

DOMAINES ET SOURCES D'INCERTITUDE EN ÉPIDÉMIOLOGIE ANIMALE *

Toma Bernard¹ et Dufour Barbara¹



RÉSUMÉ

Après avoir rappelé des notions de base sur le concept d'incertitude, quotidiennement rencontré dans la vie courante et dans le domaine scientifique, ce texte présente quatre grands types d'incertitude et leurs sources :

- L'incertitude « épistémologique » (ou épistémique), liée à une insuffisance de connaissance ;
- L'incertitude « méthodologique », résultant de la méthode utilisée ;
- L'incertitude « stochastique », en relation avec le hasard, l'hétérogénéité des populations et la variabilité de leurs sujets ;
- L'incertitude « terminologique », dépendant du vocabulaire utilisé pour communiquer.

L'incertitude globale rencontrée dans une situation peut correspondre à un type ou à la combinaison de plusieurs types d'incertitude. Une partie seulement de l'incertitude globale peut être maîtrisée.

Mots-clés : incertitude, épidémiologie animale, analyse de risque.

ABSTRACT

This paper reviews some basic notions regarding the concept of uncertainty routinely encountered in everyday's life and in the scientific field, and proceeds to characterize four major types of uncertainty and their origin:

- *Epistemic uncertainty, associated with a gap in knowledge;*
- *Methodological uncertainty, resulting from the method used;*
- *Stochastic uncertainty, related to chance, to population heterogeneity and/or to variability among individuals;*
- *Linguistic uncertainty, related to the vocabulary used to communicate.*

The overall uncertainty encountered in a given situation results from one type of uncertainty or, more frequently, from a combination of several types of uncertainty. Only a part of the overall uncertainty can be mastered.

Keywords: *Uncertainty, Animal epidemiology, Risk analysis.*



* Texte de la conférence présentée au cours de la Journée scientifique AEEMA, 19 mars 2015

¹ École nationale vétérinaire, 94704 Maisons-Alfort, France

L'incertitude est un concept aux contours... incertains que les scientifiques rencontrent fréquemment dans leur quête de vérité.

Dans deux situations très fréquentes du domaine de la santé, celle du diagnostic et celle de la recherche, leur démarche, pour aboutir à l'identification de la maladie ou à la vérification de l'hypothèse élaborée, a pour objectif de limiter l'incertitude et d'aboutir à ce qui peut être considéré comme la vérité.

Pour un personnel de santé confronté en permanence à une situation d'incertitude pour établir un diagnostic sur un sujet ou un effectif, le risque est de finir par sous-estimer l'incertitude et de surestimer la fiabilité des conclusions auxquelles il aboutit. Avec, bien sûr, de temps en temps, des situations désagréables, incitant à davantage d'humilité lorsque la preuve de l'erreur est apportée par des investigations complémentaires, par l'évolution de la maladie, voire par un compte-rendu d'autopsie.

L'incertitude fait également partie du vécu quotidien de tout décideur, en particulier dans le domaine de la santé : d'une part, celle qu'il rencontre dans les rapports et avis qu'il reçoit, qu'il doit comprendre et utiliser au mieux ; d'autre part, celle liée à la prise d'une décision en tenant compte d'un ensemble complexe d'éléments souvent divergents et eux-mêmes incertains.

L'impression ressentie par les auteurs, à l'occasion d'activités d'expertise collective au sein d'une

agence sanitaire au cours des dernières années, a été une augmentation de l'intérêt porté par les gestionnaires de santé à l'incertitude liée aux conclusions des avis, à sa prise en compte au cours de l'expertise, à son estimation et son expression.

Cette sollicitation a, très logiquement, conduit les experts à se pencher plus spécifiquement sur l'incertitude rencontrée en expertise collective, afin d'en définir les contours, d'en préciser les origines potentielles, d'en standardiser l'expression et de diminuer l'incertitude sur laquelle il est possible d'agir.

Pour l'introduction de cette Journée consacrée à l'incertitude en épidémiologie, compte tenu des conceptions variées rencontrées dans ce domaine vaste et complexe, il nous a paru souhaitable, en tant que non spécialistes de ce thème, d'en explorer les bases, sans entrer dans des arcanes, notamment statistiques ou terminologiques, et de tenter de proposer aux intervenants une terminologie commune.

Pour le concept d'incertitude, on peut utiliser la définition générale suivante :

« État de doute sur un concept, un objet, un sujet, une population ou sur la probabilité et les conséquences de la survenue d'un évènement ».

Après avoir rappelé les grands domaines et les principales sources de l'incertitude, l'attention sera focalisée sur l'incertitude en épidémiologie et en analyse de risque ainsi que sur son expression.

I - LES DOMAINES DE L'INCERTITUDE

Incertain signifie donc absence ou insuffisance de certitude. Pratiquement, on peut la rencontrer dans tous les domaines :

- La justice : les juges et les jurés sont en permanence aux prises avec l'incertitude quant à la culpabilité de l'accusé ;
- Les arts : la beauté est une notion subjective, diversement appréciée par les personnes comme en témoignent les sommes astronomiques atteintes par certaines œuvres loin de faire l'unanimité ;
- Toutes les sciences, même les plus exactes, sont confrontées à l'incertitude, aussi bien dans l'infiniment grand que dans l'infiniment petit, dans l'inerte que dans le vivant ;

- Les croyances n'y échappent pas ; il s'agit d'ailleurs sans doute du domaine pour lequel l'amplitude de l'incertitude à propos d'un même sujet est maximale en fonction des personnes, nulle pour certaines, totale pour d'autres ;
- Dans la vie courante, l'incertitude est rencontrée plusieurs fois par jour, à propos d'informations reçues, d'objets ou de personnes rencontrées, de décisions à prendre, etc.

Quel que soit le domaine, elle peut concerner des unités, prises séparément (concept, objet, sujet, évènement) ou en groupes (population).

1. UNITÉ

L'incertitude peut porter sur une unité.

⇒ S'il s'agit d'un *concept au sens large* (exemples : le big-bang créateur de l'univers, le réchauffement climatique, etc.), l'incertitude est inversement proportionnelle à la capacité de la démonstration du concept. Par suite, ce qui entre dans le champ de la raison peut ne comporter qu'une incertitude limitée.

⇒ Pour un *objet*, élément que l'on considérera comme invariable au cours du temps, l'incertitude quant à ses caractéristiques est inversement proportionnelle aux performances des outils permettant de les mesurer.

⇒ Par *sujet*, il faut entendre tout élément vivant (Homme, animal, végétal, agent pathogène biologique) dont, de ce fait, certaines caractéristiques peuvent évoluer au cours du temps. Dans ce cas, à l'incertitude liée aux performances de l'outil de mesure vient s'ajouter celle relative au moment où la mesure est faite : la valeur de nombreux paramètres d'un sujet (glycémie, diverses autres concentrations

sanguines, etc.) subit des variations intra-individu (« *variabilité* ») qui empêchent de donner une valeur unique pour ce sujet et permettent seulement d'en connaître, après coup, leur évolution au cours du temps et, donc, leur distribution de fréquence.

Notons au passage que la variabilité correspond à l'évolution de la valeur d'un paramètre au cours du temps chez un même sujet et qu'il est souhaitable de la distinguer de la « *variété* » (ou diversité ou hétérogénéité) constatée entre des sujets différents² (cf. tableau 1) : ainsi, il existe une grande variété de couleur des yeux chez l'Homme mais, chez une même personne, elle n'est que très peu variable.

⇒ Un *évènement* : la survenue d'un évènement dans le futur est, comme pour tout ce qui est « à venir », inévitablement accompagné d'incertitude et l'objet d'analyses de risque dans des domaines très divers.

Tableau 1
Différence entre « variété » et « variabilité »

Mot	Synonymes	Exemples
« Variété » Caractère de ce qui est <i>varié</i> , c'est-à-dire de ce qui présente des aspects distincts	Pluralité, hétérogénéité, diversité...	Qualitative : les races au sein d'une espèce Quantitative : le poids moyen des chiens en fonction de la race
« Variabilité » Caractère de ce qui est <i>variable</i> , c'est-à-dire de ce qui est susceptible de changer	Evolution, transformation, changement...	Evolution du poids d'un même sujet au cours de la vie

2. POPULATION

⇒ Au sein du domaine des *concepts*, on peut placer l'existence de relations entre eux ainsi que la notion de causalité entre un facteur et un évènement. Pour limiter l'incertitude, on y retrouve la notion de

démonstration, plus ou moins difficile à apporter, voire certaines fois impossible.

⇒ Pour une population d'*objets*, la variabilité est souvent exclue mais demeure la variété ou hétérogénéité.

² Cette confusion, souvent rencontrée, découle peut-être de l'usage fréquent du mot « variable » pour désigner l'élément étudié et qui, de façon qualitative ou quantitative, peut avoir un état ou une valeur différente d'un objet ou d'un sujet à un autre sans pour autant être...variable au sens propre, c'est-à-dire au cours du temps chez l'objet ou le sujet. Ainsi, de manière qualitative, on peut étudier la « variable » race au sein d'une espèce animale sans que la race d'un animal varie en quoi que ce soit, ou de façon quantitative, dans une enquête de prévalence, n'importe quelle « variable » (nombre d'animaux dans le troupeau...) sans que sa valeur change au moment de l'enquête.

Pour une population en petit nombre, il est possible de mesurer les caractéristiques de chacun des objets ; dans ce cas, demeure l'incertitude due aux

performances des outils de mesure et de ceux qui effectuent les mesures (« *mesureurs* » pour certains). En revanche, pour une population

nombreuse, les mesures ne sont faites, en général, que sur un échantillon et, par suite, à cette incertitude viennent s'ajouter d'une part, celle due au *hasard*, générateur de fluctuations d'échantillonnage et, donc, d'une estimation exprimée par un intervalle de confiance d'autant plus grand que la taille de l'échantillon est restreinte, et comportant des risques d'erreur ; d'autre part, celle pouvant être due à la méthode utilisée, génératrice de biais et, donc, d'inexactitude.

⇒ Pour une population de nombreux *sujets*, par définition vivants, la complexité (et donc le nombre de sources d'incertitude) augmente car à l'incertitude de mesure s'ajoute celle due à l'échantillonnage, relative, d'une part, à l'hétérogénéité de la population (variété des caractéristiques mesurées) et, d'autre part, à la variabilité.

Cette diversité de situations est illustrée par le tableau 2.

Tableau 2
Exemples de domaines d'incertitude et de leurs sources

Domaine	Etendue du domaine	Source(s) d'incertitude
Concept	Unité Hypothèse (exemples : existence du big-bang, du réchauffement climatique, <i>etc.</i>)	Difficulté, voire impossibilité de démonstration
	Population Relation de causalité au sein d'un ensemble (exemple : ondes électro-magnétiques et maladies)	
Objet ou sujet	Unité Caractéristiques de l'objet ou du sujet	<ul style="list-style-type: none"> • Outil de mesure • Mesureur
	Population Caractéristiques de la population d'objets ou de sujets	<i>En plus :</i> <ul style="list-style-type: none"> • Fluctuations d'échantillonnage : Hétérogénéité et variabilité → imprécision • Biais → inexactitude

II - LES SOURCES DE L'INCERTITUDE

Dès 1992, Wynne distinguait pour les questions d'environnement quatre types d'incertitude (*cf.* encadré 1).

Dans la littérature scientifique, on trouve d'autres expressions pour qualifier l'incertitude. Ainsi, de nombreux auteurs distinguent formellement la variabilité de l'incertitude et insistent pour que leur estimation soit distincte : par variabilité, ils entendent, en fait, l'hétérogénéité individuelle et

spatiale au sein d'une population et, par incertitude, ce qui découle d'une insuffisance de connaissance ou de l'imprécision de la méthode de mesure [Rai et Krewski, 1998 ; Zheng et Frey, 2005 ; Pouillot et Deline-Muller, 2010 ; Busschaert *et al.*, 2011 ; Bier et Lin, 2013]. Toutefois, certaines expressions ne simplifient pas la compréhension comme, par exemple : « *Total uncertainty is the combination of uncertainty and variability* » [Vose, 2008].

Encadré 1**Les différents types d'incertitude en environnement**

[Wynne, 1992]

- **Risque** (*risk*) : dans un système bien connu, les probabilités de survenue des conséquences peuvent être définies et quantifiées (odds connus).
- **Incertitude** (*uncertainty*) : les principaux paramètres sont connus, mais pas les distributions de probabilité.
- **Ignorance** (*ignorance*) : on ne sait pas exactement ce que l'on ne sait pas.
- **Indétermination** (*indeterminacy*) : les chaînes de causalité où les réseaux sont ouverts.

D'autres distinguent les incertitudes « faibles » (*weak uncertainties*), pouvant être exprimées en termes probabilistes (*probabilistic uncertainties*, *statistical uncertainties* ou *risks*) des incertitudes « fortes » (*strong uncertainties*, appelées également *scenario uncertainties* ou simplement *uncertainty*) lorsqu'une série de conséquences peut être connue mais sans qu'il soit possible de leur attribuer une probabilité [Christley *et al.*, 2013].

On trouve également la distinction entre incertitude réductible (*reducible or epistemic uncertainty*) et incertitude irréductible (*irreducible or ontological uncertainty*) [Fish *et al.*, 2011 ; Christley *et al.*, 2013].

Pour simplifier la terminologie de l'incertitude et rendre le raisonnement plus général, nous proposons de retenir quatre grandes catégories de sources d'incertitude (tableau 3) : l'insuffisance de connaissance, la méthode utilisée, l'imprévisibilité et le vocabulaire utilisé. Dans certains cas, ces sources peuvent alimenter simultanément la même incertitude globale.

1. L'INSUFFISANCE DE CONNAISSANCE

Sur un élément, quel qu'il soit (concept, objet, sujet) une insuffisance de connaissance, et, *a fortiori*, une absence de connaissance, entraîne une situation d'incertitude à son propos.

Tableau 3**Les grandes catégories de sources de l'incertitude**

Source	Type d'incertitude
Insuffisance de connaissance	Epistémologique (ou épistémique)
Méthode	Méthodologique
Imprévisibilité	Stochastique
Vocabulaire	Terminologique

L'incertitude est essentiellement fonction du degré de connaissance de l'élément d'intérêt : plus les informations sur celui-ci sont nombreuses et fiables, plus l'incertitude corollaire, nommée « *incertitude épistémique*³ » (ou épistémologique) (tableau 3) est réduite. A la limite, l'incertitude épistémique peut être réduite à néant en cas de connaissance parfaite de l'élément d'intérêt.

2. L'INCERTITUDE MÉTHODOLOGIQUE

Elle est fonction de la méthode utilisée.

La méthode choisie peut, selon les cas, être dépendante ou non des connaissances disponibles. On peut en effet utiliser une méthode mal adaptée, et donc génératrice d'une incertitude importante, tout en disposant d'une bonne connaissance du domaine étudié. A l'inverse, la disponibilité de connaissances précises et exactes sur une question peut faciliter la mise au point d'une méthode bien adaptée, entraînant une incertitude méthodologique restreinte.

Incertain épistémique et incertain méthodologique peuvent donc interférer ou non. On peut proposer d'appliquer l'appellation d'incertitude épistémique à ce qui a trait à l'insuffisance des connaissances disponibles sur le domaine d'étude et celle d'incertitude méthodologique à celle découlant de l'emploi de méthode(s) pour l'étude effectuée.

Les limites de précision des outils de mesure entraînent de l'incertitude. Toute mesure, quelle que soit la grandeur mesurée (longueur, masse, volume, vitesse, densité optique, etc.), fournit un résultat ne correspondant pas forcément à la valeur réelle et comportant une erreur génératrice d'*incertitude absolue* et d'*incertitude relative*. La première est l'estimation de l'écart maximal possible entre la mesure et la valeur exacte (souvent inconnue). La seconde est le rapport entre l'incertitude absolue et la mesure.

Il en est de même lors de l'emploi de tout test dans une démarche de dépistage ou de diagnostic, l'erreur individuelle n'étant plus quantitative mais qualitative (faux positif ou faux négatif).

3. L'IMPRÉVISIBILITÉ

Liée au hasard, elle est une autre source d'incertitude.

La survenue à court terme d'un événement peut être régulièrement prévue de façon exacte et précise. Ainsi, par exemple, les heures de lever et de coucher du soleil, pratiquement en chaque point du globe terrestre, peuvent être prévues avec exactitude et précision. Il en est de même pour les éclipses.

Mais, pour la plupart des événements, l'estimation de leur survenue comporte une incertitude liée au hasard ou « *incertitude stochastique* » (tableau 3).

Pour la survenue d'un événement, il convient de distinguer deux notions : la probabilité de survenue et la survenue elle-même.

3.1. PROBABILITÉ DE SURVENUE

L'estimation ou le calcul de la probabilité de survenue d'un événement est plus ou moins complexe en fonction de la situation.

Pour un système simple (dé à jouer, jeu de cartes, urne contenant des boules ou autres...), la probabilité de survenue dépend strictement du nombre de résultats favorables et du nombre de résultats possibles. Par suite, l'incertitude sur la probabilité de survenue est inversement proportionnelle à la connaissance de ces éléments. Ainsi, cette incertitude est nulle pour un dé à jouer (une chance sur six), pour un jeu de cartes (une chance sur 32 ou 56), ou pour une urne dont la composition de la population des boules est parfaitement connue. Dans une telle situation, l'incertitude sur la probabilité de survenue est indépendante de la probabilité de survenue. Elle est nulle aussi bien pour une probabilité élevée (dé à jouer) que pour une probabilité très faible (urne de 1 000 boules dont une seule rouge). En revanche, si la composition de la population dans l'urne n'est que partiellement connue (par sondage, par exemple), l'incertitude est d'autant plus grande que le taux de sondage est faible.

³ Le TLF(i) mentionne l'adjectif *épistémique* dans l'article *épistémologie*, et plus précisément dans une remarque. Le mot y est relevé en 1968 dans un traité de sociologie ainsi que le substantif féminin *épistémè*, mot grec repris notamment par Michel Foucault dans son célèbre ouvrage *Les Mots et les Choses* paru en 1966 et cité par OED² (s.v. *episteme*) ainsi que des traductions anglaises de ce livre à fort retentissement international ou des références à ce dernier. Quant à l'anglais *epistemic*, forgé sur le substantif grec επιστήμη, il est attesté dès 1922, tandis qu'*epistemological* apparaît quelques décennies plus tôt, en 1882 (OED²). Les deux adjectifs *épistémologique* et *épistémique* sont donc possibles ; le plus récent, soutenu par l'anglais *epistemic*, a l'avantage de la concision (Nadine Steinfeld).

Pour un système complexe (par exemple, évènement sanitaire dans une population), la probabilité de survenue dépend d'un ensemble complexe de facteurs interagissant au sein de la population. Son estimation est plus complexe. L'incertitude sur la probabilité de survenue dépend, outre du taux de sondage, du degré de connaissance de la population et de la fréquence de l'évènement.. Lorsque la population est imparfaitement connue, l'incertitude est élevée. C'est le cas, par exemple pour la probabilité d'introduction d'une maladie dans une région par des animaux sauvages, car le nombre et l'amplitude des déplacements de ces animaux ainsi que leur taux d'infection sont mal connus. Il en est également ainsi pour l'estimation de l'introduction d'une maladie dans un pays par des modalités illégales, desquelles la fréquence ne peut être qu'estimée et, parfois, de manière très approximative (exemple : introduction de produits d'origine animale par des voyageurs).

L'incertitude sur la probabilité de survenue d'un évènement est réductible à proportion de l'amplitude des connaissances sur la population concernée et les facteurs conditionnant sa survenue.

3.2. SURVENUE ELLE-MÊME

Quelle que soit l'incertitude liée à la probabilité de survenue d'un évènement, l'occurrence de cet évènement reste incertaine (sauf si cette probabilité de survenue était de 1 ou de 0, sans incertitude).

Ainsi par exemple, pour un lancer de dé à jouer, on connaît, sans incertitude, la probabilité d'obtenir un six (ou n'importe quelle autre valeur de un à six) lors d'un lancer, mais il est impossible de savoir le chiffre qui sera obtenu. Il en est ainsi aussi lors d'un tirage aléatoire d'une carte parmi un jeu de 32 ou 56 cartes : on connaît exactement la probabilité pour chaque carte, mais on ne peut pas prévoir la carte tirée. Il en est de même pour le nombre de sujets infectés au sein d'un échantillon tiré au sein d'une population ayant une prévalence donnée d'une

maladie : on connaît exactement les probabilités d'obtenir tel ou tel nombre de sujets infectés dans un échantillon de taille donnée, mais pas celui dans un échantillon qui vient d'être tiré. Par ailleurs, le lieu et le moment de la survenue d'un évènement sanitaire (apparition ou émergence d'une maladie) demeurent incertains. Même si la probabilité de survenue de l'évènement est bien connue (incertitude nulle ou très faible), il est impossible de connaître avec certitude à l'avance le lieu et le moment de survenue de cet évènement sanitaire. On sait par exemple que la probabilité d'introduction d'un chien en incubation de rage en France au cours d'une année est de l'ordre de 0,5 (soit une fois tous les deux ans). Mais il est tout à fait impossible d'en prévoir le lieu ni la date.

Cet aspect de la survenue d'un évènement correspond à une incertitude irréductible.

4. LE VOCABULAIRE

Le vocabulaire utilisé peut être source d'imprécision et, donc, d'incertitude. Qualifier la probabilité de survenue d'un évènement de « faible » ou « d'élevée » est, certes, informatif mais peut être compris ou interprété différemment en fonction des situations par une même personne ou, pour une même situation, par différentes personnes.

Les concepts et le vocabulaire utilisés, en particulier pour les résultats d'une analyse de risque quantitative, peuvent être en partie incompréhensibles pour le destinataire de l'analyse et entraîner une incertitude quant à l'interprétation à faire des résultats.

Ce type d'incertitude est qualifié d'« *incertitude terminologique* » (tableau 3).

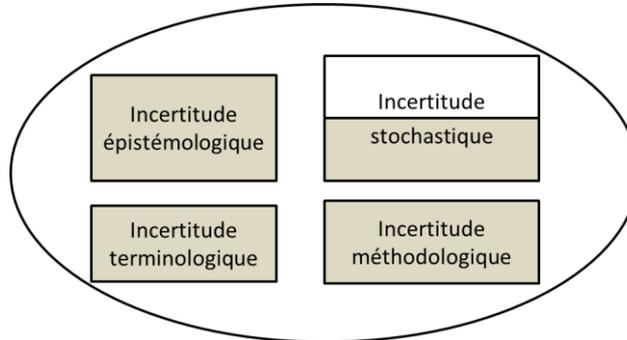
Il peut être réduit grâce à une attention particulière portée à la compréhensibilité des concepts et du vocabulaire utilisés.

La figure 1 représente les quatre principales sources d'incertitude et distingue la part réductible de la part irréductible.

Figure 1

Les quatre principales sources d'incertitude

(En gris, parties réductibles de l'incertitude ; en blanc, partie irréductible)



5. UNICITÉ OU MULTICPLICITÉ DES SOURCES D'INCERTITUDE

Selon les situations, une seule catégorie de source peut intervenir ou plusieurs (cf. figure 2).

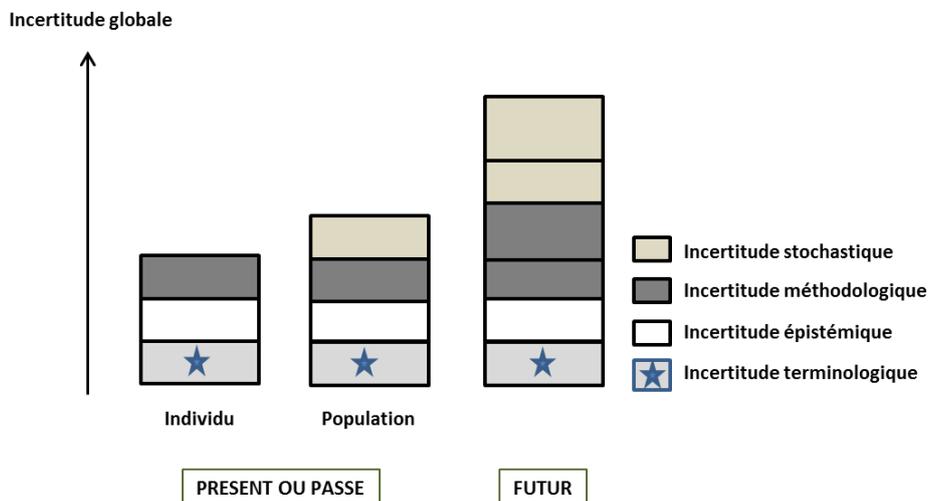
- Ainsi, pour la description du passé ou du présent d'un concept, d'un objet ou d'un sujet, l'incertitude résulte essentiellement d'un manque de connaissance plus ou moins important (incertitude épistémique) et, éventuellement, de l'emploi d'un vocabulaire insuffisamment précis (incertitude terminologique). En cas d'étude effectuée, peut s'y ajouter une incertitude méthodologique (figure 2).

- Pour la description d'une population, à ces sources d'incertitude s'ajoute le rôle du hasard (incertitude stochastique), dès lors que l'on intervient sur un échantillon (figure 2).
- Pour toute action prédictive, à ces diverses sources d'incertitude s'adjoint de l'incertitude méthodologique découlant du modèle utilisé pour la prévision.

On peut donc considérer qu'il existe un gradient d'amplitude des sources d'incertitude allant d'une simple revue bibliographique d'un concept, d'un objet ou d'un sujet jusqu'à une prévision de l'avenir en passant par une étude spécifique d'un tel type d'élément ou d'une population (figure2).

Figure 2

Exemples de représentation de l'incertitude globale d'importance croissante : individu et population (présent ou passé) et futur



III - L'INCERTITUDE EN ÉPIDÉMIOLOGIE ET EN ANALYSE DE RISQUE

L'incertitude épistémique est particulièrement importante lors d'émergence d'une maladie, non seulement quant à la nature de son agent pathogène (exemple : pour l'encéphalopathie spongiforme bovine, pendant les mois ayant suivi son apparition, personne ne connaissait la nature de l'agent pathogène), mais également pour ses principales caractéristiques : nature du réservoir (actuellement encore incertain pour le syndrome respiratoire à coronavirus du Moyen-Orient), distribution des durées de l'incubation, nature des matières virulentes, etc.

Même pour une maladie connue depuis longtemps, les connaissances peuvent être insuffisantes, à cause de la multiplicité et des changements au cours du temps des espèces animales impliquées (exemple de la tuberculose à *M. bovis* parmi les nombreuses espèces d'animaux sauvages réceptives), des différences constatées dans l'expression clinique ou épidémiologique en fonction des sérotypes d'une même espèce virale, ou en fonction de la zone géographique atteinte (différence de races d'une même espèce animale, de modes d'élevage, d'environnement, etc.), de l'évolution de la virulence de certains agents pathogènes biologiques au cours du temps (exemple : les nombreux sous-types de virus influenza), etc.

1. ÉPIDÉMIOLOGIE DESCRIPTIVE

En situation descriptive, les sources d'incertitude par rapport à la situation véritable sont diverses.

Comme pour la mesure de n'importe quelle caractéristique d'un objet, la qualité du test ou du protocole d'identification d'une maladie conditionne la qualité de la connaissance de la réalité. Elle dépend de la **sensibilité** et de la **spécificité** de cet « outil » de diagnostic ou de dépistage ainsi que de la personne ou du laboratoire le mettant en œuvre (incertitude méthodologique).

Lors de surveillance épidémiologique, notamment événementielle, le risque majeur est la **connaissance incomplète** des cas ou des foyers qui conduit à une sous-estimation, involontaire, de la situation réelle (incertitude épistémique).

Un autre facteur d'incertitude, lors d'enquête descriptive, est l'impossibilité de tester la population entière et, donc, la nécessité d'avoir à

extrapoler à la population entière les résultats obtenus sur un **échantillon**. Cette manière de procéder génère au moins deux incertitudes : celle liée à la représentativité de l'échantillon qui, si elle n'est pas parfaite, conduit à des biais pouvant entacher les résultats d'inexactitude, et celle liée à la taille de l'échantillon, conduisant à une fourchette à l'intérieur de laquelle le résultat a une probabilité élevée de se situer, cette imprécision constituant une autre source d'incertitude (incertitude stochastique).

Enfin, même avec un échantillon initial parfaitement représentatif de la population source et de grande taille (donc pouvant conduire à des résultats très précis), de nombreux autres facteurs comme une standardisation imparfaite de certains éléments du protocole (outils de mesure, ou d'analyse par exemple), peuvent être à l'origine de **biais** et entacher les résultats d'inexactitude et, donc, d'incertitude (incertitude méthodologique).

2. ÉPIDÉMIOLOGIE ANALYTIQUE

Pour une *maladie transmissible*, les différents facteurs et les mécanismes qui interviennent pour assurer le « succès » de la transmission de la maladie ne sont, le plus souvent, que partiellement connus et variables en fonction du temps, de l'espace et des populations exposées (incertitude épistémique).

Les sources de l'agent pathogène biologique sont multiples, leur intensité et leur durée de production de l'agent pathogène, variables. Certaines d'entre elles demeurent mal connues et conduisent à des interrogations comme, par exemple, le rôle possible du chien dans la transmission du virus Ebola à l'Homme en 2015 ou les modalités de contamination de l'Homme par le syndrome respiratoire à coronavirus du Moyen-Orient à partir d'animaux (incertitude épistémique).

Les modalités de la transmission peuvent être variées et dépendantes de divers facteurs environnementaux (notamment en cas de transmission vectorielle).

La réceptivité des sujets exposés n'est pas identique au sein d'une population vierge et la proportion des sujets réceptifs évolue en permanence au sein d'une population infectée en fonction d'introductions de sujets (naissances, etc.), de sorties (décès, etc.), du taux de prévalence, du

statut immunitaire (vaccinés, immunodéprimés...), etc. (incertitude stochastique).

Bref, la situation épidémiologique d'une maladie transmissible présente dans un troupeau et, à plus forte raison, dans une région, est en constante évolution sous la dépendance de nombreux facteurs intervenant dans la transmission.

Pour une *maladie non transmissible*, les études analytiques sont elles aussi exposées à de nombreux risques de biais, à détecter avant de parvenir à une conclusion. La traque attentive des facteurs de confusion peut conduire à la démonstration d'une association entre un facteur de risque supposé et une maladie, ce qui n'est qu'une étape vers l'inaccessible certitude d'une relation causale, grâce à l'analyse des critères de causalité de Hill [Desenclos et Thiébaud, 2012] (incertitude méthodologique).

3. ANALYSE DE RISQUE

Son objectif habituel est l'estimation de la probabilité de survenue d'un événement de santé et de ses conséquences. Elle utilise les connaissances plus ou moins fiables et partielles des deux domaines épidémiologiques précédents (incertitude épistémique), nécessite l'emploi d'une

méthode d'estimation (incertitude méthodologique) et se heurte aux effets du hasard (incertitude stochastique) qui amplifie l'incertitude des résultats de l'analyse.

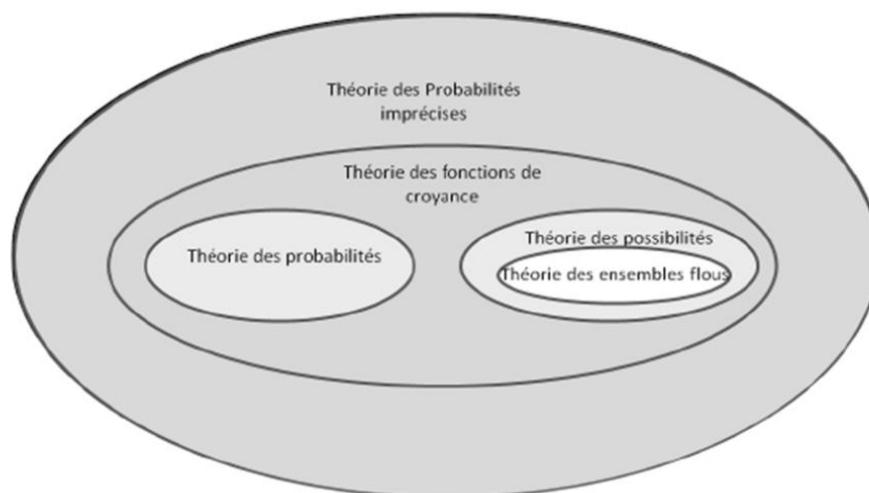
Dans le domaine de l'incertitude méthodologique, le Comité de la prévention et de la précaution a utilisé en 2010 diverses expressions pour décrire les différents types d'incertitude qui peuvent exister au sein de la chaîne de causalité :

- *L'incertitude d'identification*, en cas de doute sur l'influence d'un facteur qui n'a pas clairement été mise en évidence ;
- *L'incertitude de modélisation*, lors d'hésitation entre des représentations alternatives d'une relation causale ;
- *L'incertitude paramétrique*, lorsque l'on ne connaît pas précisément une valeur numérique entrant en jeu dans une relation [Comité de la prévention et de la précaution, 2010].

De nombreuses publications portent sur le traitement de l'incertitude en analyse de risque. Pour illustrer par un exemple, entre autres, la complexité que l'on y rencontre, la figure 3 indique les principales théories de modélisation de l'incertitude citées par Aguirre *et al.* [2015] pour les études de fiabilité.

Figure 3

Principales théories de modélisation de l'incertitude dans les études de fiabilité
[Aguirre *et al.*, 2015]



Une modalité particulière d'analyse de risque, pas seulement en santé animale, est celle qui aboutit à une hiérarchisation d'un ensemble de dangers aussi bien autochtones qu'exotiques. On y rencontre les mêmes sources d'incertitude, l'incertitude terminologique n'y jouant en général qu'un rôle très limité et, au contraire, l'incertitude méthodologique, un grand rôle.

Le tableau 4 présente les différents types d'incertitude rencontrés en épidémiologie et en analyse de risque. Il situe également les secteurs de

réduction possible de l'incertitude et les solutions pour y parvenir.

En fonction de la situation, l'incertitude peut correspondre à un seul des quatre types (par exemple, le résultat d'une enquête de prévalence ne comporte pas d'aspect prédictif et peut ne pas dépendre d'incertitude terminologique) ou résulter de la combinaison de deux ou plusieurs types (par exemple, une analyse de risque d'introduction dans un pays d'une maladie exotique mal connue, exprimée qualitativement avec des termes peu précis).

Tableau 4
Les quatre types d'incertitude en épidémiologie et analyse de risque

Type d'incertitude	Source(s) d'incertitude	Modalités d'expression	Réduction de l'incertitude
Épistémique	Insuffisance de connaissance	Mention du niveau de l'insuffisance de connaissance	Recherche bibliographique, observationnelle ou expérimentale, dires d'experts
Méthodologique	Qualité du protocole employé : limites de l'outil de mesure, travail sur échantillon, biais, appariement, facteurs de confusion, etc.	Expression du résultat avec un intervalle de confiance ou à l'aide de divers estimateurs statistiques	Amélioration de l'outil de mesure ou du système d'identification ; Augmentation de la connaissance de la population ; Augmentation de l'effectif de l'échantillon ; Utilisation de modèles adéquats ; Prise en compte des facteurs de confusion et des biais
Stochastique	Hasard, hétérogénéité d'une population et variabilité des sujets d'une population • Incertitude sur la probabilité de survenue • Incertitude sur l'identité du « sujet » touché	Expression du résultat avec un intervalle de confiance, une distribution de probabilités ou divers estimateurs statistiques	*
Terminologique	Vocabulaire imprécis ou pouvant conduire à une compréhension différente	Exprimer les qualificatifs par des intervalles ; Accompagner chaque expression qualitative de probabilité d'un intervalle numérique adapté au contexte [Hayes, 2011]	Préciser le plus possible ce que signifie chacun des termes utilisés ; Tenter d'établir une correspondance numérique des expressions qualitatives [Hayes, 2011]

*En partie non réductible

IV - L'EXPRESSION DE L'INCERTITUDE

Par définition, il est impossible de connaître de manière exacte l'importance de l'incertitude. Cependant, les scientifiques doivent faire l'effort, parfois très difficile, d'estimer et d'exprimer ou de représenter, de manière la plus compréhensible possible, l'incertitude qui entoure un résultat ou un avis fourni à un demandeur.

En épidémiologie et en analyse de risque, l'incertitude s'exprime en évoquant des niveaux de risque d'erreur, par excès (erreur de première espèce, alpha) ou par défaut (erreur de deuxième espèce, bêta), des intervalles de confiance ou des distributions de probabilités.

⇒ De manière *qualitative*, les caractères utilisés sont des mots exprimant par exemple le niveau estimé de probabilité de survenue d'un événement : « *minime* », « *très faible* », « *élevé* », etc. Dans ce cas, l'incertitude peut être exprimée par une fourchette entre des niveaux de probabilité estimés par des mots. Par exemple, dans l'échelle de valeurs de la méthode d'analyse qualitative de l'Anses en santé animale [Afssa, 2008] qui comporte une grille ordinale de dix niveaux (de 0 à 9), l'expression d'un résultat de probabilité de survenue qualifiée de « *extrêmement faible à très faible* » (soit 3 à 4 sur l'échelle logarithmique de 0 à 9) comporte une incertitude nettement plus faible que celle qualifiée de « *minime à faible* » (soit 2 à 5 sur cette échelle). Ces expressions qualitatives peuvent entraîner une incertitude terminologique lorsque différentes personnes ne comprennent pas ou n'interprètent pas de manière identique les mots utilisés comme « *faible* » ou « *élevé* » ou autre, exprimant une probabilité.

⇒ De manière *quantitative*, les résultats exprimés aussi bien en épidémiologie descriptive qu'en épidémiologie analytique comprennent une fourchette de nombres, par exemple pour la prévalence : 5 % ; IC 95 % [2 % ; 8 %] ou pour l'odds ratio ou le risque relatif : OR = 6 ; IC 95 % [4 ; 8].

Comme pour l'expression qualitative, l'amplitude de l'intervalle de confiance caractérise le niveau d'incertitude.

Lorsque la distribution des valeurs des variables utilisées dans une analyse de risque est bien connue, l'emploi des distributions de probabilités correspondantes et d'un logiciel adéquat permet d'exprimer le résultat sous forme d'une distribution de probabilités intégrant l'incertitude due à l'ensemble des variables prises en compte.

⇒ Une autre façon d'estimer et d'exprimer l'incertitude, notamment pour l'incertitude épistémique, repose sur l'emploi d'un indice d'incertitude correspondant à une échelle de niveaux. Dans un rapport sur le bien-être des bovins, l'Agence européenne de sécurité alimentaire (AESA) [EFSA, 2012] a défini pour l'incertitude épistémique une échelle simple à trois niveaux d'incertitude : faible, moyenne et élevée. Ces niveaux sont définis en fonction du nombre, de l'origine et de la fiabilité des données disponibles. Elle a utilisé également une matrice de croisement, d'une part, de l'incertitude de la probabilité de survenue de l'effet néfaste, d'autre part, de l'incertitude d'exposition (cf. tableau 5).

Tableau 5

Matrice de combinaison des niveaux d'incertitude de la probabilité de survenue de l'effet néfaste et d'incertitude de l'exposition

[EFSA, 2012]

		Incertainude d'exposition		
		Elevée	Moyenne	Faible
Incertainude de survenue de l'effet néfaste	Elevée	Elevée	Elevée	Elevée
	Moyenne	Elevée	Moyenne	Moyenne
	Faible	Elevée	Moyenne	Faible

Cette approche intéressante présente néanmoins des limites : elle ne prend en compte que l'incertitude épistémique ; on peut s'interroger pour savoir si une échelle à trois niveaux

d'incertitude suffit pour représenter la diversité des situations. Par ailleurs, la manière dont les décideurs peuvent utiliser le niveau d'incertitude associé à un résultat n'est pas précisée.

V - CONCLUSION

En épidémiologie et en analyse de risque, l'incertitude est fréquente et souvent d'étiologie complexe.

Il importe d'en identifier l'origine, d'en estimer le niveau et de fournir les résultats d'une étude ou d'un avis en exprimant le plus clairement possible l'incertitude estimée.

Une réduction partielle de l'incertitude globale est possible, fondée sur une amélioration des outils et de la connaissance, l'utilisation de modèles adéquats et l'emploi de termes évitant l'imprécision. L'incertitude stochastique est, néanmoins, en partie irréductible. Toutefois, il n'est

pas indispensable d'atteindre un niveau nul d'incertitude pour prendre une décision. Au plan pragmatique, des mesures de prévention ou de lutte sont le plus souvent décidées en situation d'incertitude. Ainsi par exemple, les décideurs sanitaires de Grande-Bretagne n'ont pas attendu d'avoir la preuve de la relation causale entre les farines de viande et l'encéphalopathie spongiforme bovine pour interdire l'emploi des farines animales dans l'alimentation des bovins. Comme l'indiquent Ducrot et Bres [2013], « on est bien obligé d'accepter le probable comme certain pour prendre des décisions ».

BIBLIOGRAPHIE

Afssa. - Une méthode qualitative d'estimation du risque en santé animale. Afssa éd., 2008, 68 p.

Aguirre F., Sallak M., Schön W. - Incertitudes aléatoires et épistémiques, comment les distinguer et les manipuler dans les études de fiabilité ? <https://hal.archives-ouvertes.fr/hal-00823114/> (consulté le 30.01.2015).

Busschaert P., Geeraerd A.H., Uyttendaele M. *et al.* - Sensitivity Analysis of a Two-Dimensional Quantitative Microbiological Risk Assessment: Keeping Variability and Uncertainty Separated. *Risk Analysis*, 2011, **31**(8), 1295-1307.

Christley R.M., Mort M., Wynne B. *et al.* - « Wrong, but Useful »: Negotiating Uncertainty in Infectious Disease Modelling. *PLoS ONE* 8(10): e76277.doi:10.1371/journal.pone.0076277.

Comité de la prévention et de la précaution. La décision publique face à l'incertitude. Clarifier les règles, améliorer les outils. Rapport, mars 2010, Paris, 50 p.

Desenclos J.C., Thiébaud R. - Causalité, incertitude et jugement. *In* Épidémiologie de terrain.

Méthodes et applications, Dabis F. et Desenclos J.C., John Libbey Eurotext, 2012, 758 p.

Ducrot C., Bres D. - Regard philosophique sur la causalité en épidémiologie-approche interdisciplinaire. *INRA Prod. Anim.*, 2013, **26**(4), 375-382.

EFSA - Panel on Animal Health and Welfare (AHAW); Scientific Opinion on the welfare of cattle kept for beef production and the welfare in intensive calf farming systems. *EFSA Journal*, 2012, 10(5), 2669. 166 p. doi:10.2903/j.efsa.2012.2669. Available online: www.efsa.europa.eu/efsajournal.

Fish R., Austin Z., Christley R. *et al.* - Uncertainties in the governance of animal disease: an interdisciplinary framework for analysis. *Phil. Trans. R. Soc. B*, 2011, **366**, 2023-2034.

Hayes K.R. - Uncertainty and uncertainty analysis methods. Technical report EP 102467, CSIRO Mathematics, Informatics and Statistics, Hobart, Australia, 2011, 132 p.

Pouillot R., Delignette-Muller M.L. - Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. *International Journal of Food Microbiology*, 2010, **142**, 330-340.

Rai S.N., Krewski D. - Uncertainty and Variability Analysis in Multiplicative Risk Models. *Risk Analysis*, 1998, **18**(1), 37-45.

Vose D. - Risk analysis, a quantitative guide. West Sussex, UK: John Wiley and sons, Ltd., 3rd edition, 2008.

Wynne B. - Uncertainty and environmental learning. Reconciling science and policy in the preventive paradigm. *Global Environ. Change*, 1992, **2**, 111-127 (doi: 10.1016/0959-3780(92)90017-2).

Zheng J., Frey H.C. - Quantitative Analysis of Variability and Uncertainty with Known Measurement Error: Methodology and Case Study. *Risk Analysis*, 2005, **25**(3), 663-675.



Remerciements

Aux personnes ayant fait des commentaires et des suggestions après lecture du projet d'article : Vincent Auvigne, Didier Calavas, Olivier Cerf, Christian Ducrot, Eric Etter, Nadine Steinfeld