'après les autres types d'études)
- **Non classé ou non identifié comme dangereux pour l'être humain** : niveau de preuve faible d'après les études chez l'être humain ET niveau de preuve faible d'après les autres types d'études

³⁴ La **cohérence** d'un élément de preuve est définie comme l'adéquation des résultats entre les différentes études sur lesquelles il s'appuie (SCENIHR 2012). Trois niveaux de cohérence sont distingués :

- *Fort* – La majorité des études présente des résultats concordants
- *Moyen* – La majorité des études présente des résultats soit allant dans la même direction, soit cohérents
- *Faible* – Peu de concordance entre les études

³⁵ L'**utilité** tient compte de la **qualité** et de la **pertinence** (SCENIHR 2012). Trois niveaux d'utilité sont distingués :

- *Forte* – La majorité des études présente un « intérêt direct » et « bonne qualité scientifique »
- *Moyenne* – La majorité des études présente soit un « intérêt indirect » et « bonne qualité scientifique », soit un « intérêt direct » et « qualité scientifique adéquate »
- *Faible* – La majorité des études présente un « intérêt indirect » et « qualité scientifique adéquate »

Étape 6 – Caractérisation de l'incertitude

L'incertitude est un élément important dans l'appréciation des éléments de preuve, « degré d'incertitude » et « niveau de preuve » peuvent d'ailleurs être vus comme deux mesures d'appréciation des éléments de preuve, inversement liées. Cependant, parce que le degré d'incertitude apporte un éclairage différent et complémentaire au niveau de preuve, son expression et sa communication ne doivent pas être négligées mais, au contraire, venir compléter celles du niveau de preuve. Trois grands modes d'expression différents sont principalement utilisés : (1) qualitative, fondée sur une échelle en quelques niveaux. Les échelles utilisées sont, en général, reprises ou adaptées de celles de l'IPCC, l'une exprimant le niveau de confiance des résultats (cf. Tableau 8) et l'autre exprimant la vraisemblance des résultats (cf. Tableau 9) (Hart et al. 2010, IPCC 2005, SCENIHR 2012) ; (2) tabulée, utilisant une matrice d'incertitude pour une caractérisation plus complète (cf. Chapitre 3) ; et, (3) quantitative, fondée sur des méthodes probabilistes, utilisée dans le cadre d'une évaluation probabiliste du risque.

Tableau 8 : Niveau de confiance quantitativement calibré – Traduit de (IPCC 2005)

Terminologie	Degré de confiance des résultats
Confiance très élevée	Au moins 9 chances sur 10 que les résultats sont corrects
Confiance élevée	Environ 8 chances sur 10
Confiance moyenne	Environ 5 chances sur 10
Confiance faible	Environ 2 chances sur 10
Confiance très faible	Moins de 1 chance sur 10

Tableau 9 : Échelle de vraisemblance – Traduit de (IPCC 2005)

Terminologie	Vraisemblance des résultats
Quasiment certain (<i>Virtually certain</i>)	> 99% de probabilité d'occurrence
Très vraisemblable (<i>Very likely</i>)	> 90% de probabilité
Vraisemblable (<i>likely</i>)	> 66% de probabilité
Vraisemblance indéfinie (<i>About as likely as not</i>)	33% à 66% de probabilité
Invraisemblable (<i>Unlikely</i>)	< 33% de probabilité
Très invraisemblable (<i>Very unlikely</i>)	< 10% de probabilité
Exceptionnellement invraisemblable (<i>Exceptionally unlikely</i>)	< 1% de probabilité

5.2.2 Prise en compte quantitative de l'incertitude dans la caractérisation du danger

Pour tenir compte des incertitudes dans la caractérisation des dangers, des méthodes quantitatives existent : elles vont d'approches déterministes simples à des approches plus sophistiquées de type probabiliste (cf. Chapitre 4). Par exemple, lors de la construction des valeurs toxicologiques de référence, la variabilité et l'incertitude liée aux limites de la connaissance sont prises en compte *via* des facteurs multiplicatifs³⁶ (souvent entre 3 et 10). Selon le type d'approche mise en œuvre, déterministe ou probabiliste, une valeur ou une distribution de probabilité est attribuée à chacun de ces facteurs. Ainsi, l'approche probabiliste permet une prise en compte plus complète de la variabilité et de l'incertitude sous-jacentes.

³⁶ En fonction du contexte (réglementaire, etc.), ces facteurs multiplicatifs sont appelés « facteur de sécurité (safety factor) », « facteur d'ajustement (*adjustment factor*) », « facteur d'incertitude (*uncertainty factor*) » ou encore plus récemment « facteur d'évaluation (*assessment factor*) »

L'IPCS a publié en 2014 un document guide pour la prise en compte de l'incertitude dans la caractérisation du danger des substances chimiques (IPCS 2014). La suite de ce paragraphe en fait la synthèse, ce document étant, à ce jour, le plus complet sur les aspects quantitatifs des incertitudes dans la caractérisation du danger.

5.2.2.1 Principe général

L'idée générale de la démarche de l'IPCS est de caractériser la relation dose-effet et son extrapolation à l'être humain sous forme de distributions statistiques pour, au final, traiter l'incertitude globale au moyen de méthodes probabilistes. Le but est d'améliorer la visibilité de l'incertitude pour permettre à l'évaluateur et au gestionnaire du risque de mieux juger de son acceptabilité, pour éventuellement mettre en place de nouvelles études expérimentales ou des approches statistiques plus adaptées.

La caractérisation du danger se ramène à un problème d'inférence sur la population cible de l'ERS, à partir d'informations sur l'effet sanitaire considéré. L'IPCS s'appuie, pour cela, sur différents concepts : la *sévérité* (ou *amplitude*) S d'un effet sanitaire et son *incidence* (ou *occurrence*) dans la population cible. Selon l'effet sanitaire considéré, S peut être exprimée de manière catégorielle (exemple : lésions hépatiques modérées) et/ou de manière continue (exemple : 5% de réduction du nombre de globules rouges). La dose humaine d'intérêt à estimer, notée DH_S^I , est définie comme la dose à laquelle une fraction I de la population cible présente l'effet sanitaire considéré à une sévérité S . Par exemple, si l'effet sanitaire considéré est la réduction de globules rouges, $DH_{0.05}^{0.01}$ correspond à la dose pour laquelle 1% de la population cible présente une réduction d'au moins 5% du nombre de globules rouges. Ainsi, l'objectif de la caractérisation du danger revient *in fine* à protéger un pourcentage – généralement important – de la population cible, de l'effet sanitaire considéré, à une sévérité donnée. En vue de l'extrapolation, le concept d'équipotence³⁷ est lui aussi introduit. Enfin, la démarche proposée par l'IPCS permet de différencier la variabilité de l'incertitude due aux limites des connaissances.

Les aspects que l'IPCS considère dans la caractérisation des dangers, comme pouvant générer une incertitude, sont les suivants (en fonction de l'étude et de la cible considérées) :

- La détermination d'un point de départ, *i.e.* de la dose à partir de laquelle s'effectue l'inférence (généralement, une NOAEL ou une Benchmark Dose (voir plus loin) à partir de l'étude de référence).
- L'extrapolation entre espèces, l'objectif de cet aspect étant d'inférer sur la toxicité humaine à partir des études de toxicité dans différentes autres espèces.
- La prise en compte de la variabilité de sensibilité au sein de la population humaine cible.
- L'extrapolation concernant la durée d'exposition, en cas d'utilisation d'une étude court-terme pour une évaluation du risque long-terme.
- L'extrapolation de voie à voie, par exemple, lors de l'utilisation de données obtenue par ingestion pour traiter des effets inhalés.
- L'extrapolation d'un scénario d'exposition à un autre, par exemple, lorsqu'on se ramène à une exposition continue, à partir d'un essai avec une exposition 8 heures par jour, 5 jours sur 7.
- L'extrapolation aux doses de plus faibles amplitudes d'effets (*low-effect levels* en anglais), l'objectif de cet aspect étant d'inférer sur les doses associées aux effets de sévérité plus faible que ceux observables (ou observés) dans les études de toxicité.
- Le manque d'études de toxicité (critères d'effet considérés, niveaux de dose).
- L'extrapolation entre substances si des approches structure-activité de type « read across » ou QSAR sont utilisées.

³⁷ Deux doses sont équipotentes si elles induisent un effet de même amplitude chez deux individus soit appartenant à deux espèces différentes, soit de sensibilités différentes au sein d'une même espèce.

- L'extrapolation entre métriques d'exposition, par exemple, lorsqu'une étude avec exposition au plomb par l'eau est utilisée pour une évaluation du danger sur la base de la plombémie.
- L'extrapolation entre des marqueurs précoces et des événements, par exemple, l'utilisation du poids à la naissance plus faible pour rendre compte d'une surmortalité néonatale.
- L'extrapolation d'un point de départ *in vitro* à un point de départ à l'échelle de l'organisme.

Aujourd'hui, ces aspects sont le plus souvent traités de façon déterministe sous la forme de facteurs de sécurité³⁶ prenant des valeurs fixes (entre 1 et 10), choisies sur la base de jugements d'experts et censées être protecteurs dans la plupart des cas. L'approche recommandée par l'IPCS revient à remplacer l'utilisation de valeurs fixes par des distributions de probabilité. L'IPCS propose, pour certains de ces facteurs, des distributions par défaut fondées sur un historique de valeurs (par exemple, des valeurs pour plusieurs substances testées chez différentes espèces pour évaluer les différences de sensibilité entre espèces).

5.2.2.2 Démarche proposée

Etape 1 : définir le point de départ

Le document de l'IPCS recommande de déterminer, dans la mesure du possible, le point de départ par le calcul d'une BMD (ou « dose repère » en français). Une BMD correspond à une dose donnant lieu à un effet donné, non nul, par rapport à une situation témoin. Il présente l'avantage de tenir compte de l'ensemble de la relation dose-réponse. Lorsque cela n'est pas possible, il est proposé d'établir une distribution du point de départ à partir d'une valeur de NOAEL, en tirant partie de l'historique des ratios NOAEL/BMDL recueillis par l'EFSA (EFSA 2009) pour différents critères d'effets (poids, poids du foie, poids du rein, nombre de globules rouges) ou se référant à une étude d'Allen et al. (1994), pour les effets sur le développement fraction de fœtus affectés (Allen et al. 1994).

Etape 2 : Evaluer l'incertitude liée à chaque extrapolation

Plusieurs extrapolations sont examinées dans le document de l'IPCS, qui propose des distributions par défaut :

- *Extrapolation entre espèces* : C'est l'article de Bokkers and Slob (2007) qui sert de référence, en proposant une mise en relation des équivalents de dose d'absorption à partir de relations allométriques (ratios de poids à la puissance 0.7, avec un intervalle à 95% de 0.66-0.74) (Bokkers and Slob 2007).
- *Variabilité de la sensibilité au sein de l'espèce humaine* : Le document se réfère essentiellement à l'étude de Hattis et Lynch (2007) qui a cherché à fournir des informations sur la variabilité de sensibilité chez l'être humain, à partir de données d'études cliniques (Hattis and Lynch 2007).
- *Autres aspects* : Pour les autres aspects, notons notamment la difficulté de fournir des règles simples pour l'extrapolation de voie à voie qui réclame, sinon un modèle toxicocinétique complet, un minimum de données de cinétique (Mielke et al. 2011, Chiu and White 2006).

Etape 3 : calcul de l'incertitude globale

L'IPCS propose trois options pour le calcul d'incertitude globale :

- Un calcul non probabiliste qui consiste à multiplier les bornes inférieures et supérieures des intervalles de confiance (processus très majorant) ;
- Un calcul quasi-probabiliste qui consiste à combiner les distributions entre elles (ce qui est immédiat si ne sont considérées que des distributions log-normales) ;

- Un calcul probabiliste complet, qui peut inclure des modèles (PBTK ou autres), un bootstrapping³⁸, sur les données de références, bref, plus de souplesse dans la réalisation.

Les incertitudes secondaires, *i.e.* sur les paramètres des distributions d'incertitudes eux-mêmes, sont évoquées, mais complexes à aborder. Le document propose un tableau avec une analyse qualitative sur ces incertitudes.

En fonction du résultat du calcul d'incertitude globale, il peut être décidé soit d'en rester là si le degré d'incertitude est jugé suffisant pour l'évaluation des dangers et des risques, soit de générer de l'information nouvelle pour réduire l'incertitude.

5.2.2.3 Limites de la démarche

S'agissant du principe général de la méthode proposée par l'IPCS, il est nécessaire pour caractériser le danger de spécifier le degré de sévérité considéré comme « acceptable » pour une fraction fixée de la population. Ceci n'a rien d'évident et demande sans doute une réflexion et une documentation spécifique pour, à la fois, relier un niveau d'effet à une réponse sanitaire « acceptable » et quantifier un pourcentage « acceptable » de la population qui serait affectée.

Les choix en matière de sévérité et d'occurrence sont parfois reliés aux valeurs de NOAEL. Ceci permet certes de faire le lien entre les évaluations sanitaires passées, basées sur ces NOAELs, et les évaluations futures, mais cela évacue la question des choix de sévérité et d'occurrence sans réelle justification sanitaire. Ainsi l'EFSA recommande, pour la mortalité, une BMD 10%, considérée comme proche en moyenne de la NOAEL, même si, dans les études humaines, une BMD de 1% a déjà été appliquée avec un nombre de sujets suivis important (exemple : étude aflatoxine) (EFSA 2009).

S'agissant du point de départ, au travers de lignes directrices et d'avis, il apparaît que les BMDs sont de plus en plus utilisées dans les instances nationales et internationales d'expertise en évaluation des risques. En particulier, l'Anses préconise, lorsque cela est possible, l'utilisation de la modélisation de la relation dose réponse avec la détermination d'une BMD, pour la construction des VTR (à seuil/ sans seuil) (AFSSET 2007, 2010).

Utiliser une distribution pour passer d'une NOAEL à une BMD, de la façon proposée par l'IPCS présente plusieurs inconvénients. Tout d'abord, les avantages de l'utilisation d'une BMD plutôt que d'une NOAEL ne sont pas conservés. Ainsi, ni l'ensemble de la dose-réponse ni l'incertitude qui pourrait être liée à un faible nombre d'animaux ou à un choix inadéquat des concentrations d'expositions ne sont prises en compte. En outre, pour utiliser une distribution connue de ratios entre NOAEL et BMD, il est nécessaire de s'assurer que l'étude considérée est du même niveau de qualité que les études qui ont servi à déterminer la distribution. Une autre limite provient du domaine d'application de la distribution. Ainsi, celle proposée par l'IPCS concerne uniquement des études extraites du NTP américain, exploitées dans l'article de Bokkers et Slob (2007). Or, les 58 substances de cet article proviennent essentiellement d'études de substances pharmaceutiques ou d'additifs alimentaires, et uniquement de données rongeurs après 90 jours d'exposition pour quelques critères d'effets (poids, poids du foie, poids du rein, nombre de globules rouges), ce qui limite la généralisation à d'autres substances/critères d'effets (Bokkers and Slob 2007). Enfin, le document de l'IPCS indique qu'il n'y avait pas de différence entre les critères d'effet dans les données de l'EFSA ce qui n'est indiqué, ni dans le document de l'EFSA ni dans l'article d'origine. La Table B2 de l'article indique même le contraire.

S'agissant des extrapolations, le choix des distributions par défaut proposées l'IPCS est critiquable. Pour l'extrapolation entre espèces, cette extrapolation n'est justifiée que pour passer de la souris au rat, pour la voie alimentaire, sur les critères d'effet et les substances de l'étude. Pour l'extrapolation au sein de l'espèce humaine, les faiblesses de l'étude d'Hattis et Lynch (2007)

³⁸ Bootstrapping : technique d'inférence statistique basée sur une succession de ré-échantillonnage ((Efron and Tibshirani 1993)).

sont à la fois sa restriction aux substances pharmaceutiques et aux critères d'effets d'études cliniques et une collection d'études pour des individus relativement homogènes (groupes d'adultes sains, ou groupes d'enfants parfois), qu'il s'agisse des considérations toxicocinétiques ou toxicodynamiques (Hattis and Lynch 2007). Le point fort reste la distinction entre l'influence de la toxicocinétique et celle de la toxicodynamique dans la réponse, la première ayant un rôle moindre que la seconde. Ainsi, si on se réfère au facteur 10 habituellement utilisé en extrapolation, selon Hattis et Lynch (2007), un facteur de 1,9 serait lié à la toxicocinétique, et un facteur de 5,3 pour la toxicodynamique. Le document de l'IPCS propose l'utilisation d'une alternative moins protectrice en considérant une variabilité liée à la toxicodynamique équivalente à celle liée à la toxicocinétique en considérant que la différence précédemment observée serait liée à des erreurs de mesures plus importantes lorsque l'on considère des effets que des concentrations.

6 Développements méthodologiques pour le traitement de l'incertitude en ERS et applications

6.1 Démarche d'analyse des données pour l'identification du danger - Poids de la preuve et Analyse d'incertitude

Le niveau de confiance qui peut être attribué à l'issue de l'étape d'identification du danger d'une substance est lié à la nature des données disponibles sur les effets de cette substance, tant d'un point de vue qualitatif que quantitatif. La démarche suivie pour analyser ces données, en vue de caractériser le danger d'une substance, doit être la plus transparente et systématique. La question à laquelle devra répondre l'ERS dans cette étape porte généralement sur l'existence ou non d'un lien de causalité entre l'exposition d'une population d'intérêt à une substance donnée et un effet (adverse ou non). Pour ce faire, l'expertise devra s'appuyer sur l'analyse des données disponibles (*chez l'être humain, chez l'animal, etc.*) et se prononcer sur le niveau de preuve associé à ces données.

Afin de rendre cette étape la plus précise possible, il est préconisé de structurer l'analyse des données selon une démarche par étapes. Cette démarche s'inspire de celles proposées par d'autres organismes d'expertise présentées au chapitre précédent (cf. paragraphe 5.2.1) (NTP-OHAT, EFSA, GRADE²⁹, FERA, SCENIHR, *etc.*). L'objectif est à la fois d'organiser les données selon une approche convergente puis d'analyser en quoi elles permettent de répondre à la question posée et enfin, d'identifier les incertitudes résiduelles.

La démarche d'analyse proposée comporte les étapes suivantes : (1) reformulation de la (ou des) question(s) posée(s), (2) conduite de la recherche bibliographique, (3) analyse de la recevabilité des études et choix des critères d'inclusion (et par conséquent d'exclusion), (4) recueil et analyse des données issues des études prises de façon individuelle, (5) analyse globale de l'ensemble des données disponibles, (6) détermination du niveau de confiance associé à la base de données disponible, et (7) conclusion sur l'ensemble des niveaux de preuves.

Étape 1 – Formulation de la question posée

Il est important avant toute ERS de bien préciser le contexte de l'évaluation, la question posée, éventuellement les options de gestion dont l'impact pourra être évalué. Concernant plus spécifiquement l'identification du danger d'une substance, il est recommandé de formuler la ou les questions posées par l'expertise. Ces questions pourront inclure les composantes suivantes : nature de l'agent et conditions d'exposition, population concernée, effet investigué. Par exemple, dans le cas BPA, l'une des questions posées aurait pu être : « *une exposition au BPA pendant la grossesse peut-elle entraîner des effets sur le système reproducteur des enfants ?* ». Les préconisations mentionnées dans les étapes suivantes s'appliquent ainsi à chacune des questions posées.

Étape 2 – Recherche bibliographique

Cette étape vise à recenser l'ensemble du *corpus* de données sur lequel il faudrait asseoir l'analyse. En général, les agents étudiés sont très peu documentés, mais parfois un *corpus* bibliographique important existe. Par exemple dans le cas du BPA, plusieurs centaines d'études ont été publiées sur les effets de cette substance au cours des dernières années. Il convient donc dans ces situations de bien définir le périmètre d'analyse des études, en prenant en compte notamment le temps disponible, les moyens d'expertise mis en œuvre ainsi que la conduite éventuelle d'expertises antérieures reconnues qui pourront alors faire date. Ainsi c'est au cours de cette étape que le profil de recherche bibliographique devra être élaboré : mots clés, combinaisons, années encadrant la période d'étude – date de début et date de fin –, langue des publications retenues, nature des support (littérature internationale, littérature « grise »...), base

des données consultées (*PubMed, Toxnet, Scopus, etc.*). Pour certaines questions posées, une revue systématique de la littérature est nécessaire. Pour déterminer si la question posée nécessite une revue systématique, il est préconisé d'identifier les éléments clés. Par exemple, pour les questions sur l'effet d'une exposition, les éléments clés sont la population d'intérêt (P), l'exposition d'intérêt (E), un comparateur (exposition de référence – C) et le résultat (*outcome* en anglais – O) ; l'acronyme PECO y fait référence (EFSA 2010). Par ailleurs, ces éléments clés aident à définir les critères d'éligibilité des études, la stratégie de recherche bibliographique, *etc.* Leur identification est donc essentielle et devrait être le point de départ de toute planification de recherche bibliographique.

Étape 3 – Sélection des études à analyser

Une fois recensées toutes les études répondant aux critères définis ci-dessus, il est nécessaire de sélectionner les études qui feront l'objet d'une analyse approfondie. Les critères de recevabilité d'une étude ou d'exclusion devront être également explicités en amont de l'analyse. Certains de ces critères pourront être spécifiques à la substance ou au contexte étudié. En général, ils conduisent à écarter les études qui présenteraient des biais majeurs par rapport au contexte de l'expertise. Par exemple, dans le cadre du BPA, pour la question posée dans l'étape 1, les études pour lesquelles les expositions ne couvrent pas toute ou partie de la grossesse pourraient être écartées. Cette étape peut être réalisée uniquement à la lecture du titre et résumé des études et ne nécessite pas, en première intention, de lire l'ensemble de l'étude. Il est recommandé, pour cette étape, que deux experts jugent séparément de la recevabilité des études. Si leurs jugements divergent pour l'une d'entre elles, un consensus doit être recherché, en faisant éventuellement appel à un tiers externe.

Étape 4 – Analyser la qualité des données étude par étude

Une fois établie la liste des études à analyser en détail, il convient d'extraire de la publication ou du rapport d'étude les données utiles à l'expertise et de juger de leur qualité. Pour apprécier la qualité d'une étude ou le risque de biais des études, il existe différents outils tels que *ToxRTool*³⁹ (Schneider et al. 2009) et *SciRAP*⁴⁰ (Molander et al. 2015). Lorsque ces outils ne sont pas adaptés à la situation d'étude, il peut être utile d'établir des grilles *ad-hoc* d'analyse des études permettant de reporter de façon systématique les informations utiles à l'expertise. Lors de l'ERS du BPA, par exemple, la grille de lecture utilisée comportait la liste des informations suivantes : espèces, souche, sexe, voie d'exposition, nombre d'animaux, effets investigués, effets rapportés, *etc.* (ANSES 2013a). Des grilles différentes pourront être élaborées selon qu'elles s'adressent aux études expérimentales ou aux études épidémiologiques. Il est préconisé, pour cette étape, que deux rapporteurs se divisent les lectures et qu'ils soient mutuellement le relecteur de l'autre. Au-delà de certains critères généraux d'analyse de la qualité d'une étude (exemple des critères de Klimisch pour les études expérimentales), au cas par cas, les experts pourront proposer des critères spécifiques liés à la thématique d'intérêt (ex. métabolisme/Obésité, cancérogenèse mammaire *etc.*) qui permettront de mieux apprécier le niveau de preuve. Il s'agit, au final, de classer les études en quatre niveaux selon la nature des biais identifiés.

Étape 5 – Analyser le niveau de preuve sur la base de chaque étude puis sur l'ensemble du corpus

L'objectif de cette étape est d'attribuer aux résultats de chacune des études analysées dans l'étape précédente un niveau de preuve par rapport à la question posée, puis d'extrapoler à l'ensemble du *corpus* de données. Il est préconisé pour cette étape de s'appuyer en première intention sur les études jugées de meilleure qualité (par exemple, études expérimentales cotées

³⁹ <https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxrtool/toxrtool-toxicological-data-reliability-assessment-tool>

⁴⁰ <http://www.scirap.org/>

Klimisch 1 ou 2) et d'extraire de ces études les informations qui apportent une réponse à la question posée puis de les qualifier en termes de robustesse. Par exemple, un effet observé dans une étude expérimentale cotée Klimisch 1 aura d'emblée une note maximale de « ++++ », contre « +++ » pour une étude cotée Klimisch 2. Puis, selon, par exemple, l'amplitude de l'effet observé, le nombre d'animaux, la cohérence des résultats *etc.*, la note initiale pourra être augmentée (rajout de « + »), ou, au contraire, dégradée (retrait de « + ») (NTP 2015). Des tables d'analyse sont proposées pour cette étape par le NTP et d'autres organismes (GRADE guidelines) (Guyatt et al. 2011, NTP 2015). Il conviendrait de les tester sur des exemples concrets et le cas échéant de les adapter. Concernant plus spécifiquement les études épidémiologiques, il est préconisé d'appliquer la hiérarchisation *a priori*, selon le type d'études, que proposent le NTP ainsi que les évaluations GRADE : étude contrôlée > étude de cohortes > étude cas-témoins > observationnelle, *etc.* Puis, individuellement selon la nature et l'intensité des effets observés, le niveau de preuve peut être renforcé ou diminué.

Étape 6 – Convertir le niveau de confiance par type d'étude en niveau de confiance global sur l'ensemble du corpus disponible

Cette étape vise à transformer les scores individuels attribués étude par étude dans l'étape précédente en un score global tenant compte des forces et faiblesses de l'ensemble du *corpus* de données. À ce stade, les données expérimentales issues d'études animales sont traitées séparément des données épidémiologiques issues d'études humaines.

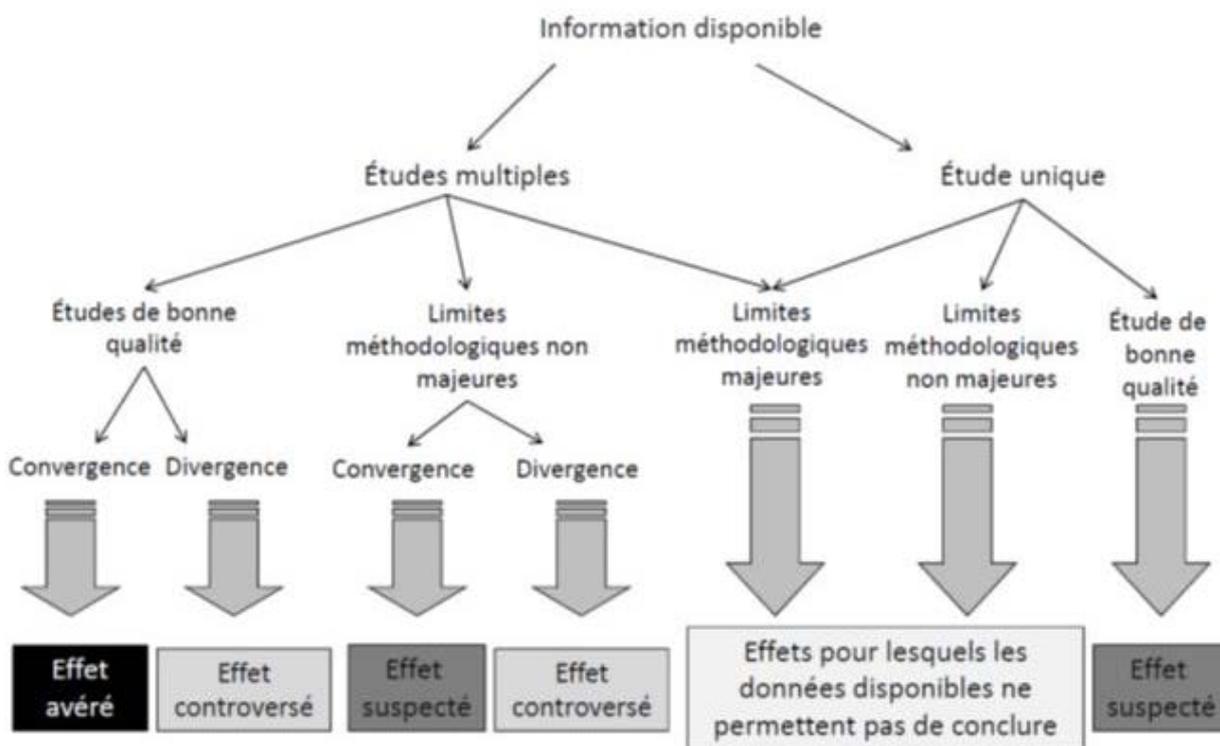


Figure 3 : Arbre décisionnel pour l'attribution d'un niveau de confiance aux corpus de données des différents niveaux de preuves

Il est préconisé de s'appuyer sur un arbre décisionnel tel que celui utilisé par l'Anses, lors de l'ERS du BPA (Figure 3), en tenant compte du nombre d'études montrant des résultats cohérents. Les études ne montrant pas d'effet, à qualité égale, devront venir pondérer le niveau de confiance global (ANSES 2013a).

Étape 7 – Intégrer les niveaux de preuves chez l'animal et chez l'être humain

Pour cette étape, il est préconisé d'utiliser une table d'analyse qui permet de mettre simplement en regard les niveaux de confiance issus des études humaines et ceux issus d'études animales. En croisant ces trois niveaux (faible, moyen, fort), une conclusion peut alors être tirée sur la nature de

l'effet : de « reconnu » par exemple si le niveau de preuve est fort chez l'être humain (et ce quel que soit le niveau de preuve chez l'animal) à « présumé » (modéré chez l'être humain et fort ou modéré chez l'animal ou faible chez l'être humain mais fort chez l'animal), « suspecté » (modéré chez l'être humain et faible chez l'animal ou faible chez l'être humain et modéré chez l'animal) à « non qualifiable » (faible chez l'être humain et l'animal). Ces niveaux de preuve pourraient être complétés par l'expression du degré d'incertitude associé telle décrit dans le chapitre précédent (cf. paragraphe 5.2.1).

6.2 Prise en compte quantitative de l'incertitude dans la caractérisation de relations dose-réponse

La démarche proposée par l'IPCS, présentée au paragraphe 5.2.2, est intéressante pour guider une démarche d'évaluation quantitative de l'incertitude. Les points forts sont notamment d'identifier et de caractériser par des distributions statistiques les différentes sources d'incertitude pour proposer un calcul de l'incertitude globale. En revanche, les propositions de choix par défaut sont contestables et pas directement transposables à la problématique des PE, par exemple. Ils méritent de plus amples réflexions.

Tout d'abord, il s'agit de déterminer des niveaux « acceptables » d'occurrence et de sévérité pour les critères d'effets de la perturbation endocrinienne. Ce choix doit être guidé par une relation entre la sévérité et des conséquences sanitaires maîtrisées ou une connaissance de la variabilité attendue au sein d'une population contrôle. Ce choix pourra faire l'objet d'un futur travail.

Ensuite, il est conseillé rechercher systématiquement comme point de départ une BMD plutôt qu'une NOAEL, compte tenu des avantages explicités dans ce document, en cherchant à l'obtenir à partir des données de l'étude de référence. En cas d'impossibilité de calculer la BMD (données sources indisponibles, par exemple), une distribution de ratios historiques entre NOAEL et BMD pour le même critère d'effet et des études de qualité comparable (notamment en termes de puissance statistique) pourrait être utilisée, de la même manière que proposé dans le document de l'IPCS. Ceci demande un travail bibliographique et statistique spécifique. En l'absence d'un tel travail, la démarche proposée par l'IPCS pourrait conduire à une évaluation quantitative de l'incertitude biaisée à la base et sans réelle pertinence.

S'agissant des extrapolations, plusieurs pistes d'améliorations sont à envisager à plus ou moins long terme, en plus de la collecte de données disponibles pour les critères d'effets sur les substances chimiques étudiées, afin de définir des distributions adéquates pour chaque extrapolation. À court terme, il est ainsi possible d'améliorer la pertinence de l'extrapolation en recentrant l'utilisation des données disponibles sur des composés proches de la substance investiguée (propriétés structurales et physicochimiques proches, mode d'action similaire), par des méthodes de type « read-across ». Il peut aussi être conseillé de combiner plusieurs sources d'informations venant renforcer une extrapolation. Par exemple, des données *in vitro* de métabolisme sur des cellules hépatiques primaires de rongeurs et d'humains peuvent conforter ou rejeter une extrapolation du rongeur à l'être humain sur la base de modèles de toxicocinétique apparentés. De même, une bonne connaissance des voies métaboliques peut permettre d'identifier des paramètres clés expliquant les différences de sensibilité entre les individus ayant un métabolisme lents et rapides en fonction de leur bagage enzymatique. Ceci devrait alors conduire à mieux prédire l'ampleur de la variabilité de la réponse au sein de la population humaine. Là encore, un travail bibliographique sur la variabilité inter-individus sur les paramètres explicatifs de la toxicocinétique, et un travail d'intégration de cette information dans des modèles toxicocinétiques, notamment ceux fondés sur la physiologie (« physiologically-based toxicokinetics models » ou PBTK) est à envisager. Les aspects toxicodynamiques, dans le cas où les effets peuvent être liés à des récepteurs ou à des processus de détoxification, qui peuvent varier d'un individu à l'autre, pourraient être abordés de manière similaire. Tous ces aspects qui prennent en compte les mécanismes en jeu, au niveau de la cinétique et du mode d'action, doivent faire l'objet d'un travail spécifique, pour lequel le document de l'IPCS, centré sur des approches probabilistes, reste peu informatif (IPCS 2014).

6.3 Démarche d'analyse qualitative de l'incertitude dans l'évaluation de l'exposition et son application au toluène

La présente démarche part de celle adoptée, par l'Anses, pour l'évaluation de l'exposition au BPA (cf. paragraphe 5.1.1) (ANSES 2013a). Elle a été consolidée et complétée à travers un cas d'étude, celui du calcul de l'exposition au toluène, en se voulant générique, *i.e.* applicable à toutes les substances chimiques.

La démarche proposée, schématisée par la Figure 4, comprend quatre étapes ; les deux premières, à savoir le recensement et la caractérisation systématique des sources d'incertitude, reprend l'approche mise en œuvre lors de l'évaluation du BPA. Deux autres étapes les complètent ; la première vise à évaluer l'influence des principales sources d'incertitude sur l'évaluation de l'exposition et la deuxième, à déterminer le niveau de confiance attribué aux résultats de l'évaluation de l'exposition. Ces quatre étapes sont décrites en détails dans les paragraphes qui suivent.

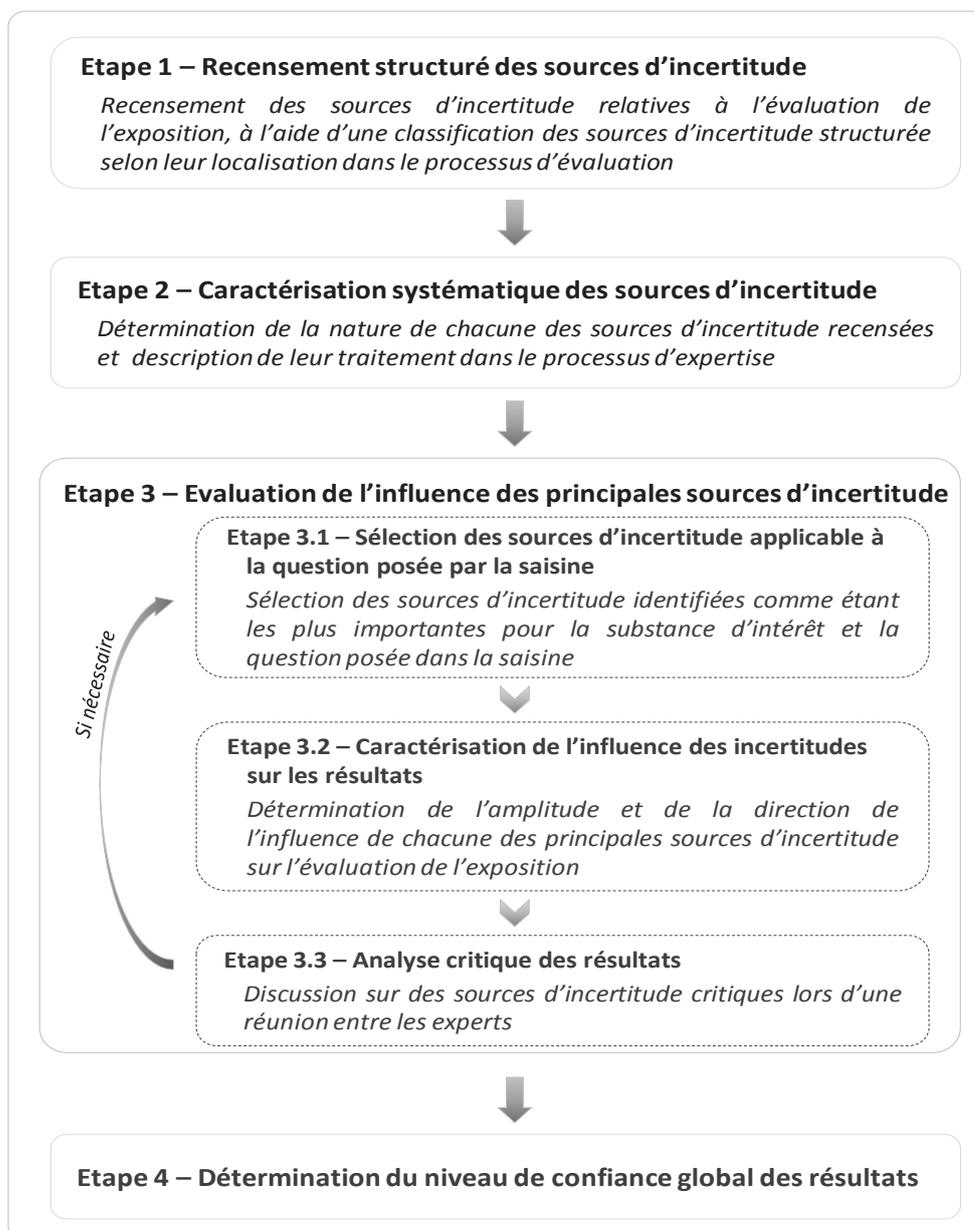


Figure 4 : Démarche proposée par le CT "Incertain" pour l'analyse qualitative des sources d'incertitude en évaluation de l'exposition aux substances chimiques, à l'Anses

6.3.1 Étape 1 – Recensement des sources d'incertitude & classification utilisée

Pour le recensement des sources d'incertitude, la classification préconisée est adaptée de celle utilisée par l'Anses pour l'ERS au BPA, et son applicabilité réévaluée à travers le cas du toluène. Une présentation complète et détaillée de cette classification figure en Annexe 1. La suite de ce paragraphe en présente l'essentiel.

La classification des sources d'incertitude proposée présente quatre classes de sources d'incertitude (contextuelle, scénarios, modèles et données d'entrée) (cf. Tableau 10).

Tableau 10 : Classification des sources d'incertitude relatives à l'évaluation de l'exposition à une substance chimique, en vue d'une ERS, structurée selon leur localisation dans le processus d'expertise

Sources d'incertitude contextuelles
- <i>Spécifications réglementaires</i>
- <i>Questions posées par l'expertise</i>
Sources d'incertitude des scénarios
- <i>Population cible considérée</i>
- <i>Source d'exposition considérée</i>
- <i>Voie d'exposition considérée</i>
- <i>Echelle de temps d'exposition</i>
Sources d'incertitude des modèles
- <i>Choix des variables du modèle</i>
- <i>Choix de l'équation mathématique</i>
- <i>Corrélation entre les variables</i>
Sources d'incertitude des données
- <i>Critères de choix des données utilisées</i>
- <i>Représentativité des données</i>
- <i>Prise en compte de la variabilité</i>
- <i>Précision des mesures des variables</i>

6.3.2 Étape 2 – Caractérisation systématique des sources d'incertitude

Pour la caractérisation systématique des sources d'incertitude, l'approche préconisée est celle adoptée par l'Anses, pour l'ERS du BPA, *i.e.* de préciser la nature de chaque source d'incertitude recensée et de décrire son traitement dans le processus d'évaluation. Le tout alimente une matrice des incertitudes. L'intérêt informatif de ces deux dimensions a, en effet, été conforté à travers le cas d'étude « Toluène ».

S'agissant de la nature des sources d'incertitude, l'information doit être recueillie de façon standardisée à l'aide d'une pré-liste de modalités possibles pour en faciliter l'analyse. Dans le cas du Toluène, la pré-liste proposée inclut les modalités suivantes : *Manque de données ; Hypothèse ; Extrapolation ; Facteur par défaut ; Manque de connaissances scientifiques ; Limites métrologiques ; Limites imposées par la saisine ; Limites imposées par le champ de compétences de l'Anses ; Autres*. Cette pré-liste, plus exhaustive que celle utilisée pour le BPA, pourra être complétée si nécessaire, ou même être le point de départ d'une typologie sur la nature des sources d'incertitude spécifique à l'ERS de substances chimiques.

S'agissant du traitement des sources d'incertitude dans le processus d'évaluation, la description s'effectue de façon narrative. En effet, les différents traitements possibles des sources d'incertitude dans le cadre d'une ERS et leur niveau de complexité rendent la standardisation très délicate ; cet effort mérite toutefois d'être entrepris pour gagner en communication et en transparence.

6.3.3 Étape 3 – Évaluation de l'influence des principales sources d'incertitude sur l'évaluation de l'exposition

L'objectif de cette étape est d'évaluer l'impact de l'incertitude rémanente après son traitement lors du calcul de l'exposition, sur le résultat de ce calcul.

La procédure préconisée fait exclusivement appel au jugement d'expert, elle se décline en trois sous-étapes (cf. Figure 4) ; la première consiste à sélectionner des sources principales d'incertitude, parmi celles rapportées dans la matrice des incertitudes ; la deuxième consiste à en caractériser leur influence en terme de direction⁹ et d'amplitude¹⁰ sur l'évaluation de l'exposition ; et la troisième discute des résultats obtenus. Cette discussion peut conclure à la nécessité de revenir à la sous-étape 1, *i.e.* revenir sur la sélection des sources principales d'incertitude. Les sous-étapes 2 et 3 sont alors reprises, et ainsi de suite jusqu'à obtenir une liste stabilisée de sources principales d'incertitude et une caractérisation de leur influence sur l'évaluation de l'exposition.

Pour chaque source d'incertitude sur le résultat final, l'évaluation de l'amplitude de son influence s'appuie sur une échelle de Likert (Likert 1932). Largement répandue dans les questionnaires psychométriques, l'échelle de Likert permet aux personnes interrogées d'exprimer leur degré d'accord ou de désaccord vis-à-vis d'une affirmation (Loewenthal 1996) (Loewenthal 1996). Ainsi, pour chaque source d'incertitude considérée, il est demandé à chaque participant de répondre à la question suivante : « *Quelle est l'influence de cette source d'incertitude sur les résultats ?* ». L'échelle utilisée comprend sept modalités de réponse (cf. Tableau 11) dont trois permettent de nuancer une réponse en faveur d'une influence plutôt forte, trois autres, en faveur d'une influence plutôt faible et une modalité neutre *i.e.* non polarisée sur une influence à tendance forte ou faible ; cette échelle n'est donc pas à choix forcé.

Tableau 11 : Échelle de jugement utilisée pour déterminer l'amplitude de l'impact d'une source d'incertitude sur l'évaluation de l'exposition au toluène

Réponse	Score
Très forte	1
Forte	2
Assez forte	3
Assez faible	4
Faible	5
Très faible	6
Je ne peux pas répondre	/

Un score est attribué *a priori* à chaque modalité de réponse. Il n'est pas obligatoire d'attribuer un score élevé aux réponses en faveur d'une influence à tendance forte plutôt que l'inverse. Par exemple, cette dernière option a été choisie dans l'étude de cas « Toluène », les scores les plus élevés ayant été attribués aux réponses en faveur d'une influence à tendance faible. Cependant, une fois choisie, la logique doit prévaloir sur toutes les sources d'incertitude considérées. L'avantage d'une telle échelle est qu'elle permet de s'assurer que l'ensemble des participants comprennent de manière homogène les modalités (ou niveaux) de réponse.

Par ailleurs, il est demandé à chaque participant évaluer la direction de l'influence de chaque source d'incertitude sur les résultats, dans le sens d'une sous-estimation ou d'une surestimation. Comme pour la première question, l'option « je ne peux pas répondre » est disponible.

La matrice doit être complétée par chaque participant, individuellement et de manière confidentielle, *i.e.* indépendamment les uns des autres. Une fois renseignées, les matrices sont toutes transmises à une tierce personne en charge de synthétiser les résultats. Le caractère confidentiel d'un tel procédé présente l'intérêt de permettre à chaque expert d'exprimer individuellement son propre jugement, sans qu'il soit infléchi, dans un premier temps, par les autres membres du groupe. En effet, il arrive que la dynamique d'un groupe favorise une idée ou une autre en fonction de l'influence de l'expert qui la porte, de sa capacité à la défendre devant ses collègues, y compris lorsqu'il est confronté à des avis critiques.

Pour le cas d'étude « Toluène », huit participants ont renseigné la matrice d'incertitude (cinq agents de l'Anses et trois experts). La durée de renseignement du questionnaire varie entre 40 minutes et trois heures (moins d'une heure pour 6 experts et 2 à 3 heures pour deux autres). Les réponses individuelles ont ensuite été anonymisées et synthétisées graphiquement par la Figure 5, avant d'être discutées collégalement par l'ensemble du groupe. À l'issue de cette discussion, chaque participant a eu la possibilité de réviser son score initial.

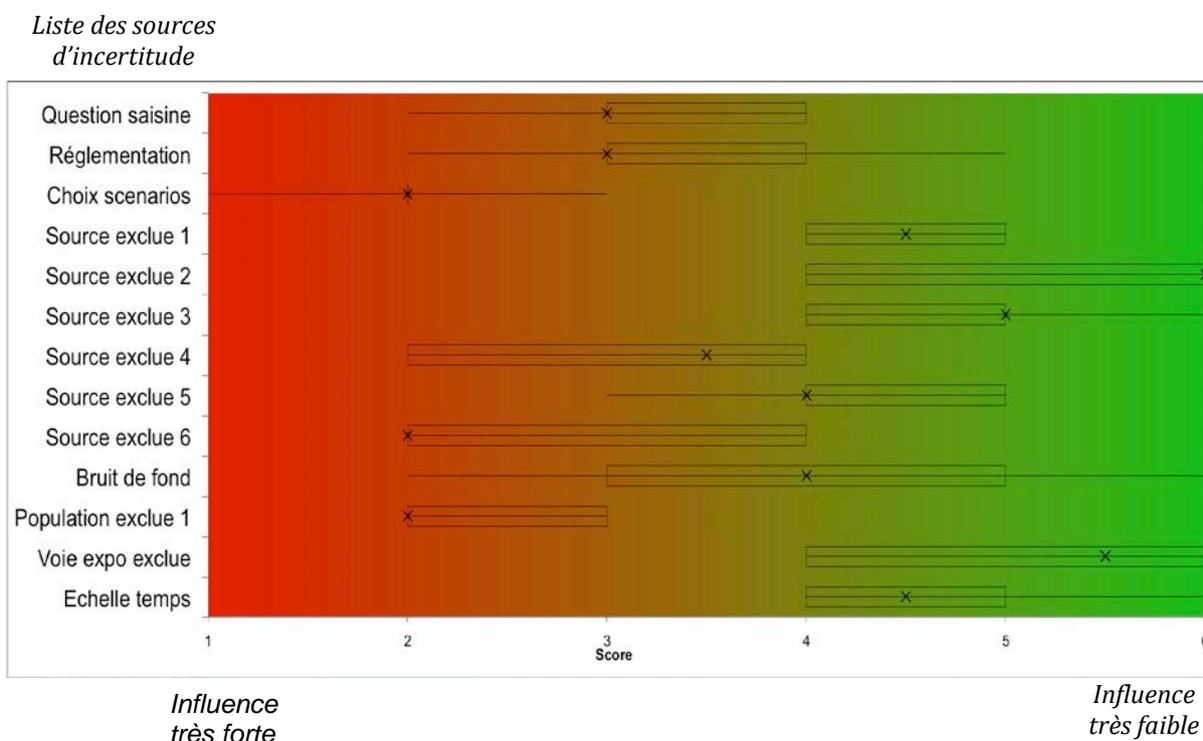


Figure 5: Incertitudes contextuelles et dans les scénarios

L'objectif premier de la représentation graphique n'est pas de fournir une vision finale de l'incertitude dès la première évaluation faite par chaque expert individuellement, mais de servir de support pour une discussion entre les experts, qui s'assureront ainsi qu'aucun aspect important n'a été oublié, que le consensus est solide et que les avis minoritaires ou les points de dissensus ne sont pas oubliés. Les divergences qui subsistent après cette discussion pourront être représentées, après une réévaluation par chacun des experts, dans une version graphique finale rapportant les sources d'incertitude discutées.

Dans la Figure 5, la médiane des scores donnés par les répondants est représentée par un « x », l'intervalle interquartile, représenté par un rectangle, et pour chaque source d'incertitude, une droite relie le score minimum et le score maximum. Les zones rouges correspondent à des scores indiquant une influence forte de la source d'incertitude, alors que le vert correspond à une influence faible. Cette représentation graphique permet de repérer facilement les sources

d'incertitude qui ont une influence forte ou très forte sur les résultats finaux du calcul de l'exposition. La représentation de l'intervalle interquartile aide à visualiser, pour chaque source d'incertitude, l'hétérogénéité des réponses des experts, afin de se focaliser sur celles où l'hétérogénéité entre les participants est importante. Les résultats de cette partie de l'étude de cas « Toluène » figurent en Annexe 2.

Pour conclure, on peut observer que cette méthode peut jouer un rôle d'analyse de sensibilité qualitative, car elle permet de prioriser les sources d'incertitude les plus importantes. Par ailleurs, un niveau de confiance global dans le calcul de l'exposition peut être calculé à partir de cette évaluation.

6.3.4 Étape 4 – Détermination d'un niveau de confiance global

Le niveau de confiance globale des résultats est fonction du niveau d'influence évalué au cours de l'étape 3 et l'hétérogénéité des réponses des participants à cette évaluation. Pour standardiser les réponses, il est préconisé de s'aider d'une échelle telle que celle utilisée dans le cas d'étude « Toluène ». Elle se décline en trois modalités – niveau de confiance « haut », « moyen » et « réduit » – définies comme suit :

Les résultats d'une évaluation de l'exposition seront associés à :

- un **haut degré de confiance**, si moins d'un tiers des sources principales d'incertitude sont critiques,
- un **niveau moyen de confiance**, si un tiers à deux tiers des sources principales d'incertitude sont critiques
- un **niveau réduit de confiance**, si plus de deux tiers des sources principales d'incertitude sont critiques

Une source d'incertitude est dite « critique » si elle se trouve dans la situation suivante :

- la médiane des scores est inférieure à 4, ce qui exprime une influence évaluée à relativement « forte ».
- l'évaluation de son influence induit une très grande hétérogénéité des réponses parmi les experts interrogés (4 ou 5 points de différence entre deux experts, la plus grande différence possible entre deux experts étant de 5 points), ce qui exprime un certain manque de consensus dans le groupe, peut-être une opinion minoritaire ;
- son influence est considérée comme très forte au moins par un expert (score de 1), exprimant potentiellement une opinion minoritaire subsistant même après la discussion des différentes sources d'incertitude dans le groupe,

A titre d'illustration, pour le toluène, 57 sources principales d'incertitude ont été recensées. Parmi elles, 27 ont été classées comme étant « critiques ». On peut donc conclure, selon la règle de décision établie, à un « niveau moyen de confiance » dans le calcul de l'exposition.

6.3.5 Discussion sur la démarche proposée

Le domaine disciplinaire de chacun des experts, ainsi que leur niveau de compétence sur le sujet spécifique traité, sont des aspects importants qui peuvent influencer les résultats de l'évaluation qualitative de l'incertitude mais qui n'ont pas été inclus dans la démarche dans ce rapport, notamment pour des raisons de faisabilité. En effet, il s'agissait pour le cas d'étude « Toluène » de vérifier la démarche méthodologique proposée, en évitant de demander un nombre important d'informations supplémentaires aux participants. Mais l'origine disciplinaire et la compétence sur un sujet particulier peuvent être facilement incluses dans la démarche, en demandant à chacun des participants de cocher une ou plusieurs cases qui correspondraient à leur discipline. Pour illustrer ce type d'approche, l'annexe 3 reprend le questionnaire d'identification des participants qui

a été utilisé dans une démarche similaire à celle proposée ici mais appliquée à l'évaluation de la qualité des études *in vivo*, dans un projet financé dans le cadre du PNRPE⁴¹ (intitulé « Dico-Risque »).

Ainsi, il est possible de comparer les réponses fournies par des experts ayant coché les mêmes cases disciplinaires, en réalisant deux représentations graphiques des réponses au lieu d'une seule (comme dans la Figure 5). Les deux représentations graphiques illustreraient les réponses des experts du domaine disciplinaire d'intérêt, et ceux des autres domaines disciplinaires du groupe. Une autre possibilité est de réaliser deux représentations graphiques pour deux disciplines d'intérêt, ou pour deux niveaux de compétence sur le sujet différents (ex. : sous-groupe expologie et sous-groupe dangers). La comparaison des deux figures permettrait d'identifier des critères où les différences sont importantes entre les deux groupes disciplinaires ou de compétences. Les niveaux de confiance pourraient également être différents pour chacun des deux sous-groupes disciplinaires/de compétence. Un tel exercice pourrait être utilisé dans une démarche itérative afin d'identifier et d'explicitier les sources d'incertitude objets discutées entre différents experts (cf. Figure 4).

6.4 Proposition pour la prise en compte quantitative de l'incertitude dans le calcul de l'exposition

L'évaluation de l'exposition est généralement réalisée en s'appuyant sur un modèle mathématique qui relie des paramètres de transferts et des données d'exposition dans plusieurs compartiments. L'analyse quantitative de l'incertitude (*i.e.* interaction de l'incertitude et du poids des facteurs d'entrée du modèle sur la précision des prédictions) consiste d'abord à effectuer une analyse de sensibilité de la sortie d'intérêt du modèle (par exemple, la concentration moyenne de la substance étudiée dans l'organisme) par rapport aux facteurs d'entrée du modèle. Ensuite, une représentation schématique telle que celle proposée par le Tableau 12, qui classe les facteurs d'entrée suivant le degré d'incertitude sur la valeur (à distinguer de la variabilité) et les résultats de l'analyse de sensibilité, permet d'identifier les facteurs d'entrée qui ont à la fois le plus d'impact et l'incertitude la plus élevée. Ceci permet ensuite d'orienter les études à réaliser pour réduire cette incertitude.

Tableau 12 : Représentation schématique de l'analyse quantitative de l'incertitude. Dans l'exemple, le paramètre 3, avec une incertitude élevée et un impact fort sur les prédictions devrait faire l'objet d'études complémentaires.

		INCERTITUDE		
		Forte	Moyenne	Faible
SENSIBILITE	Forte	paramètre 3	paramètre 1	
	Moyenne			paramètre 2
	Faible			paramètre 4 paramètre 5

Nous recommandons ainsi la systématisation de l'analyse de sensibilité pour l'analyse quantitative de l'incertitude sur l'évaluation de l'exposition. Parmi les différentes méthodes pouvant être appliquées, nous recommandons le calcul des indices de Sobol qui est adapté à tout type de modèle, de complexité variable, et dont l'implémentation, qui nécessite un nombre important de calculs, ne pose pas de problème compte tenu des capacités informatiques d'aujourd'hui.

⁴¹ Programme national de recherche sur les perturbateurs endocriniens (PNRPE)

7 Recommandations

Ce chapitre propose les prochaines étapes du processus au regard des questions méthodologiques qui se posent, des données d'entrée nécessaires, et du déploiement des démarches méthodologiques proposées dans ce document. Il se décline en trois grandes parties : des recommandations sur le déploiement des méthodes proposées, à mettre en œuvre dès que possible, et à moyens termes, des recommandations sur de nouveaux développements méthodologiques et des recommandations sur l'acquisition de données.

Recommandations sur le déploiement d'un traitement standardisé de l'incertitude en ERS au sein de l'Anses :

- Il est recommandé de conduire systématiquement une analyse d'incertitude et la rapporter, avec le résultat de l'ERS, dans les rapports d'expertise collective, pour que l'ensemble des parties prenantes ait une pleine visibilité des éléments clés qui influencent l'estimation du risque ;
- dans la perspective d'une mise en routine ultérieure, il est nécessaire d'éprouver, à travers des cas d'étude, la démarche de traitement de l'incertitude proposée dans ce rapport, dans le cadre des ERS des substances chimiques. En particulier, son applicabilité dans le cadre d'une expertise collective, les ressources nécessaires en temps et en personnel, devront être évaluées ;
- dans tous les cas, il conviendrait de proportionner la conduite de l'analyse d'incertitude au regard de la méthode d'ERS retenue (exemple : si une évaluation déterministe fondée sur des hypothèses conservatrices indique une absence de risque, une analyse approfondie (qualitative ou quantitative) de l'incertitude a peu d'utilité) ;
- la conduite de l'analyse d'incertitude doit être de niveau de détail comparable entre les différents volets de l'ERS (danger, exposition et risque), et suivre les spécificités méthodologiques qui leur sont propres ;

Recommandations méthodologiques sur le traitement de l'incertitude en ERS :

- en vue d'une harmonisation des pratiques au sein de l'agence, il est recommandé d'adopter un référentiel méthodologique unique couvrant les quatre étapes de l'ERS et les outils qualitatifs et quantitatifs pertinents ;
- pour le cas où il est fait appel à un jugement d'expert, celui-ci devra être recueilli en utilisant une méthode formalisée d'élicitation de ce jugement telle que celle proposée par l'EFSA (EFSA 2014) ;
- il est nécessaire de choisir et, le cas échéant, développer des outils d'évaluation de la qualité et de la pertinence (grilles de lecture) des données et des modèles sélectionnés pour l'ERS (danger et exposition) ;
- concernant l'étape de caractérisation des sources d'incertitude, il est proposé d'utiliser, en première intention, les deux dimensions recommandées dans ce rapport – « nature » de la source d'incertitude et son « traitement » dans le cadre de l'expertise ; ces critères pourront être affinés par la suite et leur liste complétée selon les besoins (cf. Chapitre 6) ;
- Les méthodes bayésiennes peuvent être utiles en raison de leur capacité à intégrer l'incertitude dans des hypothèses de départ. Par exemple, elles permettraient de combiner les incertitudes à partir de données spécifiques à certaines substances chimiques avec des données génériques dérivées de données sur d'autres substances chimiques présentant les mêmes caractéristiques. Il est donc recommandé d'évaluer, à travers des cas d'étude, leur applicabilité dans un contexte d'expertise collective à l'Anses ;

- les sources d'incertitude recensées n'ayant pas toutes le même impact sur les résultats de l'ERS, il apparaît nécessaire de les hiérarchiser et éventuellement de les pondérer pour une meilleure appréciation de l'impact global ;
- l'évaluation de l'impact agrégé de toutes les sources d'incertitude identifiées, visant à définir un niveau de confiance global des résultats, doit faire l'objet d'un travail méthodologique spécifique. Le point de départ pour la réflexion peut être le calcul du niveau de confiance pour l'exposition (cf. paragraphe 6.3.4) ou l'approche utilisée dans le cadre l'identification du danger fondée sur l'appréciation des éléments de preuve (cf. paragraphe 6.1) ;
- il est recommandé une démarche itérative fondée sur l'appréciation du niveau de confiance global des résultats pour proportionner les recommandations destinées aux gestionnaires du risque.

Recommandation spécifique relative aux données d'entrée :

- pour les données utilisées de façon récurrente lors des ERS menées par l'Anses (e.g. base de données CONTAMINE, données des études EAT⁴² et INCA⁴³, future base de données sur les facteurs d'exposition), il est souhaitable d'identifier l'ensemble des sources possibles d'incertitude, pour toutes les futures analyses d'incertitude.

⁴² Etude de l'alimentation totale (EAT)

⁴³ Individuelle Nationale de Consommation Alimentaire (INCA)

8 Conclusions

Ce rapport est le résultat d'un premier travail de réflexion sur le traitement de l'incertitude en ERS, au sein de l'Anses. Le travail a été initié dans le cadre du GT PE. Il amorce un recensement et une analyse critique des principaux outils et méthodes faisant actuellement référence, et propose une démarche standardisée pour le traitement de l'incertitude en ERS, adaptée au contexte d'expertise collective en place à l'Anses.

L'incertitude « fait référence à une situation où le corpus de connaissances disponibles sur un sujet particulier (exemple : les effets toxiques des PE) est perçu comme « problématique » (à savoir, peu connu ou inconnu, incomplet, incorrect, biaisé, pas assez convaincant, ...) par divers individus. Il s'agit donc d'une définition qui va au-delà d'un « manque de connaissances », pour inclure le potentiel d'erreur ou de biais, les insuffisances techniques des instruments de recherche, les insuffisances méthodologiques des protocoles de recherche, les choix par jugement d'expert des hypothèses qui remplacent les connaissances scientifiques manquantes, les divergences d'interprétation des données brutes disponibles et le caractère incomplet, incorrect ou ambigu de la communication des résultats. » (ANSES 2013b).

Le présent document s'intéresse à l'ensemble du processus d'ERS et considère toutes les sources d'incertitude possibles – quantifiable et non quantifiable – lors d'une ERS des substances chimiques. Il montre qu'une analyse d'incertitude permet de (ré-)évaluer l'ensemble des éléments intervenant dans une évaluation : le contexte de l'expertise, les scénarios choisis, la modélisation (structure des modèles et données d'entrées), etc. et considère leur influence sur les résultats de l'évaluation.

Dans la mesure où une méthode peut difficilement répondre à elle seule à toutes les questions qui peuvent se poser lors d'une analyse d'incertitude, et qu'en outre une première analyse peut conduire à de nouvelles questions, le document met en exergue l'avantage d'une démarche standardisée, fondée sur une approche par paliers, à quatre niveaux d'analyse, avec un premier niveau dit « Screening » (Niveau 0), une analyse qualitative (Niveau 1) et deux niveaux d'analyse quantitative, utilisant des données déterministes (Niveau 2) et probabilistes (Niveau 3). Au final, il propose, pour chaque étape de l'ERS, une première démarche standardisée pour le traitement qualitatif et quantitatif de l'incertitude, compatible avec les procédures d'expertises collectives en vigueur à l'Anses.

Les limites des démarches méthodologiques présentées sont discutées tout le long du document, et des axes de progrès identifiés. Enfin, des recommandations sont formulées ci-dessus pour initier un déploiement au sein de l'Anses. Le document souligne ainsi que le travail réalisé ne vise pas à répondre à toutes les questions qui se posent en matière d'analyse d'incertitude, mais s'inscrit plutôt dans un processus dynamique d'évolution méthodologique, d'acquisition de données, de communication, etc. pour optimiser la prise de décision en matière de gestion des risques sanitaires.

9 Bibliographie

- AFSSET. 2007. Valeurs toxicologiques de référence pour les substances reprotoxiques. Méthode de construction de VTR fondées sur des effets toxiques pour la reproduction et le développement (Saisine Afsset n° 2003/AS03). Maisons-Alfort: Afsset.
- AFSSET. 2010. Valeurs toxicologiques de référence pour les substances cancérigènes. Méthode de construction de VTR fondées sur des effets cancérigènes (Auto-Saisine 2004/AS16) Maisons-Alfort: Afsset.
- AHRQ. 2012a. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. AHRQ.
- AHRQ. 2012b. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Rockville, MD: AHRQ.
- Allen, B. C., R. J. Kavlock, C. A. Kimmel, and E. M. Faustman. 1994. "Dose-response assessment for developmental toxicity: II. Comparison of generic benchmark dose estimates with no observed adverse effect levels." *Toxicological Sciences* 23 (4):487-495. doi: 10.1093/toxsci/23.4.487.
- ANSES. 2013a. Avis et rapport de l'Anses relatif à l'Évaluation des risques du bisphénol A (BPA) pour la santé humaine. - Tome 1 : Évaluation des risques du bisphénol A (BPA) pour la santé humaine et aux données toxicologiques et d'usage des bisphénols S, F, M, B, AP, AF, et BADGE. Maisons-Alfort: Anses.
- ANSES. 2013b. Rapport d'appui scientifique et technique de l'Anses relatif à "Perturbateurs endocriniens - Synthèse et étude des auditions, contribution à la saisine n°2009-SA-0331 : Expertise sur les risques sanitaires pour le consommateur liés à des substances reprotoxiques et/ou perturbateurs endocriniens présents dans des produits et/ou articles mis sur le marché". Maisons-Alfort: Anses.
- Balshem, H., M. Helfand, H. J. Schünemann, A. D. Oxman, R. Kunz, J. Brozek, G. E. Vist, Y. Falck-Ytter, J. Meerpohl, S. Norris, and G. H. Guyatt. 2011. "GRADE guidelines: 3. Rating the quality of evidence." *Journal of Clinical Epidemiology* 64 (4):401-406. doi: 10.1016/j.jclinepi.2010.07.015.
- Bernier, J., E. Parent, and J.J. Boreux. 2000. *Statistique pour l'environnement: traitement bayésien de l'incertitude*: Tec & Doc.
- Bogen, K. T., A. C. Cullen, H. C. Frey, and P. S. Price. 2009. "Probabilistic exposure analysis for chemical risk characterization." *Toxicological Sciences* 109 (1):4-17. doi: 10.1093/toxsci/kfp036.
- Bogen, K. T., and R. C. Spear. 1987. "Integrating uncertainty and interindividual variability in environmental risk assessment." *Risk Analysis* 7 (4):427-436.
- Bokkers, B. G. H., and W. Slob. 2007. "Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach." *Critical Reviews in Toxicology* 37 (5):355-373. doi: 10.1080/10408440701249224.
- Bruchou, C., J. Couteau, N. Dumoulin, R. Faivre, T. Faure, B. Iooss, S. Lehuta, S. Mahévas, D. Makowski, H. Monod, B. Poussin, E. Ramat, H. Richard, L. Rouan, J-C. Soulié, and J. Wang. 2013. *Analyse de sensibilité et exploration des modèles – Application aux sciences de la nature et de l'environnement* QJAE.

- Chiu, W. A., and P. White. 2006. "Steady-state solutions to PBPK models and their applications to risk assessment I: Route-to-route extrapolation of volatile chemicals." *Risk Analysis* 26 (3):769-780. doi: 10.1111/j.1539-6924.2006.00762.x.
- Ciffroy, P., M. Keller, and A. Pasanisi. 2013. "Estimating hazardous concentrations by an informative Bayesian approach." *Environmental Toxicology and Chemistry* 32 (3):602-611. doi: 10.1002/etc.2096.
- Ciffroy, P., T. Tanaka, E. Johansson, and C. Brochot. 2011. "Linking fate model in freshwater and PBPK model to assess human internal dosimetry of B(a)P associated with drinking water." *Environmental Geochemistry and Health* 33 (4):371-387. doi: 10.1007/s10653-011-9382-6.
- Craye, M., S. O. Funtowicz, and J. P. van der Sluijs. 2005. "A reflexive approach to dealing with uncertainties in environmental health risk science and policy." *International Journal of Risk Assessment and Management* 5:216-236. doi: 10.1504/ijram.2005.007169.
- 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.
- Delmaar, J.E. , M.V.D.Z. Park, and J.G.M. van Engelen. 2005. ConsExpo 4.0 - Consumer Exposure and Uptake Models - Program Manual. RIVM.
- Droesbeke, J.J. , J. Fine, and G. Saporta. 1997. *Plans d'Expériences: Applications à l'Entreprise*. Paris: Technip.
- ECHA. 2012. "Uncertainty analysis (ECHA-12-G-25-EN)." In *Guidance on information requirements and chemical safety assessment*. Helsinki: ECHA.
- 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. 2008. Scientific Opinion of the Panel on Plant Health on a request from the European Commission on *Guignardia citricarpa* Kiely. In *EFSA Journal*: European Food Safety Authority (EFSA).
- EFSA. 2009. "Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessment (EFSA-Q-2005-232)." *EFSA Journal* 7 (6):1150, 72 pp. doi: 10.2903/j.efsa.2009.1150.
- 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).
- Fang, K-T. , R. Li, and A. Sudjianto. 2006. *Design and modeling for computer experiments*. London Chapman et Hall/CRC.
- Frey, H. C. 1992. *Quantitative analysis of uncertainty and variability in environmental policy making*. Washington, DC: American Association for the Advancement of Science.
- Gaylor, D., L. Ryan, D. Krewski, and Y. Zhu. 1998. "Procedures for calculating benchmark doses for health risk assessment." *Regulatory Toxicology and Pharmacology* 28 (2):150-164. doi: 10.1006/rtp.1998.1247.
- Gee, D. 2006. "Late lessons from early warnings: Toward realism and precaution with endocrine-disrupting substances." *Environmental Health Perspectives* 114 (SUPPL.1):152-160.

- Gee, D. 2009. "Late Lessons from Early Warnings: Towards realism and precaution with EMF?" *Pathophysiology* 16 (2-3):217-231.
- Greenberg, M., C. Haas, A. Cox, K. Lowrie, K. McComas, and W. North. 2012. "Ten Most Important Accomplishments in Risk Analysis, 1980-2010." *Risk Analysis* 32 (5):771-781. doi: 10.1111/j.1539-6924.2012.01817.x.
- Guimaraes Pereira, Ângela, Sofia Guedes Vaz, and Sylvia Tognetti. 2006. *Interfaces between Science and Society*: Geenleaf publishing.
- Guyatt, G. . 2012. Tools to Assess Risk of Bias in Cohort Studies. McMaster University.
- Guyatt, G., A. D. Oxman, E. A. Akl, R. Kunz, G. Vist, J. Brozek, S. Norris, Y. Falck-Ytter, P. Glasziou, H. Debeer, R. Jaeschke, D. Rind, J. Meerpohl, P. Dahm, and H. J. Schünemann. 2011. "GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables." *Journal of Clinical Epidemiology* 64 (4):383-394. doi: 10.1016/j.jclinepi.2010.04.026.
- Hart, A. , J.P. Gosling, A. Boobis, D. Coggon, P. Craig, and D. Jones. 2010. Development of a Framework for Evaluation and Expression of Uncertainties in Hazard and Risk Assessment. Fera.
- Hattis, D. , and M. K. Lynch. 2007. Empirically Observed Distributions of Pharmacokinetic and Pharmacodynamic Variability in Humans—Implications for the Derivation of Single Point Component Uncertainty Factors Providing Equivalent Protection as Existing RfDs. edited by Toxicokinetics in Risk Assessment: Informa Healthcare USA, Inc.
- Hill, A. B. 1965. "The environment and disease: Association or Causation?" *Proceedings of the Royal Society of Medicine* 58:295-300.
- Hoffman, F. O., and J. S. Hammonds. 1994. "Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability." *Risk Analysis* 14 (5):707-712.
- 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.
- IPCS. 2014. Guidance Document on Evaluating and Expressing Uncertainty in Hazard Characterization In *Harmonization Project Document*. Geneva.
- Jourdan, A. 2005. "Planification d'expérience numériques." *Revue MODULAD* (33):63–73.
- Kalberlah, F., K. Schneider, and U. Schuhmacher-Wolz. 2003. "Uncertainty in toxicological risk assessment for non-carcinogenic health effects." *Regulatory Toxicology and Pharmacology* 37 (1):92-104. doi: 10.1016/s0273-2300(02)00032-6.
- Klimisch, H. J., M. Andreae, and U. Tillmann. 1997. "A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data." *Regulatory Toxicology and Pharmacology* 25 (1):1-5.
- 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).

- Likert, R. 1932. *A technique for the measurement of attitudes*. Edited by Archives of psychology. No. 40 vols. New York.
- Linkov, I., D. Loney, S. Cormier, F. K. Satterstrom, and T. Bridges. 2009. "Weight-of-evidence evaluation in environmental assessment: Review of qualitative and quantitative approaches." *Science of the Total Environment* 407 (19):5199-5205.
- Loewenthal, K.M. 1996. *An introduction to psychological tests and scales*. London: University College London.
- Makowski, D. 2013. "Uncertainty and sensitivity analysis in quantitative pest risk assessments; practical rules for risk assessors." *NeoBiota* (18):157-171.
- Makowski, D., and M. N. Mittinty. 2010. "Comparison of scoring systems for invasive pests using ROC analysis and Monte Carlo simulations." *Risk Analysis* 30 (6):906-915. doi: 10.1111/j.1539-6924.2010.01393.x.
- Mekel, O., O. Mosbach-Schulz, M. Schümann, P.K. Okken, C. Peters, J. Herrmann, Hehl O., M Bubenheim, R Fehr, and Timm J. 2007. Evaluation of standards and models for probabilistic exposure assessment.
- Mielke, H., K. Abraham, M. Götz, B. Vieth, A. Lampen, A. Luch, and U. Gundert-Remy. 2011. "Physiologically based toxicokinetic modelling as a tool to assess target organ toxicity in route-to-route extrapolation-The case of coumarin." *Toxicology Letters* 202 (2):100-110. doi: 10.1016/j.toxlet.2011.01.022.
- Molander, Linda, Marlene Ågerstrand, Anna Beronius, Annika Hanberg, and Christina Rudén. 2015. "Science in Risk Assessment and Policy (SciRAP): An Online Resource for Evaluating and Reporting In Vivo (Eco)Toxicity Studies." *Human and Ecological Risk Assessment: An International Journal* 21 (3):753-762. doi: 10.1080/10807039.2014.928104.
- Morgan, M.G., and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. New York: Cambridge University Press.
- NRC. 2009. *Science and Decision: Advancing Risk Assessment*. Washington, DC: The National Academies Press.
- NTP. 2015. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration OHAT Approach for systematic review and evidence integration for literature-based health assessments. Research Triangle Park, N.C.: NIEHS.
- Ott, W.R. . 1990. "A Physical Explanation of the Lognormality of Pollutant Concentrations." *J. Air Waste Manage. Assoc* (40):1378-1383.
- Parent, E., and J. Bernier. 2007. *Le raisonnement bayésien - Modélisation et inférence*. Paris: Springer-Verlag France.
- Phillips, K. P., and W. G. Foster. 2008. "Endocrine toxicants with emphasis on human health risks." *Journal of Toxicology and Environmental Health - Part B: Critical Reviews* 11 (3-4):149-151. doi: 10.1080/00927870701873115.
- Robert, C.P. 2006. *Le choix bayésien: Principes et pratiques, Statistique et probabilités appliquées*. Paris: Springer-Verlag.
- Rojas, R., L. Feyen, and A. Dassargues. 2008. "Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and Bayesian model averaging." *Water Resources Research* 44 (12). doi: 10.1029/2008WR006908.

- Rotmans, J., and M. B. A. van Asselt. 2001. "Uncertainty management in integrated assessment modeling: Towards a pluralistic approach." *Environmental Monitoring and Assessment* 69 (2):101-130.
- 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., M. Ratto, S. Tarantola, and F. Campolongo. 2012. "Update 1 of: Sensitivity analysis for chemical models." *Chemical Reviews* 112 (5):PR1-PR21. doi: 10.1021/cr200301u.
- SCENIHR. 2012. Memorandum on the use of the scientific literature for human health risk assessment purposes – weighing of evidence and expression of uncertainty. Brussels: European Commission.
- Schneider, Klaus, Markus Schwarz, Iris Burkholder, Annette Kopp-Schneider, Lutz Edler, Agnieszka Kinsner-Ovaskainen, Thomas Hartung, and Sebastian Hoffmann. 2009. "'ToxRTool', a new tool to assess the reliability of toxicological data." *Toxicology Letters* 189 (2):138-144. doi: <http://dx.doi.org/10.1016/j.toxlet.2009.05.013>.
- Schünemann, H. J., A. D. Oxman, G. E. Vist, J. P. T. Higgins, J. J. Deeks, P. Glasziou, and G. H. Guyatt. 2012. Chapter 12: Interpreting results and drawing conclusions. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, edited by J.P.T. Higgins and S. Green.
- Tressou, J. 2008. "Bayesian nonparametrics for heavy tailed distribution. Application to food risk assessment." *Bayesian Analysis* 3 (2):367-392. doi: 10.1214/08-ba314.
- U.S.EPA. 2011. U.S. EPA Exposure Factors Handbook 2011 Edition. Washington, DC: U.S. Environmental Protection Agency.
- van Asselt, M. B. A., and J. Rotmans. 2002. "Uncertainty in Integrated Assessment modelling. From positivism to pluralism." *Climatic Change* 54 (1-2):75-105.
- van der Sluijs, J. P., M. Craye, S. Funtowicz, P. Kloprogge, J. Ravetz, and J. Risbey. 2005. "Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: The NUSAP system." *Risk Analysis* 25 (2):481-492.
- van der Sluijs, J.P., J.S. Risbey, P. Kloprogge, J.R. Ravetz, S.O. Funtowicz, S. Corral Quintana, A. Guimarães Pereira, B. De Marchi, A.C. Petersen, P.H.M. Janssen, R. Hoppe, and S.W.F. Huijs. 2003. RIVM/MNP Guidance for Uncertainty Assessment and Communication: Detailed Guidance. In *RIVM/MNP Guidance for Uncertainty Assessment and Communication Series*. Utrecht: Utrecht University.
- Verdonck, F. A., A. Souren, M. B. van Asselt, P. A. van Sprang, and P. A. Vanrolleghem. 2007. "Improving uncertainty analysis in European Union risk assessment of chemicals." *Integrated environmental assessment and management* 3 (3):333-343.
- Vermeire, T., T. Jager, G. Janssen, P. Bos, and M. Pieters. 2001. "A probabilistic human health risk assessment for environmental exposure to dibutylphthalate." *Human and Ecological Risk Assessment (HERA)* 7 (6):1663-1679.
- Walker, W. E., P. Harremoës, J. Rotmans, J. P. van der Sluijs, M. B. A. van Asselt, P. Janssen, and M. P. Kreyer von Krauss. 2003. "Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support." *Integrated Assessment* 4 (1):5-17. doi: 10.1076/iaij.4.1.5.16466.

Wynne, B. 1992. "Uncertainty and environmental learning. Reconceiving science and policy in the preventive paradigm." *Global Environmental Change* 2 (2):111-127.

Zheng, J., and H. C. Frey. 2005. "Quantitative analysis of variability and uncertainty with known measurement error: Methodology and case study." *Risk Analysis* 25 (3):663-675. doi: 10.1111/j.1539-6924.2005.00620.x.

ANNEXES

Annexe 1 : Caractérisation de l'incertitude dans le calcul de l'exposition – Glossaire

La définition de l'incertitude utilisée par le GT dans ses travaux antérieurs « fait référence à une situation où le corpus de connaissances disponibles sur un sujet particulier (exemple : les effets toxicologiques des PE) est perçu comme « problématique » (à savoir, peu connu ou inconnu, incomplet, incorrect, biaisé, pas assez convaincant...) par divers individus. Il s'agit donc d'une définition qui va au-delà d'un « manque de connaissances », pour inclure le potentiel d'erreur ou de biais, les insuffisances techniques des instruments de recherche, les insuffisances méthodologiques des protocoles de recherche, les choix par jugement d'expert des hypothèses qui remplacent les connaissances scientifiques manquantes, les divergences d'interprétation des données brutes disponibles et le caractère incomplet, incorrect ou ambigu de la communication des résultats. » (ANSES, Synthèse et étude des auditions, 2013, p. 14).

La majeure partie des sources d'incertitude identifiées pour le BPA s'est avérée transposable au cas du toluène. Cependant des adaptations ont été nécessaires, les modifications apportées visaient essentiellement à éviter la redondance dans l'appellation et le contenu des sous-classes. Ainsi, dans cette nouvelle classification les sous-classes « choix des sources d'exposition » et « choix des populations exposées » sont réservées à la classe « Incertitude des scénarios », et n'apparaissent plus dans la classe « incertitudes contextuelles ». À la place, deux nouvelles sous-classes s'avèrent nécessaires, les sous-classes « Spécifications réglementaires » et « Questions posées par la saisine ». De même, une sous-classe « Critères de choix des scénarios » s'est avérée nécessaire dans la classe « Incertitudes des scénarios », une sous-classe « Choix méthodologiques applicables au calcul de l'exposition », dans la classe « Incertitude des modèles » ; cette dernière sous-classe pourrait aussi trouver sa place dans une classification spécifique au niveau « caractérisation du risque ». Enfin, une dernière sous-classe « Critères de choix des données utilisées » s'est avérée nécessaire dans la classe « Incertitudes des données d'entrée des modèles ». La classification modifiée est présentée ci-après.

Incertitudes contextuelles

Les incertitudes contextuelles relèvent de la définition du périmètre de l'expertise, par des choix qui ne sont pas dans le pouvoir du GT, mais d'autres instances (habituellement, le ou les ministères qui formulent une question, dans le cas des saisines, ou d'instances réglementaires).

Note : Toutes les autres classes, i.e., incertitudes liées aux scénarios, aux modèles et aux données d'entrée, concernent des incertitudes associées aux choix du GT lui-même lors du processus de calcul de l'exposition.

Question posée dans la saisine, hypothèses réglementaires

Spécifications / hypothèses réglementaires

Cette source d'incertitude concerne le respect des mesures réglementaires. Vérifier que ces mesures sont effectivement respectées n'est pas dans le pouvoir du GT, mais le fait qu'elles existent peut influencer certaines hypothèses faites dans le calcul de l'exposition. Par exemple : respect des mesures de gestion du risque indiquées dans les fiches de sécurité, respect des concentrations réglementaires d'une certains substance dans des produits, etc.

Questions posées par la saisine

La définition du périmètre de l'expertise peut influencer la pertinence des résultats finaux de l'expertise par rapport à la réalité elle-même. Même si le GT répond bien à une question, si la question est mal posée dès le départ, les résultats seront incapables de donner une idée correcte de la réalité objective et finalement de conduire à des décisions incorrectes (Ramsey, 2008). Typiquement, la question posée dans une saisine peut restreindre le champ des usages pris en compte dans le calcul de l'exposition. Même s'il n'a pas le pouvoir de modifier la question posée dans la saisine, le GT peut signaler que la question peut exclure un certain nombre d'expositions possibles, par exemple.

Incertitudes liées au scénario d'exposition

Cette classe d'incertitudes est liée à l'identification des scénarios, notamment à leur description (e.g. information erronée ou incomplète) et aux approximations, extrapolations ou facteurs par défaut qu'ils incluent. En général, les incertitudes liées aux scénarios concernent l'identification et la quantification des sources d'exposition considérées, des populations exposées, des échelles temporelles et spatiales de l'exposition (e.g. applicabilité géographique et saisonnière), des microenvironnements, des activités des populations concernées, des sources de l'agent à risque, des routes, des voies et des événements d'exposition, etc. (IPCS, 2008, p. 17). Si ces éléments ne sont pas complets ou pas pertinents pour répondre à la question posée, l'estimation de l'exposition est susceptible d'être biaisée.

Choix des scénarios retenus

Critères de choix des scénarios

Pour caractériser cette source d'incertitudes, les questions à se poser sont :

- comment a-t-on identifié les scénarios ?
- Selon quels critères les a-t-on jugés pertinents ?
- Est-ce qu'il y aurait pu avoir d'autres scénarios retenus ?

Choix des populations cibles incluses dans l'ERS

Populations exclues

Les questions à se poser pour caractériser cette source d'incertitudes sont :

- quelles sont les populations exclues de l'ERS (e.g. certaines classes d'âges, populations vulnérables, catégories professionnelles, etc.) ?
- Pour quelles raisons (limites de la saisine, manque de données, disponibilité des VTR, etc.) ?

Choix des sources d'exposition incluses dans l'ERS

Source d'exposition exclue

Les questions pertinentes pour caractériser cette source sont :

- Est-ce possible qu'on ne soit pas au courant de certaines sources d'exposition ? Est-ce que la méthode utilisée pour identifier des sources d'exposition (e.g. étude de filière) a permis une identification exhaustive ?
- Si on a exclu certaines sources d'exposition, quelles sont les raisons ?
- Comment traite-t-on le cumul des expositions selon différentes sources ?
- Comment traite-t-on le cumul des sources d'exposition considérées et du bruit de fond ?

Choix des voies d'exposition incluses dans l'ERS

Voie d'exposition exclue

Les questions pertinentes pour caractériser cette source sont :

- Est-ce possible qu'on ne soit pas au courant de certaines voies d'exposition ? Est-ce que la méthode utilisée pour identifier les voies d'exposition a permis une identification exhaustive ?
- Si on a exclu certaines voies d'exposition, quelles sont les raisons ?
- Comment traite-t-on le cumul des expositions selon différentes voies ?

Choix des échelles de temps de l'exposition (aigue, chronique) incluses dans l'ERS

Échelle de temps concernée

La question pertinente pour caractériser cette source est : est-ce qu'on a respecté la concordance entre le temps d'exposition considéré pour fixer la VTR (ou autre valeur toxicologique retenue pour caractériser le danger) et le temps d'exposition considéré dans les scénarios d'exposition ?

Note : La question peut se poser aussi par rapport à l'échelle spatiale d'un scénario. Si par exemple l'échelle spatiale est trop large alors que les expositions concernent des individus ou des populations plus concentrées à des échelles spatiales plus réduites, l'exposition sera diluée par l'inclusion des individus ou populations peu exposés. Dans ce cas, les individus ou populations les plus exposés ne seront pas identifiés dans le calcul de l'exposition, et finalement ne seront pas protégés par l'évaluation du risque. Les microenvironnements ou les concentrations géographiques de certaines expositions doivent être évaluées pour comprendre s'ils sont correctement inclus dans les scénarios retenus (IPCS, 2008).

Incertitudes liées aux modèles d'exposition

L'incertitude liée aux modèles concerne la capacité de ceux-ci à décrire de manière correcte et complète la réalité qu'ils sont censés caractériser, à travers la quantification des relations entre des données d'entrée et des réponses de l'ensemble ou d'une partie d'un système. Un modèle représente donc une simplification de la réalité, dont le niveau de détail doit être en accord avec les objectifs de l'évaluation. L'utilisation des modèles doit susciter une réflexion sur plusieurs aspects :

- la pertinence du modèle conceptuel sélectionné pour le scénario pour lequel il est utilisé
- les dépendances entre les paramètres
- les hypothèses faites lors de la modélisation ou de l'utilisation (ou l'extrapolation) de modèles déjà prêts pour une nouvelle situation d'exposition ou une nouvelle substance
- la pertinence du niveau de détail du modèle choisi (simple ou complexe) pour la situation d'exposition considérée (IPCS, 2008).

Choix méthodologiques applicables au calcul de l'exposition

Choix méthodologique pour le calcul de l'exposition : travailler en dose externe ou interne

Avant toute modélisation, certains choix méthodologiques détermineront le niveau de l'exposition où le calcul sera fait. Ainsi, choisir de travailler en dose interne ou externe pourrait conduire à des résultats différents et à des niveaux de robustesse différents des résultats finaux.

Modèle 1

Choix des paramètres du modèle

Paramètre exclu

Si certains paramètres ayant une influence importante sur la réalité simplifiée dans le modèle, ne sont pas considérés, le résultat est susceptible de représenter la réalité d'intérêt de manière inadéquate (incomplète ou incorrecte). Il est donc important de se poser la question si l'équation (le modèle) choisi(e) est complète ou est-ce qu'il y aurait des paramètres pertinents qu'on n'a pas considérés par manque de données, par manque de connaissances, etc.

A ce niveau, un certain nombre d'hypothèses doivent parfois être faites pour pouvoir utiliser un modèle déjà disponible ou légèrement adapté à une situation nouvelle. Il est important de garder la trace de ces hypothèses, pour pouvoir y revenir si les résultats de la modélisation s'avèrent finalement être en désaccord avec des mesures disponibles ou la connaissance de la réalité modélisée par les experts.

Choix de l'équation mathématique

Degré de nouveauté / d'expérience accumulée avec l'équation

Le niveau de confiance dans un modèle peut dépendre d'un certain nombre de facteurs, dont sa vérification antérieure pour d'autres situations d'exposition. Si l'expérience gagnée avec un modèle est grande,

notamment lorsqu'il a été déjà utilisé par des équipes différentes ou il a été publié dans la littérature à comité de lecture, on l'utilisera avec plus de confiance que si le modèle est nouveau.

Critère de choix du modèle pour les scénarios

Le développement ou le choix d'un modèle pour décrire un scénario peuvent impliquer des décisions portant sur la concordance entre les échelles de temps et d'espace utilisés - dans la modélisation et dans le scénario.

Pour caractériser cette source d'incertitude, les questions qu'on peut se poser sont : Comment a-t-on jugé que le modèle (l'équation) choisi(e) est adapté(e) aux scénarios ? Est-ce que la concordance avec les données mesurées a été un critère pour choisir le modèle ? Est-ce qu'il y a eu d'autres critères ? On s'est inspiré des expériences d'autres structures d'expertise ou de la littérature scientifique, par exemple ?

Corrélation entre les paramètres

Corrélations possibles en réalité mais potentiellement exclues dans le modèle

Selon l'**IPCS (2008)**, les « erreurs de dépendance » (*dependency errors*) sont associés au fait que les corrélations entre certains paramètres ne sont pas considérées dans l'équation mathématique du modèle, ce qui peut conduire à une modélisation incorrecte de l'exposition et/ou, par exemple, des extrapolations incorrectes des adultes aux enfants. Par exemple, des propriétés anthropométriques comme le poids corporel, la surface de la peau et le taux d'inhalation sont corrélées. D'autres exemples donnés par l'**IPCS (2008)** sont :

- ne pas inclure la corrélation possible entre la fréquence d'usage d'un produit de consommation et la quantité consommée par utilisation
- utiliser un taux d'émission constant pour décrire l'évaporation d'une substance sans prendre en considération le fait que le taux d'évaporation va dépendre de la concentration de la substance dans l'air et dans la matrice

Incertitudes liées aux données d'entrée qui alimentent les paramètres des modèles

Cette source d'incertitude concerne les valeurs numériques (ponctuelles ou des distributions) utilisées pour les paramètres du modèle (**IPCS, 2008**).

Critères de choix des données utilisées

Données mesurées

Pour évaluer l'incertitude associée aux données mesurées, plusieurs questions sont pertinentes :

- Est-ce que toutes les données publiques et privées ont été disponibles ? Par exemple, est-ce que le pourcentage d'entreprises ayant répondu à l'enquête filière est significatif, par rapport au nombre total d'entreprises ?
- Pour les données issues de la littérature parmi toutes celles disponibles, comment avoir choisi celles incluses finalement dans les calculs ?
- Est-ce qu'on a privilégié les données françaises ? Est-ce que ce choix a limité le nombre de données disponibles (alors que pour d'autres pays un nombre plus important de données était disponible) ?
- Y a-t-il une étude de littérature faite pour l'exposition ? Si oui, dans quelles bases de données, selon quels mots clés et selon quels critères on a considéré que les références sont valides pour l'ERS ?
- Si pertinent : comment a-t-on traité les désaccords existant entre diverses sources bibliographiques ?

Données modélisées

Pour les données modélisées, on peut se poser la question : Comment a-t-on choisi quand utiliser des données mesurées et quand des données modélisées ?

Modalités de vérification / d'ajustement de la qualité des données

Données modélisées

En général cette source n'est pertinente que pour les données modélisées, les données mesurées ne peuvent pas forcément être adaptées, et leur choix en fonction de leur qualité entre dans la source d'incertitude précédente.

Pour les données modélisées, on peut néanmoins revenir sur la modélisation si elles sont par exemple en décalage trop important avec les données mesurées disponibles, ou si les experts les considèrent en désaccord flagrant avec leur propre expérience des situations modélisées. La question pertinente pour cette source est : quels critères de qualité a-t-on utilisé pour améliorer la pertinence des données modélisées par rapport aux situations réelles considérées ? (à nouveau, il peut s'agir d'établir un certain facteur entre données modélisées et mesurées, d'utiliser le jugement d'expert, de faire varier le nombre de fois qu'on fait tourner le modèle, faire une analyse de sensibilité et peaufiner certains paramètres, etc.).

Paramètre x

Données mesurées pour le paramètre du modèle

Représentativité des données

Cette source est pertinente pour les données mesurées. La représentativité des données utilisées pour caractériser un paramètre dépend de la taille de l'échantillon, du nombre d'échantillons et de la méthode d'échantillonnage. Il s'agit ici donc de la représentativité au sens statistique du terme.

Si les données étrangères ou obtenues pour d'autres situations d'exposition sont utilisées, la pertinence de l'extrapolation doit être évaluée (et donc la représentativité », au sens non-statistique du terme, de ces données).

Selon l'IPCS, « *Si l'échantillon est représentatif, alors la deuxième considération est de savoir comment le choix au hasard des échantillons conduisent à des variations aléatoires dans l'estimation des statistiques d'intérêt, tels que l'écart type de la distribution d'échantillonnage de la variabilité interindividuelle. Il est souvent supposé qu'un petit échantillon ne peut pas être représentatif, mais il s'agit d'une idée fautive très répandue. La question de la représentativité dépend de la conception de l'étude et de la stratégie d'échantillonnage. Si l'on dispose d'un échantillon représentatif, même s'il est très petit, alors les méthodes statistiques classiques peuvent être utilisées pour tirer des conclusions concernant les distributions de certains éléments statistiques tels que la moyenne. Les implications d'un échantillon petit, aléatoire et représentatif sont que les intervalles de confiance estimés pour statistiques d'intérêt, tels que la moyenne, peuvent être très larges. La portée d'un intervalle de confiance pour une statistique est généralement influencée par la taille de l'échantillon, la variabilité de l'échantillon observé (i.e., l'écart type de l'échantillon) et de tout autre élément d'intérêt statistique. Par exemple, la distribution d'échantillonnage d'une valeur moyenne tend d'être asymptotiquement normale lorsque la taille de l'échantillon augmente. Cependant, la distribution d'échantillonnage de la variance est positivement dissymétrique* » (IPCS, 2008, p. 24)

Prise en compte de la variabilité

Pour évaluer cette source d'incertitude, on peut se poser plusieurs questions, telles que : comment a-t-on défini les distributions pour les paramètres des modèles dont les valeurs sont probabilistes ? Comment a-t-on décidé quels paramètres seront considérés de manière probabiliste et lesquels de manière déterministe ?

“Uncertainty in the determination of the statistical distribution used to represent distributed parameter values

Probability distribution models can be used to represent frequency distributions of variability or uncertainty distributions. When the data set represents variability for a model parameter, there can be uncertainty in any non-parametric statistic associated with the empirical data.

For situations in which the data are a random, representative sample from an unbiased measurement or

estimation technique, the uncertainty in a statistic could arise because of random sampling error (and thus be dependent on factors such as the sample size and range of variability within the data) and random measurement or estimation errors. The observed data can be corrected to remove the effect of known random measurement error to produce an “error-free” data set (Zheng & Frey, 2005).

If a parametric distribution (e.g. normal, lognormal, log logistic) is fit to empirical data, then additional uncertainty can be introduced in the parameters of the fitted distribution. If the selected parametric distribution model is an appropriate representation of the data, then the uncertainty in the parameters of the fitted distribution will be based mainly, if not solely, on random sampling error associated primarily with the sample size and variance of the empirical data. Each parameter of the fitted distribution will have its own sampling distribution. Furthermore, any other statistical parameter of the fitted distribution, such as a particular percentile, will also have a sampling distribution. However, if the selected model is an inappropriate choice for representing the data set, then substantial biases in estimates of some statistics of the distribution, such as upper percentiles, must be considered.

For some empirical data sets for variability in a quantity, there can be one or more values that are either below a limit of detection of the measurement instrument or zero for structural reasons (e.g. zero consumption rate of a particular food by a particular person on a particular day). The former case is an example of missing data, whereas the latter is an example of a mixture distribution that is composed of one component that describes the frequency of zero values and a second component that might be a continuous distribution (e.g. consumption rates in this case).

For distributions that contain missing data, there are a variety of techniques that can be used, depending on the type of missingness. Perhaps the most common case of missing data is when one or more attempted measurements are below the detection limit, which are often referred to as “non-detects”. This is referred to as a “left-censored” data set, in that small values below one or more threshold values (detection limits) are unobservable. Left-censored data are commonly dealt with using simplifying assumptions that, unfortunately, introduce bias into estimates of key statistics. For example, non-detects are often replaced with one half of the detection limit, and then mean values of the entire data set are estimated, taking into account both the non-detects and the detected values. For most data sets, the use of one half of the detection limit as a surrogate for the unknown non-detected value leads to a biased estimate of the mean of the empirical data set. Instead, statistical methods, such as maximum likelihood estimation, that use information regarding the proportion of the data set that is below one or more detection limits, the numerical values of the detection limits and the values of the observed data can provide more accurate estimates of a fitted distribution and its statistics (e.g. Zhao & Frey, 2004).

Structural uncertainties in a distribution can be represented by empirical distributions or a mixture of a discrete and parametric distribution, as described previously. Unless there is a need to partition the distribution into zero and non-zero components, failure to account for the frequency of zero values can lead to overestimation of the mean of the overall distribution.” (IPCS, 2008, p. 27-28)

Précision des mesures des paramètres 1 : méthodes analytiques de mesure

Les données d'entrée pour les paramètres du modèle peuvent être entachées d'erreurs de mesure, liées à la méthode analytique utilisée, au choix des limites de quantification et de détection, des imprécisions dans le protocole de collecte de données, etc.

Certaines de ces erreurs peuvent être incluses dans deux catégories :

- des erreurs de mesure aléatoires, qui donnent lieu à une variation autour de la « vraie » valeur : par exemple, la lecture d'une échelle avec une certaine précision, ou la communication de l'usage d'un produit par un consommateur, ont des chances égales de sous- ou surestimer la réalité. Les erreurs de mesure aléatoires sont souvent aggravées par la variabilité réelle sous-jacente dans la quantité mesurée. Par exemple, si l'on mesure la concentration ambiante d'un polluant de l'air à un endroit précis, il existe une variabilité temporelle de la concentration réelle. Si la méthode de mesure a une composante aléatoire de l'erreur de mesure, alors la variabilité des mesures dans le temps sera plus grande que la variabilité réelle de la concentration ambiante. Ne pas prendre en considération l'erreur de mesure aléatoire à partir des données observées peut conduire à une surestimation de la variabilité.

- Des erreurs de mesure systématiques, où la moyenne des mesures d'un facteur d'exposition peut différer de la « vraie » valeur. Cette différence s'appelle le biais et peut être le résultat, par exemple, l'étalonnage incorrect de l'appareil de mesure ou de la mauvaise compréhension des questions dans des questionnaires (IPCS, 2008, p. 24).

D'autres erreurs sont liées aux choix méthodologiques faits par l'expérimentateur, par exemple par le manque d'adaptation de la méthode (ex. : LD et LQ) aux caractéristiques de la matrice analysée.

Précision des mesures des paramètres 2 : données de substitution, extrapolations, facteurs par défaut, jugement d'expert

Dans certaines situations, l'utilisation de données de substitution, issues du jugement d'expert, des facteurs ou données par défaut ou des données extrapolées est nécessaire.

Selon l'IPCS (2008, p. 25-26) :

- Les données de substitution sont des substituts pour un paramètre qui sont utilisées pour estimer les valeurs analogues ou correspondantes d'un autre paramètre. Si les données de substitution ne sont pas représentatives du paramètre d'intérêt, alors leur utilisation peut conduire à des biais dans le calcul final de l'exposition.
- Le processus pour obtenir le jugement d'expert comprend généralement trois étapes : la sélection des experts, la sollicitation du jugement des experts et la communication des résultats. La sélection des experts est généralement destinée à identifier les personnes qui ont des connaissances approfondies des thèmes pour lesquels les jugements sont recherchés. Le protocole de sollicitation est une procédure formelle qui implique une conversation structurée avec un expert. Deux ou plusieurs jugements d'experts pour le même paramètre peuvent être utilisés, séparément ou combinés. Les incertitudes peuvent survenir pour plusieurs raisons. La première est que les experts choisis n'auraient pas été les personnes les plus appropriées pour le paramètre estimé. Enfin, les heuristiques cognitives classiques et bien connus associés au jugement en incertitude peuvent apparaître (ex. : l'ancrage à la première valeur énoncée, l'heuristique de la disponibilité⁴⁴ (Tversky et Kahneman, 1982), la tendance de être trop confiant par rapport à ses propres capacités d'estimation, etc.). Enfin, la manière dont on combine les différents jugements d'experts peuvent conduire à des incertitudes ».
- Les valeurs par défaut (qui peuvent être des distributions) sont des valeurs de référence recommandées pour une utilisation à des fins spécifiques. Le niveau d'incertitude associé à l'utilisation de valeurs par défaut dépend la qualité des données de base définie utilisées pour les calculer.
- Les données issues de l'extrapolation sont issues de l'inférence à partir de données connues (par exemple, les données sur une situation future à partir de données sur une situation passée). L'extrapolation procède par l'analyse des tendances et des hypothèses. L'incertitude concerne la représentativité des données extrapolées pour la situation pour laquelle elles sont utilisées.

L'étape de recensement du cas d'étude « Toluène » s'est appuyée sur cette classification. Un premier recensement a été réalisé par un expert membre du CT Incertitude spécifiquement en charge de ce point, avant d'être discuté avec l'ensemble de l'équipe de l'Anses en charge de l'évaluation de l'exposition au toluène, puis soumis au CT Incertitude.

⁴⁴ Les heuristiques de jugement sont des opérations mentales automatiques, intuitives et rapides, des raccourcis cognitifs utilisés par les individus afin de simplifier leurs opérations mentales dans le but de répondre aux exigences d'un milieu complexe. Par exemple, l'heuristique de disponibilité désigne un mode de raisonnement qui se base uniquement ou principalement sur les informations immédiatement disponibles, sans chercher à en acquérir de nouvelle concernant la situation. Un autre exemple est l'heuristique de l'ancrage, qui est la difficulté cognitive naturelle à se départir d'une première impression.

Références

IPCS (International Programme on Chemical Safety), 2006. Uncertainty and data quality in exposure assessment, WHO, IPCS harmonization project document no. 6. URL : <http://www.who.int/ipcs/methods/harmonization/areas/exposure/en/>

Ramsey, M.H., 2008. Uncertainty in the assessment of hazard, exposure and risk. *Environmental Geochemistry and Health*, 31(2):205-217.

Tversky, A. and Kahneman, D., 1982. Judgment under uncertainty: heuristics and biases. In: Kahneman, D., Slovic, P. and Tversky, A., 1982. *Judgment under uncertainty: heuristics and biases*, Cambridge University Press, New York, pp. 3 - 20

Annexe 2 : Analyse qualitative de l'incertitude dans le calcul de l'exposition au toluène – Résultats

Evaluation du jugement d'expert, représentation graphique de l'incertitude agrégée et de l'hétérogénéité inter-experts

Source
d'incertitude

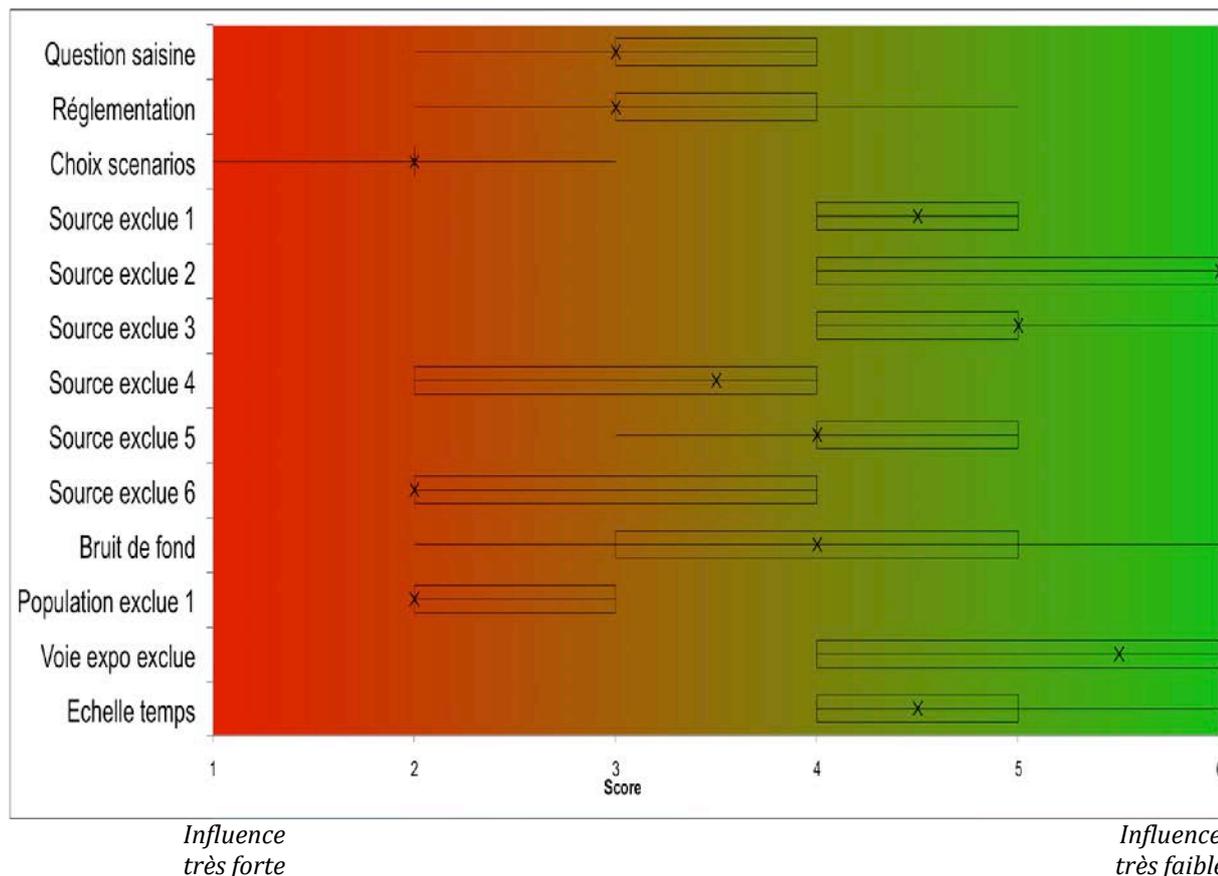


Fig. 3.1 Incertitudes contextuelles et dans les scénarios

Pour deux classes, contextuelles et dans les scénarios, on peut remarquer une relative homogénéité des réponses des experts, à l'exception du « bruit de fond » où la différence entre le minimum et le maximum, ainsi que l'intervalle interquartile, sont assez importants.

Pour les catégories « incertitudes contextuelles » et « incertitudes liées aux scénarios », les sources d'incertitude qui ont une influence très importante sur les résultats, selon les experts de manière agrégée (médiane en zone rouge ou orange), sont :

- Question posée dans la saisine
- Spécifications réglementaires
- Critères de choix des scénarios
- *Source d'exposition exclue 4* : mastics, durcisseur, insecticides, lubrifiants
- *Source d'exposition exclue 6* : cosmétiques
- *Populations exclues 1* : classes d'âge particulières : enfants

Pour une de ces sources d'incertitude, à savoir « *Source d'exposition exclue 6 : cosmétiques* », la raison de l'incertitude est de nature institutionnelle. En effet, le champ de responsabilités de l'Anses ne couvre pas les cosmétiques. Il est donc important d'accompagner la représentation semi-quantitative (à l'aide des scores) des incertitudes par des explications textuelles qui analysent les raisons des résultats fournis par les experts.

Source d'incertitude

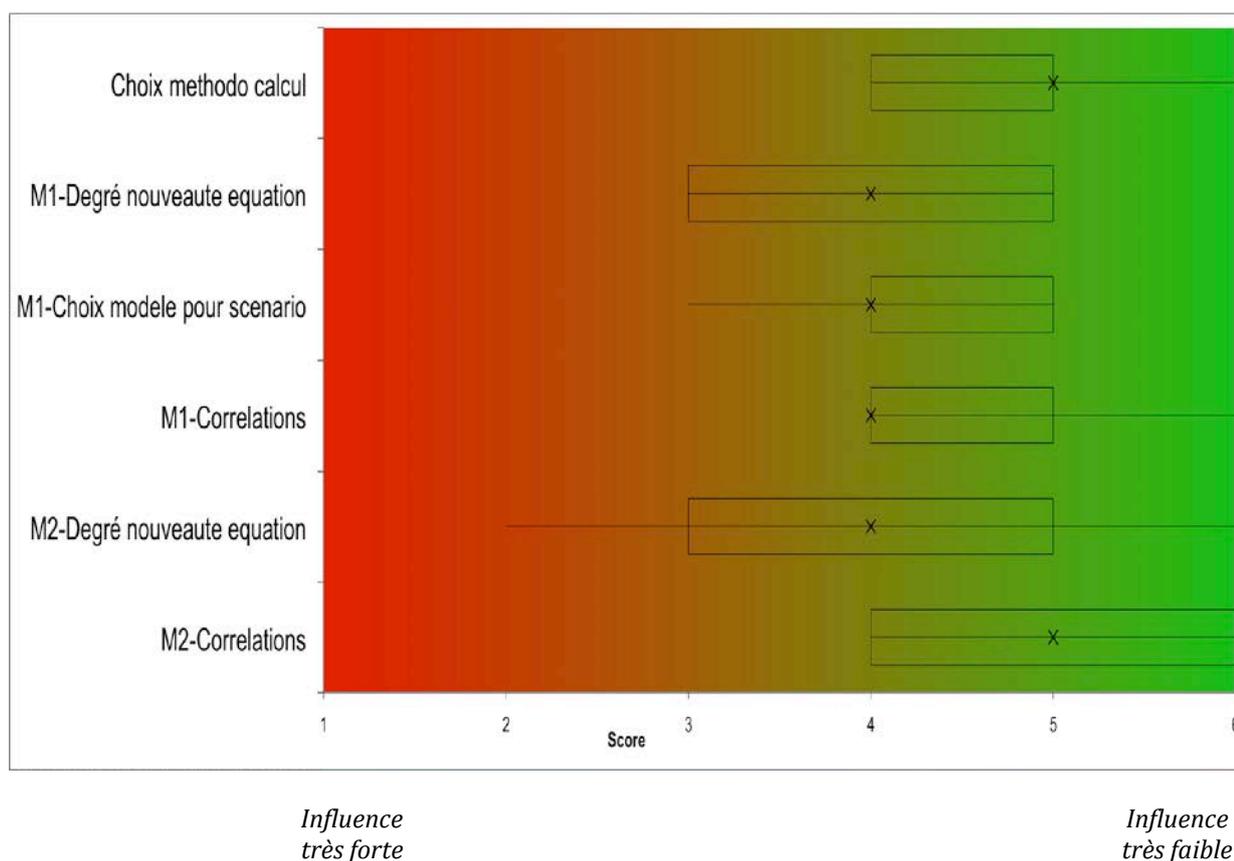


Fig. 3.2. Incertitudes liées aux modèles d'exposition

Pour la catégorie des incertitudes liées aux modèles d'exposition, la majorité des sources d'incertitude ont une influence relativement faible sur les résultats finaux (scores élevés). On observe également un niveau plutôt élevé d'homogénéité entre les experts.

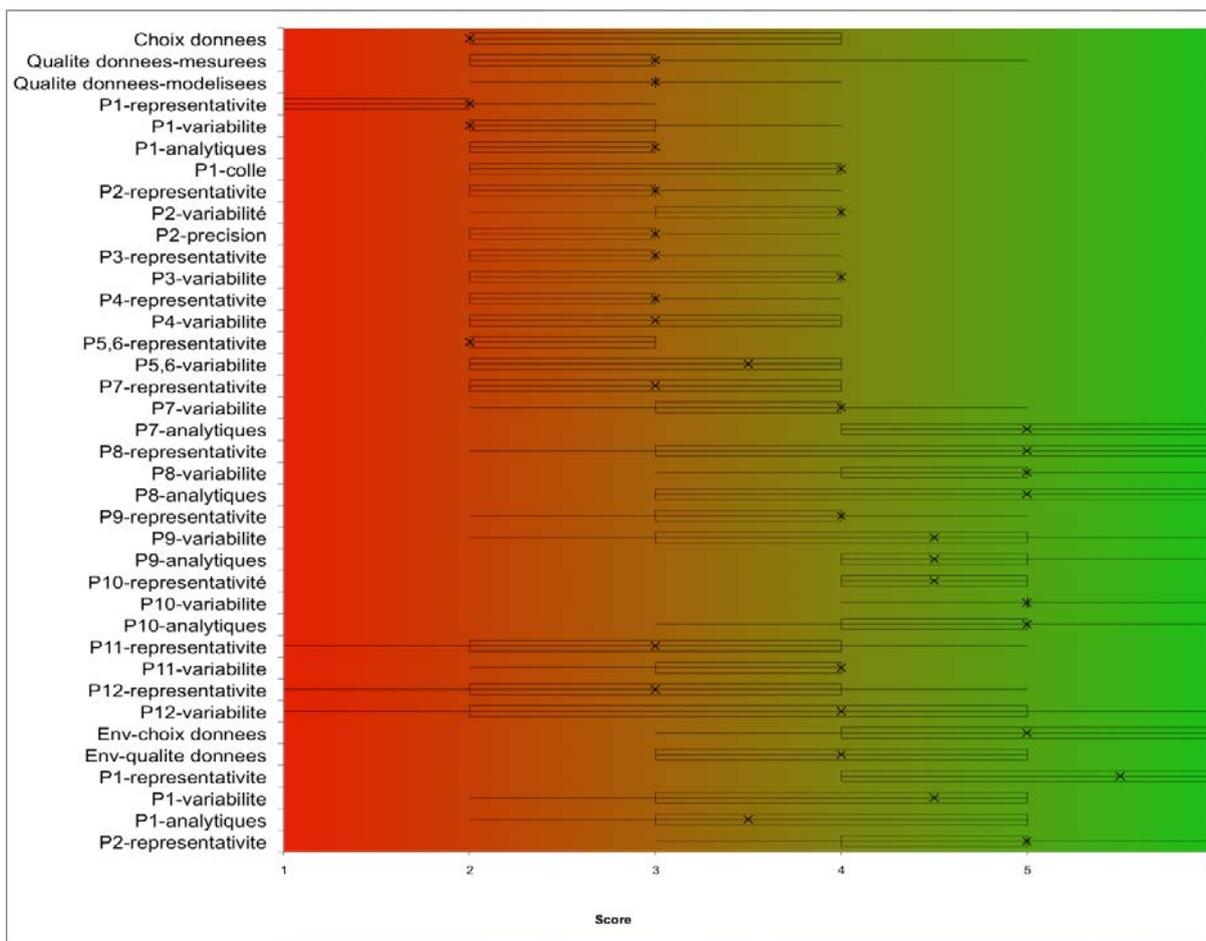


Fig. 3.3. Incertitudes dans les données d'entrée qui alimentent les paramètres des modèles

Un nombre assez important de sources d'incertitudes dans les données d'entrée, plus précisément 15 sur 38, se trouvent dans la zone rouge. Trois sources d'incertitude se caractérisent aussi par des intervalles interquartiles de plus de 2 points, ce qui indique une hétérogénéité importante des jugements d'experts. C'est en effet la catégorie d'incertitudes qui est la plus problématique, comme le montre également le fait que la plupart des experts n'ont pas su dire si ces incertitudes sous-estiment ou surestiment les résultats finaux.

Pour conclure, on peut observer que cette méthode peut jouer un rôle d'analyse de sensibilité qualitative, car elle permet de prioriser les sources d'incertitude les plus importantes. Par ailleurs, un niveau de confiance global dans le calcul de l'exposition peut être calculé à partir de cette évaluation.

Calcul d'un niveau de confiance global

Dans l'ensemble des sources d'incertitude, celles qu'on peut considérer comme étant les plus CRITIQUES sont celles :

- qui montrent une très grande hétérogénéité (4 ou 5 points de différence entre deux experts, la plus grande différence possible entre 2 experts étant de 5)
- qui reçoivent des scores très bas (au moins un expert donne un score de 1), donc on peut considérer qu'une opinion minoritaire forte subsiste même après la discussion des différentes sources d'incertitude dans le groupe,
- qui ont la médiane < 4 (« l'incertitude agrégée ») (zone orange ou rouge).

Pour le toluène, la matrice d'incertitudes contient 57 sources.

On peut établir un seuil de décision pour dire qu'une étude reçoit :

- un **haut degré de confiance**, s'il n'y a pas plus d'un tiers (19) des sources d'incertitude qui sont critiques,
- un **niveau moyen de confiance**, si entre un tiers et deux tiers des critères (20 – 37) sont dans les catégories mentionnées ci-dessus
- un **niveau réduit de confiance**, si plus de deux tiers des critères (38 ou plus) sont dans les catégories mentionnées

Pour le toluène, 27 sources d'incertitude peuvent être classées comme étant « critiques ». On peut donc conclure, selon la règle de décision établie, à un niveau moyen de confiance dans le calcul de l'exposition.

Le domaine disciplinaire de chacun des experts, ainsi que leur niveau de compétence sur le sujet spécifique traité, sont des aspects importants qui peuvent influencer les résultats de l'évaluation qualitative de l'incertitude mais qui n'ont pas été inclus dans la démarche dans ce rapport, pour des raisons de faisabilité. En effet, il s'agissait pour le toluène de vérifier la démarche méthodologique proposée sans demander un nombre important d'informations aux experts, et donc un effort trop important, ce qui aurait pu provoquer une attitude de rejet de leur part même avant d'avoir tenté l'exercice. Mais l'origine disciplinaire et la compétence sur un sujet particulier peuvent être facilement incluses dans la démarche, en demandant à chacun des experts participants de cocher une ou plusieurs cases qui correspondrait à leur discipline. Pour exemplifier ce type d'approche, l'annexe 5 reprend le questionnaire d'identification des experts qui a été utilisé dans une démarche similaire à celle proposée ici mais appliquée à l'évaluation de la qualité des études in vivo, dans un projet financé dans le cadre du PNRPE (intitulé Dico-Risque).

Ainsi, il est possible de comparer les réponses fournies par des experts ayant coché les mêmes cases disciplinaires, en réalisant deux représentations graphiques des réponses au lieu d'une seule (comme dans les figures 3.1 à 3.3). Les deux représentations graphiques illustreraient les réponses des experts du domaine disciplinaire d'intérêt, et ceux des autres domaines disciplinaires du groupe. Une autre possibilité est de réaliser deux représentations graphiques pour deux disciplines d'intérêt, ou pour deux niveaux de compétence sur le sujet différents (ex. : sous-groupe « expologie » et sous-groupe « dangers »). La comparaison des deux figures permettrait d'identifier des critères où les différences sont importantes entre les deux groupes disciplinaires ou de compétences. Les niveaux de confiance pourraient également être différents pour chacun des deux sous-groupes disciplinaires/de compétence.

Influence de chaque source d'incertitude sur la sur- ou sous-estimation des résultats finaux de l'exposition

Dans les tableaux 3.1 à 3.4, à la proposition des membres du CT Incertitude, uniquement les sources avec au moins deux réponses indiquant une sous- ou surestimation ont été retenues, les autres experts indiquant qu'ils ne peuvent pas se prononcer.

Les retours des experts comportaient un grand nombre de « je ne peux pas répondre ». Cette réponse concernait notamment les sources d'incertitude spécifiques aux données d'entrée qui alimentent les paramètres des modèles, notamment « représentativité des données », « prise en compte de la variabilité » et « précision des mesures des paramètres ». En effet, pour pouvoir calculer l'exposition en dépit des données manquantes, on utilise les données disponibles même si elles ne sont pas toujours représentatives des situations qu'elles mesurent ou mesurées avec les outils les plus précis disponibles, ou on fait des hypothèses sur la forme de leur distribution probabiliste. Il s'avère très difficile pour les experts de juger l'influence de ces contraintes sur le résultat final de l'exposition.

Tableau 3.1. Direction de l'influence des incertitudes contextuelles sur les résultats (le chiffre représente le nombre d'occurrences et le sens de la flèche indique s'il s'agit d'une sous ou d'une surestimation).

Localisation des sources d'incertitude	Evaluations	
Question posée dans la saisine, hypothèses réglementaires		
Question posée dans la saisine	8 x ↓	
Spécifications réglementaires	2 x ↓	4 x ↑

Tableau 3.2. Direction de l'influence des incertitudes liées aux scénarios sur les résultats

Localisation des sources d'incertitude	Evaluations
Choix des scénarios retenus	
Critères de choix des scénarios	4 x ↓
Choix des sources d'exposition (préparations) incluses dans l'ERS	
<i>Sources d'exposition exclues 1 :</i> <ul style="list-style-type: none"> - textiles, isolants thermiques pour équipements frigorifiques ou pompes à chaleur, matériels de sport, matelas, garnissage de sièges automobiles : fabriqués à partir de la synthèse de polyuréthane - colorants, pigments, caoutchoucs, plastiques, résines, fabriqués à partir de nitrotoluène, éthylbenzène et phénol - colorants, certaines huiles et paraffines extraites à partir du toluène 	6 x ↓
<i>Sources d'exposition exclue 2 :</i> appareils de mesure type thermomètre	5 x ↓
<i>Source d'exposition exclue 3 :</i> produits finis/articles destinés au grand public, notamment meubles, journaux...	6 x ↓
<i>Source d'exposition exclue 4 :</i> mastics, durcisseur, insecticides, lubrifiants	7 x ↓
<i>Source d'exposition exclue 5 :</i> lessive en poudre, additif pour essence, nettoyant pour carburateur, allume feu liquide	6 x ↓
<i>Source d'exposition exclue 6 :</i> cosmétiques	5 x ↓
<i>Source d'exposition non cumulée :</i> le « bruit de fond »	7 x ↓

Localisation des sources d'incertitude	Evaluations
Choix des populations ciblées incluses dans l'ERS	
Populations exclues 1 : classes d'âge particulières : enfants	3 x ↓
Choix des voies d'exposition incluses dans l'ERS	
Voies d'exposition exclues : exposition par ingestion et par contact cutané direct et indirect	8 x ↓

Tableau 3.3. Direction de l'influence des incertitudes liées aux modèles sur les résultats

Localisation des sources d'incertitude	Evaluations	
Modèle d'exposition au Toluène par inhalation : Modèle de diffusion dans une pièce avec air uniformément mélangé suite à une émission constante + Modèle de diffusion dans une pièce avec air uniformément mélangé suite à une émission instantanée		
Choix méthodologiques applicables à l'évaluation des expositions		
Choix méthodologique pour le calcul de l'exposition 1 : travailler en dose externe ou interne	↓	↑
Choix de l'équation mathématique		
Degré de nouveauté ou d'expérience accumulée avec l'équation / vérification de la capacité de l'équation de décrire les phénomènes concernés	2 x ↓	2 x ↑
Critère de choix du modèle pour les scénarios	2 x ↑	
Modèle d'exposition au Toluène par contact : Modèle de contact direct avec la substance + Modèle de contact constant avec la substance		
Choix de l'équation mathématique		
Degré de nouveauté ou d'expérience accumulée avec l'équation / vérification de la capacité de l'équation de décrire les phénomènes concernés	2 x ↑	

Tableau 3.4. Direction de l'influence des incertitudes liées aux données d'entrée des modèles sur les résultats

Localisation des sources d'incertitude	Evaluations
EXPOSITION LIEE À L'UTILISATION DE PRÉPARATIONS CONTENANT DU TOLUENE	
Critères de choix des données utilisées	
<i>Données mesurées vs modélisations</i>	2 x ↑
Modalités de vérification / d'ajustement de la qualité des données	
<i>Données mesurées</i>	2 x ↑
<i>Données modélisées</i>	4 x ↑
PARAMETRES SPECIFIQUES AUX MODELES	
Modèle d'exposition au Toluène par contact: Modèle de contact direct avec la substance + Modèle de contact constant avec la substance	
EXPOSITION PAR LES ENVIRONNEMENTS DOMESTIQUES OU EXTÉRIEURS	
Paramètre 2 : Concentration d'exposition via les poussières sédimentées	
<i>Représentativité des données</i>	5 x ↓

Les tableaux 3.1 à 3.4 montrent qu'il existe un taux élevé d'homogénéité entre les experts concernant la direction d'influence des sources d'incertitude. Pour peu de sources, les experts ont des avis différents, certains indiquant une surestimation et d'autres une sous-estimation, pour la même source (ex. : Choix de l'équation mathématique).

Annexe 3 : Expert Profile

Name (for recording, the analysis will treat the answers anonymously):

Institution (for recording, the analysis will treat the answers anonymously):

Function (for recording, the analysis will treat the answers anonymously):

Please respond to the following questions:

1. What are your main research discipline(s) and area(s)?

** multiple answers possible;*

** you can chose either the overall disciplinary domain (e.g. BIOLOGY), or select one subclass / several subclasses (e.g., cellular biology, molecular biology)*

BIOLOGY

- Genetics (including epigenetics)
- Reproductive biology (e.g., biology of reproductive hormones)
- Cellular biology
- Molecular biology
- Developmental biology
- Human physiology
- Anatomy (human)
- Neuroscience (human)
- Study of human behavior, psychobiology
- Endocrinology
- Epidemiology
- Andrology
- Zoology
- Chemical Biology
- Biochemistry
- Ecology
- Evolutionary Biology
- Other (*please specify*)

MEDICINE

- Gynecology
- Obstetrics
- Paediatrics
- Pathology
- Oncology
- Clinician
- Pharmacology
- Physical medicine
- Embryologist
- Other (*please specify*)

(HUMAN) TOXICOLOGY

- Reproductive toxicology
- Computational Toxicology
- Molecular toxicology
- Carcinogenesis
- Teratogenicity
- Toxicogenomics
- Pharmacokinetics and pharmacodynamics
- Other (*please specify*)

CHEMISTRY

- environmental chemistry
- analytical chemistry
- expology (exposure to contaminants)
- Other (*please specify*)

PHARMACY

ECO-TOXICOLOGY

VETERINARY SCIENCES

STATISTICS

CHEMICAL RISK ASSESSMENT

ENVIRONMENTAL HEALTH SCIENCES

REGULATORY TOXICOLOGY

ENDOCRINE TOXICITY (TOXICITY OF ENDOCRINE DISRUPTERS)

COMPUTER SCIENCE

- bioinformatics
- Other (*please specify*)

OTHER : *please specify*

2. Please attach your list of publications; please underline those on BPA and/or endocrine disrupters
3. Your work on BPA is:
 - experimental (e.g., in laboratory)
 - theoretical (e.g., expertise)
 - both
 - none (never worked on BPA and or endocrine disruptors)
4. Do you define yourself as having:
 - a general or interdisciplinary knowledge of bisphenol A; *please explain (one sentence)*
 - a specialized knowledge of bisphenol A; *please explain (one sentence)*

5. Expert affiliation (your main employer):

- safety agency
- academic scientists
- private scientist (industry)
- non-gouvernemental organization
- independent consultant / consultancy company

6. Do you have contracts with the industry?

- yes
- no

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](https://twitter.com/Anses_fr)