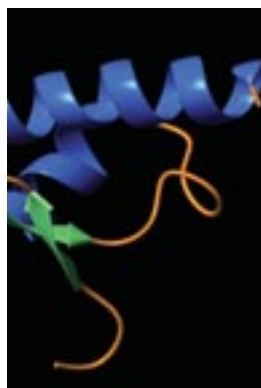


> Les modélisations de l'épidémie du variant de la la maladie de Creutzfeldt-Jakob (vMCJ) dû à l'agent de l'encéphalopathie spongiforme bovine (ESB) font toutes appel à un « rétro-calcul » qui relie par une équation mathématique simple: (1) le nombre de personnes infectées; (2) la distribution des dates d'infection; et (3) la distribution de la durée d'incubation. Les modèles se distinguent par les hypothèses et les techniques utilisées pour définir ces trois paramètres inconnus. En ce qui concerne la durée d'incubation, certains modèles font appel à des scénarios fixant arbitrairement sa valeur moyenne et sa variabilité; d'autres, plus récents, cherchent à les estimer en s'appuyant sur l'incidence du vMCJ entre 1995 et 2001. Malgré de nombreuses différences méthodologiques, les résultats des modélisations sont très cohérents. Pour des durées d'incubation de 15 à 20 ans, valeur obtenue par tous les modèles ayant cherché à estimer ce paramètre à partir de l'incidence du vMCJ, les cinq modèles prédisent une épidémie de 500 à 2000 cas au Royaume-Uni. <

Épidémie de nouveau variant de la maladie de Creutzfeldt-Jakob au Royaume-Uni : anatomie et physiologie des modèles prédictifs

Annick Alpérovitch, Jérôme Huillard d'Aignaux



A. Alpérovitch: Inserm U.360, Recherches épidémiologiques en neurologie et psychopathologie, Hôpital de la Salpêtrière, 47, boulevard de l'Hôpital, 75651 Paris Cedex 13, France. alperovi@chups.jussieu.fr.
J. Huillard d'Aignaux: London School of Hygiene and Tropical Medicine, Department of Infectious Diseases, Londres, Royaume-Uni.

assez importants. Ces différences ne nous étonnent plus, même si nous ne disposons pas de tous les éléments permettant de les comprendre et les interpréter, et elles ne remettent pas en question l'intérêt de ces travaux. La prédiction de l'épidémie de nouveau variant de la maladie de Creutzfeldt-Jakob (vMCJ) n'est que l'un des exemples, récent et très médiatisé, d'apparentes divergences entre modèles. Cet article a pour objectif de montrer que ces divergences ne sont pas majeures et sont assez faciles à expliquer. Pour développer un modèle, il faut partir des données et revenir aux données. Partir des données, qui aident à définir les hypothèses et paramètres du modèle. Revenir aux données, pour s'assurer de leur compatibilité avec les résultats de la modélisation. Nous commencerons donc par un bref résumé des données épidémiologiques sur le vMCJ.

Le nouveau variant de la maladie de Creutzfeldt-Jakob : épidémiologie descriptive

Au début de l'année 1996, Will *et al.* décrivaient dans le *Lancet* [1] une nouvelle forme sporadique de la maladie de Creutzfeldt-Jakob (MCJ) apparue en 1994 au

Les modèles mathématiques sont utilisés dans de nombreuses disciplines scientifiques, le plus souvent pour décrire et aider à comprendre un phénomène observé, mais aussi pour prédire son évolution. Les développements de l'informatique et des logiciels de statistique ont fait des modèles descriptifs un banal outil de l'analyse des données biomédicales. La majorité des utilisateurs de ces logiciels n'a pas conscience qu'une analyse statistique implique souvent une certaine part de modélisation.

Les modèles généraux très simples sont insuffisants lorsqu'il s'agit d'analyser des phénomènes économiques, génétiques ou épidémiologiques complexes. Il faut alors développer des modèles *ad hoc*. La modélisation d'un problème multifactoriel et complexe n'est pas univoque, et la réponse à la question posée varie en fonction du modèle adopté. En économie, domaine où les résultats des modélisations sont commentés par de nombreux médias, il n'est pas rare que les écarts entre les prévisions fournies par différents organismes soient

Royaume-Uni. Cette première publication faisait état de 11 cas de vMCJ, qui se distinguaient des MCJ sporadiques classiques par l'âge, et la localisation et le type des lésions neuropathologiques. Le premier cas français était diagnostiqué en 1996 [2]. D'emblée, sur des arguments épidémiologiques, Will *et al.* faisaient l'hypothèse d'un lien entre le vMCJ et l'encéphalopathie spongiforme bovine (ESB). Les travaux expérimentaux réalisés dans les mois suivants ont renforcé cette hypothèse [3-5], qui fait aujourd'hui l'objet d'un large consensus. L'hypothèse d'une contamination d'origine alimentaire est considérée comme la plus probable, mais les données sur l'association entre consommation de produits bovins et risque de vMCJ ne sont pas encore très nombreuses [6-8], et n'apportent que des arguments indirects en faveur de l'origine alimentaire de l'infection. Fin décembre 2001, le nombre de décès par vMCJ était de 104 au Royaume-Uni, 3 en France, 1 en République d'Irlande. Aucun cas n'avait encore été diagnostiqué en dehors de ces trois pays. Quelques cas de vMCJ probable étaient en cours d'évolution (une dizaine environ au Royaume-Uni et 2 en France), le délai entre les premiers symptômes et le décès dépassant rarement 18 mois. La *Figure 1* représente le nombre de décès par vMCJ au Royaume-Uni, par année et cumulé sur la période 1995-2001. Une augmentation de l'incidence d'environ 23 % par an a été observée au cours de la période 1994-2000 [9]. Les patients atteints de vMCJ ont en moyenne 29

ans; les trois-quarts ont entre 20 et 40 ans, 5 % seulement ont plus de 50 ans. Pour mémoire, l'âge moyen des cas sporadiques est 65 ans. Comme le montre la *Figure 2*, l'âge moyen des cas n'a pas significativement varié depuis le début de l'épidémie, et en particulier il n'a pas augmenté: cette stabilité est une autre caractéristique importante de la distribution des âges des vMCJ.

À ce jour, tous les cas de vMCJ sont homozygotes méthionine-méthionine sur le codon 129 du gène de la protéine prion (PRNP). Le gène *PRNP*, situé sur le chromosome 20, présente sur le codon 129 un polymorphisme qui est associé au risque de MCJ [10]. Quel que soit le type de MCJ, les individus homozygotes (méthionine-méthionine ou valine-valine) sont plus à risque de développer la maladie que les hétérozygotes méthionine-valine.

Une étude sur les cas iatrogènes imputables à un traitement par hormone de croissance extractive a montré que ce polymorphisme influence aussi la durée d'incubation de la maladie, plus longue chez les hétérozygotes [11]. En ce qui concerne le vMCJ, deux hypothèses peuvent être envisagées: soit les individus méthionine-méthionine, qui représentent 40 % de la population générale, sont les seuls qui risquent de développer la maladie, soit la durée d'incubation est plus longue chez les individus d'un autre génotype. Des travaux expérimentaux suggèrent que risque et durée d'incubation sont négativement corrélés. Selon cette hypothèse, les cas homozygotes valine ou hétérozygotes s'ils surviennent plus tardivement, seraient probablement en plus petit nombre que les cas méthionine-méthionine. Les données épidémiologiques sur le vMCJ qui viennent d'être très brièvement résumées (incidence, âge, génotype) sont prises en compte dans toutes les modélisations de l'épidémie.

Modélisation de l'épidémie de vMCJ au Royaume-Uni: méthodologie générale

La population britannique a été massivement exposée à l'agent de l'ESB. Environ 1 000 000 de bovins infectés sont entrés dans la chaîne alimentaire au Royaume-Uni, dont près de la moitié avant la fin de 1989, date des mesures d'interdiction d'utilisation de certains tissus dits « à risque spécifié » dans l'alimentation humaine. Il est possible, compte tenu des procédés de fabrication de certains produits à base de viande, que des millions de personnes aient été en contact avec cet agent infectieux. La crainte d'une épidémie de vMCJ importante, voire très importante, repose sur ces faits. La première prédiction de l'épidémie de vMCJ a été publiée au début

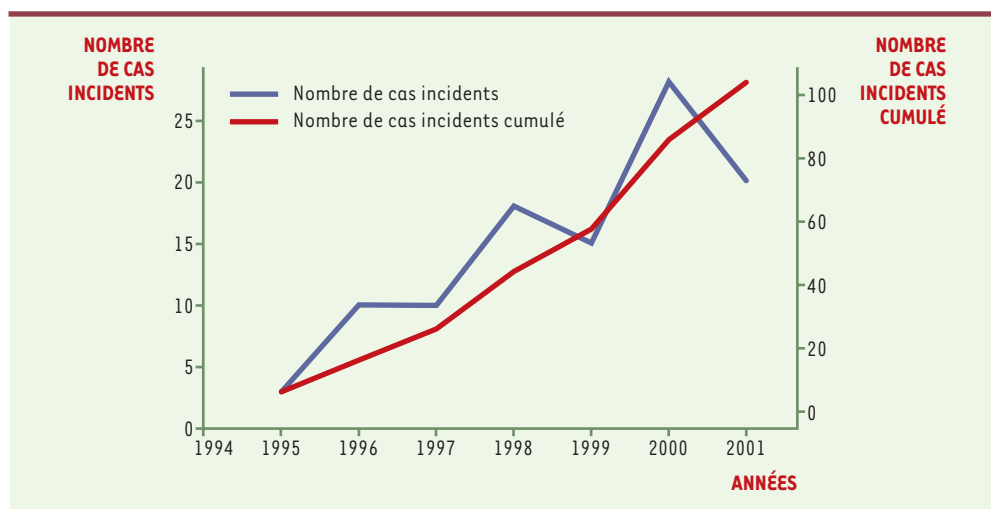


Figure 1. Incidence du vMCJ au Royaume-Uni, par année et cumulée, entre 1995 et 2001.

de 1997, alors que le nombre de décès par vMCJ était seulement de 14 [12]. À ce jour, cinq modèles de prédiction de l'épidémie de vMCJ ont été développés [12-16], dont trois au cours du dernier trimestre 2001. Dans cet article, nous nous limiterons à une analyse des quatre premières modélisations [12-15]. La dernière [16] se distingue essentiellement par la prise en compte d'une possible transmission de l'ESB au mouton. Le modèle indique qu'une éventuelle épidémie d'ESB ovine n'entraînerait pas d'augmentation considérable du nombre de cas de vMCJ à venir. Pour le reste, cette modélisation confirme, globalement, le travail antérieur de la même équipe [13], avec une révision à la baisse de ses prédictions les plus pessimistes.

On peut ajouter à ces cinq articles, deux modélisations [17, 18] ayant pour objectif d'estimer la prévalence de l'infection dans la population, prévalence dont dépend évidemment le nombre de cas de vMCJ à venir. On peut détecter de la protéine prion anormale dans l'appendice et les amygdales de patients atteints de vMCJ, alors que ce test est négatif dans les autres formes de MCJ [19]. Ce résultat a conduit à envisager des études en population générale dont le but serait la recherche de protéine prion anormale dans des pièces opératoires d'amygdalectomie ou d'appendicectomie. Une étude pilote portant sur des pièces opératoires provenant d'interventions réalisées entre 1995 et 1998 sur des personnes âgées de 10 à 30 ans a été complètement négative [20]. Mais les modèles [17, 18] montrent que pour estimer avec précision la prévalence de l'infection dans la population, il faudrait examiner des dizaines, voire des centaines de milliers de pièces opératoires. On peut déjà noter au passage que dans le travail de Cooper *et al.* [17], la meilleure adéquation entre modèle et données est obtenue pour une durée d'incubation du vMCJ de l'ordre de 15 ans, valeur très proche des estimations proposées récemment par Huillard d'Aignaux *et al.* [14] et Valleron *et al.* [15].

Tous les modèles de prédiction de l'épidémie de vMCJ font appel au rétro-calcul, outil majeur de l'épidémiologie théorique des maladies infectieuses [21], utilisés notamment pour prédire l'épidémie de SIDA au début des années 1980. La méthode de rétro-calcul est fondée sur le fait que le nombre de cas d'une maladie infectieuse à un instant donné dépend de trois paramètres: le nombre de personnes infectées, la distribution des dates d'infection, la distribution du délai entre l'infection et la maladie (ou durée d'incubation) (Figure 3). La relation entre ces paramètres peut être décrite par une équation mathématique simple. La difficulté est que tous ces paramètres sont inconnus, et ce sont les hypothèses et techniques utilisées pour donner formes et valeurs à ces paramètres qui vont distinguer les différents modèles.

Modélisation de la durée d'incubation

La distribution des intervalles entre la date de l'infection et la date des premiers signes suit une loi mathématique dont la forme et les paramètres (moyenne, variabilité) sont inconnus dans le vMCJ. Les modèles font donc appel à des lois de probabilité connues pour décrire la durée d'incubation de différentes maladies infectieuses. La plus classique est la loi lognormale, mais les modèles utilisent en général plusieurs formes mathématiques (lois gamma, Weibull, F généralisée,...)*. Outre l'équation mathématique, il faut déterminer les valeurs numériques des paramètres de la distribution: moyenne, médiane, et percentiles. Ces valeurs numériques sont fixées arbitrairement dans les modèles « scénarios », mais peuvent être estimées à partir des données d'incidence, lorsque l'effectif est suffisant. Avec 14 cas de vMCJ fin 1996, Cousens *et al.* [12] ne pouvait utiliser qu'une approche « scénarios ». Huillard d'Aignaux *et al.* [14] et Valleron *et al.* [15], s'appuyant respectivement sur 82 et 97 cas incidents de vMCJ, ont pu estimer les paramètres de la distribution de la durée d'incubation.

Les lois utilisées pour décrire la durée d'incubation du vMCJ sont toutes unimodales, c'est-à-dire n'ont qu'un seul pic (ou mode). Les modèles font donc tous l'hypothèse que la durée d'incubation suit la même loi quelles que soient les caractéristiques des individus, en particulier quel que soit leur âge. La rareté des cas de plus de 50 ans est une caractéristique épidémiologique majeure du vMCJ. Les explications possibles de cette observation sont au nombre de trois: (1) les personnes les plus

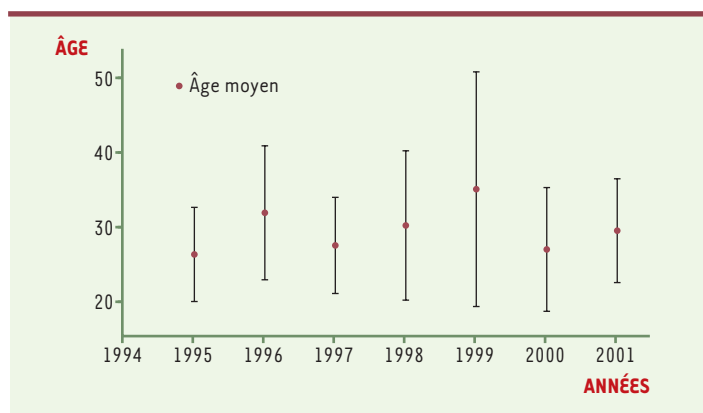


Figure 2. Âge moyen des cas de vMCJ entre 1995 et 2001.

* Ces distributions sont définies par des équations mathématiques différentes. Des travaux ont montré qu'elles étaient bien adaptées pour des maladies infectieuses à durées d'incubation variables. A moyenne et variance constantes, ces lois diffèrent principalement aux extrêmes de la distribution (durées d'incubation les plus courtes ou les plus longues). Le choix de l'une ou l'autre des distributions va donc influencer l'allure de la courbe au début de l'épidémie ou lors de sa phase décroissante.

jeunes ont été les plus exposées à l'agent infectieux, en raison de leurs habitudes alimentaires; (2) les personnes les plus jeunes sont les plus à risque de développer la maladie pour des raisons biologiques encore inconnues; (3) la durée d'incubation dépend de l'âge, et est plus longue quand l'infection survient à un âge plus tardif. Des travaux expérimentaux et l'analyse des données sur le kuru [22] suggèrent que la durée d'incubation peut varier en fonction de l'âge. Mais les variations observées ou estimées sont faibles et peuvent difficilement expliquer la distribution d'âge des cas de vMCJ. Aussi, tous les modèles font l'hypothèse que la durée d'incubation est indépendante de l'âge, ce qui revient à rejeter la troisième explication pour ne retenir que les deux premières. S'appuyant sur 97 cas de vMCJ, Valleron *et al* [15] écartent aussi l'hypothèse de la plus forte exposition des adolescents et adultes jeunes pour ne retenir que celle d'un facteur biologique lié à l'âge.

Outre l'âge, le codon 129 du gène *PRNP* pourrait aussi être responsable d'une bi- ou tri-modalité de la distribution des durées d'incubation. On peut même envisager une loi plurimodale si, comme le suggèrent les travaux de l'équipe de Collinge [23], d'autres gènes que le gène *PRNP* jouent un rôle dans la maladie. Les données épidémiologiques actuelles ne permettent ni d'écarter cette hypothèse ni d'inclure cette possibilité dans un modèle

de durée d'incubation bi- ou pluri-modale. La solution adoptée dans toutes les modélisations est de ne considérer qu'une seule et même loi de probabilité pour toutes les personnes infectées.

En conclusion, pour ce qui est de la durée d'incubation, la différence majeure entre les modèles est que certains utilisent des « scénarios » [11, 12, 16], approche permettant de faire varier arbitrairement les paramètres (moyenne, variabilité) de la distribution, et d'autres [14, 15] estiment ces paramètres à partir des données observées. Quelle que soit l'approche, un intervalle de prédiction va être obtenu à partir d'un grand nombre de simulations. L'interprétation de cet intervalle dépend de la méthode. Dans les modèles « estimation », on cherche à déterminer de vrais intervalles de confiance. Dans les modèles « scénarios », on obtient un intervalle de valeurs compatibles avec

les données d'incidence observées au moment de la simulation, sans qu'il soit possible de définir un niveau de confiance, ou d'associer une quelconque probabilité à la prédiction.

Modélisation de la distribution des dates d'infection

Second élément nécessaire au rétro-calcul: la distribution des dates d'infection. Tous les modèles font l'hypothèse que, pour l'ensemble de la population, le risque d'infection à l'instant t est proportionnel au niveau d'exposition à l'agent de l'ESB, qui lui-même dépend principalement de deux paramètres: (1) le nombre d'animaux infectés par l'agent de l'ESB entrant dans la chaîne alimentaire à l'instant t ; et (2) l'efficacité des mesures de réduction du risque prises au Royaume-Uni à la fin de 1989. Le nombre d'animaux infectés entrés dans la chaîne alimentaire à un instant donné n'est pas connu, et les modèles considèrent habituellement que ce nombre est directement proportionnel au nombre de cas d'ESB diagnostiqués à la même période, en introduisant en général un facteur de sous-estimation au début des années 1980. Certains modèles [13] tiennent compte des variations de la force de l'infection dans le cheptel bovin en fonction du temps (très liée évidemment à l'incidence des cas d'ESB), et de la distribution d'âge des animaux abattus. Les travaux expérimentaux indiquent que la quasi-totalité de l'infectivité est localisée dans les tissus dits « à risque spécifié » interdits dans l'alimentation humaine au Royaume-Uni depuis la fin de 1989. L'efficacité de cette mesure dépend aussi évidemment de la rigueur de son application. Cousens *et al.* [12] considèrent deux scénarios d'efficacité des mesures de

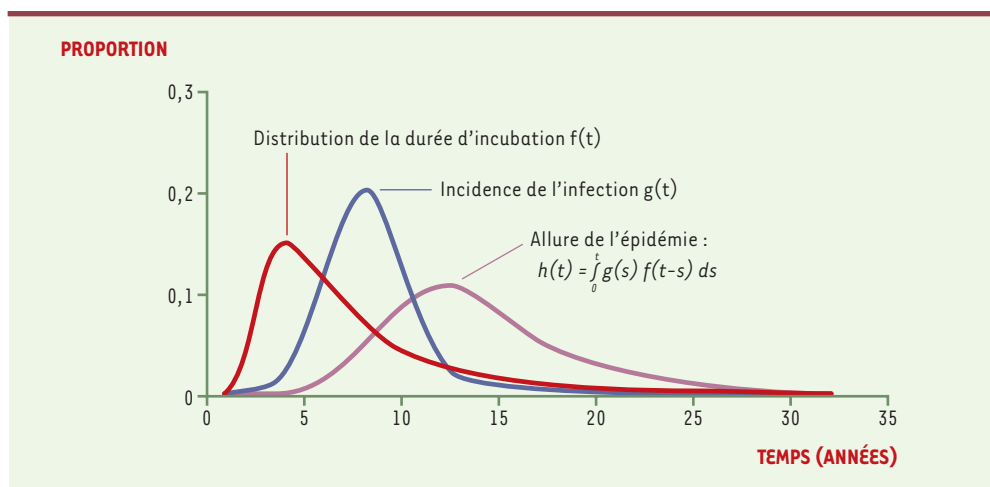


Figure 3. Schéma du modèle de rétro-calcul. Les modèles utilisent différentes distributions $g(t)$ s'appuyant toutes sur l'incidence de l'ESB au Royaume-Uni depuis le début des années 1980. Ils simulent aussi différentes distributions $f(t)$ dont les paramètres (moyenne, variabilité) peuvent être soit fixés arbitrairement, soit estimés si les données le permettent. A chaque $f(t)$ et $g(t)$ correspond une courbe épidémique $h(t)$.

réduction du risque après 1989 : 90 % et 100 %. Huillard d'Aignaux *et al.* [14] font varier cette efficacité entre 0 et 90 %. Valleron *et al.* [15] ne prennent pas en considération la probabilité d'une exposition après 1989 (ce qui revient à supposer 100 % d'efficacité aux mesures prises fin 1989). Ghani *et al.* [13] font varier ce paramètre sans préciser dans son article l'intervalle de variation. On peut schématiquement considérer que si les mesures prises fin 1989 réduisent le risque de plus de 80 %, le pic de la distribution du risque d'infection est en 1989, pour des réductions de risque inférieures à 80 %, le pic est en 1992. La *Figure 4* donne quelques exemples des distributions du risque d'infection en fonction du temps utilisées dans ces modélisations.

En conclusion, tous les modèles considèrent, à quelques nuances près, que la force d'infection à l'instant t est proportionnelle à l'incidence de l'ESB au même instant. Les modèles varient sur le facteur de réduction à appliquer à partir de 1990. Certains ne retiennent que des réductions de risque très importantes, supérieures ou égales à 90 % [12, 15]. D'autres examinent aussi des hypothèses plus pessimistes, allant jusqu'à la totale inefficacité des mesures de réduction du risque [14].

Nombre de personnes infectées

Du point de vue épidémiologique, il faut distinguer plusieurs étapes dans le processus aboutissant à la maladie. La première est le contact avec l'agent infectieux. La probabilité qu'un individu soit en contact avec l'agent infectieux dépend probablement de différents facteurs environnementaux encore mal connus (niveau socio-économique, habitudes alimentaires, etc.), mais elle est indépendante de caractéristiques biologiques individuelles, en particulier génétiques. Une autre étape correspond à l'infection, c'est-à-dire à la diffusion de

l'agent dans l'organisme et sa réplication, terme classique mais qui n'est peut-être pas tout à fait adapté aux infections à prions. La probabilité d'infection dépend de la dose ingérée et aussi, vraisemblablement, de facteurs biologiques individuels liés à l'âge et de facteurs génétiques. Ces facteurs peuvent aussi influencer la dynamique du processus d'infection, et donc la durée d'incubation, intervalle entre le début de l'infection et l'apparition des symptômes cliniques, dernière étape du processus. On peut concevoir le rôle de ces facteurs, en particulier du codon 129 du gène *PRNP*, d'une manière déterministe: les individus homozygotes valine ou hétérozygotes ne peuvent pas être infectés par l'agent de l'ESB. On peut aussi envisager une relation probabiliste: à dose égale, la probabilité d'être infecté et la dynamique de l'infection dépendent du génotype. Tous les cas de vMCJ observés à ce jour étant homozygotes méthionine, la solution adoptée dans toutes les modélisations est de considérer que seuls les individus méthionine-méthionine (soit 40 % de la population exposée) sont à risque d'infection, et donc de vMCJ. Là encore, les données existantes ne permettent pas de construire un modèle où d'autres génotypes pourraient être infectés. Les données existantes sont cependant suffisantes pour affirmer que les personnes homozygotes valine ou hétérozygotes ont un risque beaucoup plus faible de développer la maladie ou/et une durée d'incubation beaucoup plus longue que les homozygotes méthionine (*voir plus haut*).

Pour estimer le nombre de personnes infectées, les modèles se basent aussi sur la distribution d'âge des cas de vMCJ observés. Sous l'hypothèse que la durée d'incubation ne dépend pas de l'âge (*voir plus haut*), un modèle simple permet de calculer soit la distribution d'âge au moment de l'infection, soit les années de naissance des personnes exposées. Cette distribution reflète la relation entre le risque d'infection et l'âge. Valleron [15] suppose que le

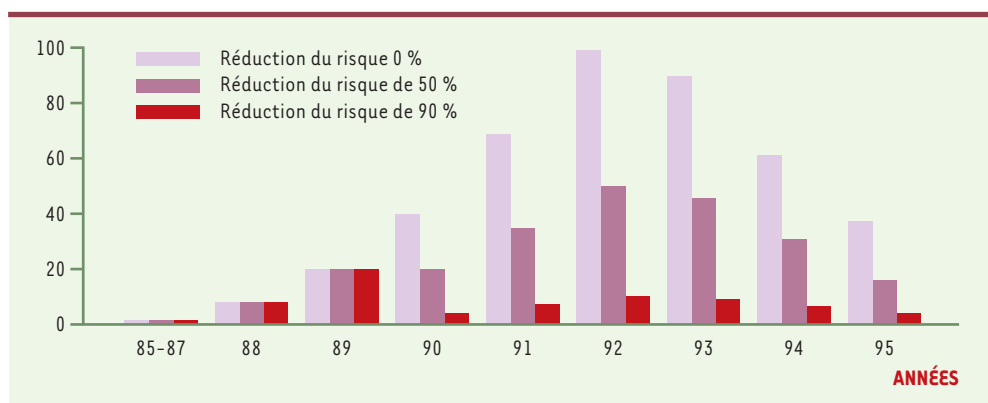


Figure 4. Distribution du risque d'infection entre 1985 et 1995 pour trois hypothèses d'efficacité des mesures de réduction du risque prises fin 1989. Les valeurs de risque sont présentées en affectant arbitrairement la valeur 100 au pic de la distribution correspondant à une absence totale d'efficacité des mesures de réduction de risque.

risque d'infection est constant entre 6 mois et 15 ans, puis décroît exponentiellement avec l'âge. Ghani [13] et Huillard d'Aignaux [14] font l'hypothèse que le risque d'infection varie avec la date de naissance: maximum pour les individus nés entre 1970 et 1975, il se calcule pour les autres cohortes d'âge en se fondant sur la distribution des dates de naissance

des cas de vMCJ observés. On peut facilement passer d'un modèle à l'autre (c'est-à-dire du modèle âge au modèle année de naissance) en faisant intervenir la distribution des dates d'infection. On réalise alors que ces hypothèses se recourent largement: dans un modèle comme dans l'autre, la très grande majorité des personnes infectées sont les adolescents de la fin des années 1980.

Cependant, aucun modèle ne permet de dire si la relation supposée entre âge et risque d'infection est liée à l'exposition ou à une susceptibilité plus élevée à niveau d'exposition constant. L'interprétation de cette relation reste encore assez arbitraire. Les études épidémiologiques en cours visant à estimer le niveau d'exposition en fonction de l'âge aideront à clarifier la nature de la relation entre âge et risque d'infection.

Une synthèse des différentes modélisations

Ce ne sont pas tant les différences entre modèles qui rendent difficiles la synthèse des prédictions, mais plutôt l'absence d'homogénéité dans la description des résultats obtenus. Cette hétérogénéité est en partie due à de réelles, mais subtiles, différences dans les objectifs de la modélisation, reflétées par le titre de l'article qui contient le mot « prédiction » [12, 13], ou « prédictibilité » [14] ou « estimation » [15]. Mais elle tient surtout à l'énorme quantité de résultats produite par les modélisations. Résumer cette masse d'informations en deux ou trois tableaux ou/et figures n'est pas simple. Les critères de sélection des résultats et leur mode de présentation vont varier d'une équipe à l'autre. De cette hétérogénéité se dégagent néanmoins quelques invariants importants. D'abord, quel que soit le modèle, qu'il soit du type « scénario » ou du type « estimation », le nombre de cas prédit est d'autant plus grand que la durée d'incubation est plus longue. À autres hypothèses constantes, le modèle de Cousens *et al.* [12] prédit 1,595 cas pour une durée moyenne d'incubation de 15 ans, et 80 000 cas si la durée d'incubation moyenne est de 25 ans. Les auteurs de modèles « scénarios » soulignent que la durée moyenne d'incubation est le paramètre majeur. Aucune des hypothèses et aucun des paramètres qui ont été discutés dans le paragraphe précédent n'a le même impact sur la prédiction. Ce constat suggère que pour comparer valablement les différentes prédictions, on peut essayer de se placer à durée d'incubation égale, que cette durée d'incubation soit estimée à partir des données [14, 15] ou fixée par « scénario » [12, 13]. Les deux modèles fournissant une estimation de la durée d'incubation du vMCJ, aboutissent à des valeurs moyennes très proches: entre 10 et 20 ans, selon

les hypothèses, pour Huillard d'Aignaux *et al.* [14], et 17 ans pour Valleron *et al.* [15]. Le travail de Cooper [17] cité précédemment, est en accord avec cette estimation. Pour des durées d'incubation comparables, disons inférieures à 20 ans, les scénarios étudiés par Cousens *et al.* [12] fin 1996 prédisaient un nombre total de vMCJ inférieur à 2 000, et ceux de Ghani [13] en 1999 moins de 1 000. Au total, pour des durées d'incubation inférieures à 20 ans, les cinq modèles prédisent une épidémie de vMCJ allant de 500 à 2 000 cas. Il y a donc une grande cohérence entre ces modèles.

Modélisation prédictive : une aide à la décision en santé publique ?

Les maladies à prions ont été le sujet de quelques tempêtes médiatiques, dont celle de l'été 2000 a fait suite à la publication de la modélisation de Ghani *et al.* [13], sous la forme d'une très brève communication dans la revue *Nature*. Ghani utilisait des scénarios couvrant un très large champ d'hypothèses, notamment sur la durée moyenne d'incubation de la maladie qui pouvait varier dans ce modèle de moins de 20 ans à plus de 60 ans. Sans attendre les estimations récentes qui situent cette durée autour de 15 ans [14, 15, 17], on pouvait considérer que les scénarios envisageant une durée moyenne d'incubation de plus de 60 ans n'étaient pas les plus vraisemblables. Les médias ont rapporté les résultats de Ghani *et al.* [13] en présentant essentiellement la limite supérieure de l'intervalle correspondant à une durée d'incubation moyenne de plus de 60 ans, soit le pire (limite supérieure de la prédiction) du scénario le pire (plus de 60 ans de durée d'incubation), conduisant à une épidémie de 136 000 cas de vMCJ au Royaume-Uni. Certains insistaient aussi sur l'imprécision de la modélisation, en ne retenant que les valeurs extrêmes: la limite inférieure du scénario le plus optimiste (60 cas de vMCJ au total, correspondant à une durée moyenne d'incubation inférieure à 20 ans) et la limite supérieure du scénario le plus pessimiste (136 000 cas). D'où une perplexité compréhensible sur l'intérêt des modèles de prédiction.

La modélisation d'un problème aussi complexe que l'épidémie de vMCJ mérite mieux qu'une critique sommaire. La prédiction de Cousens *et al.* [12], quelques mois après la description des premiers cas de vMCJ, cadrait déjà assez bien la taille de l'épidémie, en éloignant le scénario catastrophe de centaines de milliers de cas envisagé, sans aucune base sérieuse, quelques mois auparavant. Sur les 64 scénarios étudiés, couvrant des durées moyennes d'incubation allant de 15 à 25 ans, 9 prédisaient une forte épidémie (entre 5 000 et 80 000

cas), tous les autres scénarios étant plus optimistes. Pour chaque scénario, les auteurs indiquaient non seulement le nombre total de cas à venir, mais aussi l'incidence attendue de la maladie en 1996, 1997, 1998. En confrontant ces toutes premières prédictions à l'incidence réellement observée, on voit aujourd'hui que les scénarios pessimistes (5000 cas ou plus) impliquaient pour ces années des incidences beaucoup plus élevées que les nombres observés. Par exemple, selon le scénario « 80 000 vMCJ », on aurait dû avoir 32 cas incidents en 1996, 79 cas en 1997, 166 cas en 1998, alors que seulement 38 cas sont survenus au cours de la période 1996-1998. Les scénarios de Cousens *et al.* compatibles avec l'incidence observée du vMCJ entre 1996 et 1998 prédisaient, pour des durées d'incubation entre 15 et 20 ans, un nombre total de cas aux alentours de 1000 à 2000, valeurs très proches des récentes estimations de Huillard d'Aignaux *et al.* [14] et Valleron *et al.* [15] s'appuyant sur près de 7 ans de surveillance épidémiologique du vMCJ. Seule donc la modélisation de Ghani *et al.* [13] semble sortir de l'épuration. Encore que, pour le pire des scénarios de cette modélisation (durée d'incubation moyenne supérieure à 60 ans), le nombre de cas prédit se situe dans un intervalle allant de 1300 à 136 000 cas, dont la limite inférieure recoupe donc les estimations précédentes.

En conclusion, quelques connaissances sur les techniques de modélisation aident à interpréter les résultats fournis par un modèle, et discuter cohérence et incohérence entre modèles. Sur un problème de santé publique aussi préoccupant que le vMCJ, il est important que les auteurs (et les revues) fournissent à leurs lecteurs l'ensemble des informations nécessaires à une interprétation correcte de leurs résultats. Les modélisations qui se sont succédées depuis 1996 permettent aujourd'hui aux responsables de santé publique de disposer d'une estimation du nombre de cas de vMCJ attendus, dont la précision doit certes être améliorée, mais est néanmoins suffisante pour évaluer le poids épidémiologique de cette maladie dans les années à venir. Suffisante aussi pour aider à définir des mesures de réduction du risque adéquates. On peut parfois entendre des jugements très critiques sur les prédictions de l'épidémie de vMCJ, considérant que ces prédictions ont démontré surtout l'impossibilité de prédire l'épidémie. Une lecture attentive des publications ne conduit pas au même jugement. Quelques mois après la première description du vMCJ, la modélisation [12] permettait d'écarter l'hypothèse d'une épidémie de plusieurs centaines de milliers de cas au Royaume-Uni. Dès 1998, une ré-analyse des résultats publiés par Cousens *et al.* [12] aurait indiqué une épidémie de moins de 5000 cas. En 2001, les dernières esti-

mations [14, 15] situent cette épidémie aux alentours de 1000 à 2000 cas. On ne peut évidemment ignorer la souffrance que représente chaque cas, pour la personne atteinte et sa famille. Mais pour la population dans son ensemble, au Royaume-Uni comme dans les autres pays exposés à l'ESB, on peut envisager l'avenir avec beaucoup moins d'inquiétude qu'en 1996. Les modélisations, quelles que soient leurs limites, ont eu un rôle majeur dans cette évolution. ♦

SUMMARY

vCJD epidemic in the UK: anatomy and physiology of predictive models

The back-calculation technique is used in all studies aimed at predicting the future incidence of the variant Creutzfeldt-Jakob disease (vCJD) caused by the agent of bovine spongiform encephalopathy (BSE). The back-calculation method uses a simple mathematical equation which relates (1) the number of infected persons; (2) the distribution of infection times; and (3) the incubation period distribution. The different studies considered in this review differ essentially in their assumptions and in the techniques used to assign values to the unknown parameters. For the incubation period distribution, some studies use scenario based-technique analysis where the mean and the variance are arbitrarily fixed. In other recent studies, these parameters are estimated using the observed incidence of vCJD between 1995 and 2001. Despite several methodological differences, the results obtained from these different studies are consistent. Albeit the methodology used, for mean incubation periods between 15 and 20 years - the best estimates found in the « estimation » studies - the five models considered in this review predict an epidemic size between 500 and 2000 cases in the United Kingdom. ♦

RÉFÉRENCES

1. Will RG, Ironside JW, Zeidler M, *et al.* A new variant Creutzfeldt-Jakob disease in the UK. *Lancet* 1996; 347: 921-5.
2. Chazot G, Brousolle E, Lapras C, Blättler T, Aguzzi A, Kopp N. New variant of Creutzfeldt-Jakob disease in a 26-year-old French man. *Lancet* 1996; 347: 1181-8.
3. Bruce M, Will RG, Ironside JW, *et al.* Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent. *Nature* 1997; 389: 498-501.
4. Lasmézas CI, Deslys JP, Demalmay R, *et al.* BSE transmission to macaques. *Nature* 1996; 381: 743-4.
5. Hill AF, Desbruslais M, Joiner S, *et al.* The same prion strain causes vCJD and BSE. *Nature* 1997; 389: 448-50.
6. Gore SM, Bingham S, Day NE. Age related dietary exposure to meat products from dietary surveys of teenagers and adults in the 1980s and 1990s. *Br Med J* 1997; 315: 404-5.
7. Cousens S, Smith PG, Ward H, *et al.* Geographical distribution of variant Creutzfeldt-Jakob disease in Great Britain, 1994-2000. *Lancet* 2001; 357: 1002-7.

8. Bryant G, Monk P. Summary of the final report of the investigation in the North Leicestershire cluster of variant Creutzfeldt-Jakob disease. Leicestershire NHS Health Authority, 2001. <http://www.leics-ha.org.uk> (consulté le 7 octobre 2002).
9. Andrews NJ, Farrington CP, Cousens SN, *et al.* Incidence of variant Creutzfeldt-Jakob disease in the UK. *Lancet* 2000; 356: 481-2.
10. The EuroCJD Group. Genetic epidemiology of Creutzfeldt-Jakob disease in Europe. *Rev Neurol (Paris)* 2001; 157: 633-7.
11. Huillard d'Aignaux J, Costagliola D, Maccario J, *et al.* Incubation period of Creutzfeldt-Jakob disease in human growth hormone recipients in France. *Neurology* 1999; 53: 1197-201.
12. Cousens SN, Vynnycky E, Zeidler M, Will RG, Smith PG. Predicting the CJD epidemic in humans. *Nature* 1997; 385: 197-8.
13. Ghani AC, Ferguson NM, Donnelly CA, Anderson RM. Predicted vCJD mortality in Great Britain. *Nature* 2000; 406: 583-4.
14. Huillard d'Aignaux J, Cousens SN, Smith PG. Predictability of the UK variant Creutzfeldt-Jakob disease epidemic. *Science* 2001; 294: 1729-31.
15. Valleron AJ, Boelle PY, Will RG, Cesbron JY. Estimation of epidemic size and incubation time based on age characteristics of vCJD in the United Kingdom. *Science* 2001; 294: 1726-8.
16. Ferguson NM, Ghani AC, Donnelly CA, Hagensars TJ, Anderson RM. Estimating the human health risk from possible BSE infection of the British sheep flock. *Nature* 2002; 415: 420-4.
17. Cooper JD, Bird SM, De Angelis D. Prevalence of detectable abnormal prion protein in persons incubating vCJD: plausible incubation periods and cautious inference. *J Epidemiol Biostat* 2000; 5: 209-19.
18. Ghani AC, Ferguson NM, Donnelly CA, Hanegaars TJ, Anderson RM. Estimating the number of people incubating variant CJD. *Lancet* 1998; 352: 1353-4.
19. Hill AF, Butterworth RJ, Joiner S, *et al.* Investigation of variant Creutzfeldt-Jakob disease and other human prion diseases with tonsil biopsy samples. *Lancet* 1999; 353: 183-9.
20. Ironside JW, Hilton DA, Ghani AC, *et al.* Retrospective study of prion-protein accumulation in tonsil and appendix tissues. *Lancet* 2000; 355: 1693-4.
21. Brookmeyer R. Reconstruction and future trends of the AIDS epidemic in the United States. *Science* 1991; 253: 37-42.
22. Huillard d'Aignaux J, Cousens SN, Maccario J, *et al.* An analysis of the incubation period of kuru by sex and age at infection. *Epidemiology* 2002 (sous presse).
23. Lloyd SE, Onwuazor ON, Beck JA, *et al.* Identification of multiple quantitative trait loci linked to prion disease incubation period in mice. *Proc Natl Acad Sci USA* 2001; 98: 6279-83.

TIRÉS À PART

A. Alperovitch

PUB PROMOCHEM CHEZ CORLET